United States Patent Application Publication 20250263691
Kind Code A1
Publication Date August 21, 2025
Inventor(s) HAAS; Heinrich et al.

POTENCY ASSAY FOR THERAPEUTIC POTENTIAL OF CODING NUCLEIC ACID

Abstract

The invention provides potency assays for measuring, determining, identifying, quantifying, confirming, and/or validating the therapeutic potential of nucleic acid such as RNA encoding a pharmaceutically active peptide or polypeptide. The potency assays may be performed with nucleic acid such as RNA encoding various types of peptides or polypeptides, including pharmaceutically active peptides or polypeptides comprising one or more antigens or one or more epitopes. Nucleic acid such as RNA having therapeutic potential may be useful in downstream clinical applications, e.g., for eliciting an immune response against one or more antigens or one or more epitopes encoded by the nucleic acid in a subject which immune response may be therapeutic or partially or fully protective. Thus, the nucleic having therapeutic potential may be useful for vaccination.

Inventors: **HAAS; Heinrich (Mainz, DE), SCHUMACHER; Jens (Mainz, DE), TENZER; Stefan (Mainz, DE)**

Applicant: **BioNTech SE** (Mainz, DE)

Family ID: **1000008590758**

Appl. No.: **18/256510**

Filed (or PCT Filed): **September 01, 2022**

PCT No.: **PCT/EP2022/074395**

Foreign Application Priority Data

WO	PCT/EP2021/074304	Sep. 02, 2021

Publication Classification

Int. Cl.: **C12N15/10** (20060101)

U.S. Cl.:

CPC **C12N15/1082** (20130101);

Background/Summary

CROSS-REFERENCE TO RELATED APPLICATIONS [0001] This application is a National Stage Entry of International Application Number PCT/EP2022/074395, which was filed on Sep. 1, 2022 and claimed priority to International Application Number PCT/EP2021/074304, which was filed on Sep. 2, 2021. The contents of each of the aforementioned applications are incorporated herein by reference in their entireties.

SEQUENCE LISTING
[0002] The computer-readable Sequence Listing submitted on Jun. 8, 2023 and identified as follows: 28,842 bytes ST.26 XML document file named “028320-8052 Sequence Listing.xml,” created Jun. 8, 2023, is incorporated herein by reference

in its entirety.

TECHNICAL FIELD

[0003] The invention provides potency assays for measuring, determining, identifying, quantifying, confirming, and/or validating the therapeutic potential of nucleic acid (such as RNA and/or DNA) encoding a pharmaceutically active peptide or polypeptide. The potency assays may be performed with nucleic acid (such as RNA and/or DNA) encoding various types of peptides or polypeptides, including pharmaceutically active peptides or polypeptides comprising one or more antigens or one or more epitopes. Nucleic acid (such as RNA and/or DNA) having therapeutic potential may be useful in downstream clinical applications, e.g., for eliciting an immune response against one or more antigens or one or more epitopes encoded by the nucleic acid in a subject which immune response may be therapeutic or partially or fully protective. Thus, the nucleic acid having therapeutic potential may be useful for vaccination.

BACKGROUND

[0004] Apart from their well-known ability to encode biologically active proteins, nucleic acids such as DNA and RNA have other remarkable properties that make them attractive therapeutic agents. Nucleic acid-based therapeutics are easy to manufacture and relatively inexpensive.

[0005] Generally, DNA is more stable than RNA, but has some potential safety risks such as the induction of anti-DNA antibodies and the integration of the transgene into the host genome.

[0006] The use of RNA to deliver foreign genetic information into target cells offers an attractive alternative to DNA. The advantages of RNA include transient expression and non-transforming character. RNA does not require nucleus infiltration for expression and moreover cannot integrate into the host genome, thereby eliminating the risk of oncogenesis.

[0007] Potency tests are used to measure product attributes associated with product quality and manufacturing controls, and are performed to assure identity, purity, strength (potency), and stability of products used during all phases of clinical study. Similarly, potency measurements are used to demonstrate that only product lots that meet defined specifications or acceptance criteria are administered during all phases of clinical investigation and following market approval.

[0008] Thus, defining potency of biopharmaceuticals is a central figure during product development and thereafter.

[0009] Potency assays involve the quantitative measure of certain criteria that should describe the ability of a product to achieve a defined biological effect. The criteria measured should be closely related to the product's intended biological effect and ideally, it should be related to the product's clinical purpose. Measurement of the potency of a product is not the same as measuring clinical efficacy. Rather, it is a means to control product quality and provide appropriate release criteria, in particular under GMP. Normally, for each and every product which is to be administered to a subject, a separate potency assay must be developed. In the rapidly evolving nucleic acid world, with potential hundreds of different constructs and where mostly no antibodies are available for detection, a potency assay which can be easily adapted to a new product would be of great benefit.

[0010] There is a need of providing potency assays to measure, determine, identify, quantify, confirm, and/or validate the therapeutic potential of coding nucleic acid (such as RNA and/or DNA), in particular the therapeutic potential of particulate formulations comprising said coding nucleic acid, e.g., a nucleic acid drug product, and related uses thereof.

SUMMARY

[0011] It has been observed according to the invention that the ability of nucleic acid (such as RNA and/or DNA) to express an encoded pharmaceutically active peptide or polypeptide in a cellular system in vitro is associated with therapeutic potential in vivo. Based on this observation a rapid, cost-effective, reliable, and easy to use and interpret potency assay to measure, determine, identify, quantify, confirm and/or validate the therapeutic potential of nucleic acid (such as RNA and/or DNA) encoding a pharmaceutically active peptide or polypeptide is provided. The potency assay provided herein can be easily adapted to a new product.

[0012] Specifically, an LC-MS potency assay for nucleic acid products, e.g., RNA-lipoplex (RNA-LPX) or DNA-LPX or RNA-lipid nanoparticle (RNA-LNP) products, is described herein which is based on testing the key steps of the product-specific mechanism of action (MoA) that are directly coupled to the biological activity, e.g., activation of T cells. The readout, namely the measure of translation of the delivered nucleic acid (such as RNA and/or DNA) into peptide, is directly indicative of the product quality of the nucleic acid product, e.g., RNA-LPX, DNA-LPX, or RNA-LNP drug product, and is stability indicating. The assay allows insight into the potency, covering the successful uptake of the nucleic acid (such RNA and/or DNA) and translation into the respective peptides, e.g., antigens, and is predictive for the biological activity, e.g., T cell activation.

[0013] For measuring cellular uptake and translation, cells from an animal cell line can be used, in particular those which take up nucleic acid products, e.g., RNA-LPX, DNA-LPX, or RNA-LNP, using the same mechanism as cells of a recipient (i.e., the target cells of the recipient which are to take up the nucleic acid products, such as dendritic cells (DCs)) and, optionally, which are suitable for routine testing in a QC-environment (such as GMP QC-environment) for batch release. In some embodiments, the cells may be selected from the group consisting of Chinese hamster ovary (CHO), K562, HepG2, HEK293T, RAW, and C2C12 cells. In some embodiments, CHO cells were chosen given that these cells take up nucleic acid products, e.g., RNA-LPX or RNA-LNP, using the same mechanism as dendritic cells (DCs) and based on their suitability for routine testing in a QC-environment (such as in a GMP QC-environment) for batch release. Upon translation, CHO cells are able to present antigenic epitopes to cognate T cells, resulting in their stimulation. Thus, this assay is reflective of the biological function of the drug product.

[0014] In order to quantify the translated peptide, a unique MITD (MHC class I trafficking domain) peptide common to the C-terminal end of the nucleic acid-encoded (such as RNA-encoded) sequence may be quantified from the total cell lysate

using LC-MS/MS analysis. In some embodiments, however, also another peptide of the nucleic acid-encoded (such as RNA-encoded) sequence can be used for the quantification of the expressed peptide. In some embodiments, the another peptide can be quantified using LC-MS/MS. In some embodiments, the another peptide is or comprises a reporter peptide or polypeptide, such as peptide or polypeptide producing bioluminescence (like Green Fluorescent Protein (GFP) or luciferase).

[0015] In various embodiments, a potency assay for the therapeutic potential of nucleic acid (such as RNA and/or DNA) is contemplated. The potency assay is clinically important because it can be used to rapidly and reliably validate a clinical nucleic acid-based such RNA-based therapy product prior to administration to a subject. In effect, the framework for the potency assays for therapeutic potential contemplated herein is likely to become the “gold-standard” validation assay for therapeutic nucleic acid such as RNA products.

[0016] In some aspects, the invention provides a method for analyzing nucleic acid encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprising the following steps:

[0017] (i) providing the nucleic acid; [0018] (ii) transfecting cells in vitro with the nucleic acid; [0019] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; and [0020] (iv) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof.

[0021] In some embodiments, the method further comprises the following step: [0022] (v) using the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof as an indication for the potency of the nucleic acid to induce the biological activity in a biological system.

[0023] In some embodiments, the nucleic acid is DNA (e.g., one or more DNAs), RNA (e.g., one or more RNAs), or a mixture of DNA and RNA (e.g., one or more DNAs and one or more RNAs).

[0024] In some aspects, the invention provides a method for analyzing DNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprising the following steps: [0025] (i) providing the DNA; [0026] (ii) transfecting cells in vitro with the DNA; [0027] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; and [0028] (iv) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof.

[0029] In some embodiments, the method further comprises the following step: [0030] (v) using the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof as an indication for the potency of the DNA to induce the biological activity in a biological system.

[0031] In some embodiments, the DNA is present in the form of a vector, e.g., a vector comprising DNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity. In some embodiments, the vector is a DNA vector.

[0032] In some aspects, the invention provides a method for analyzing a mixture of DNA and RNA each encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprising the following steps: [0033] (i) providing the mixture of DNA and RNA; [0034] (ii) transfecting cells in vitro with the mixture of DNA and RNA; [0035] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; and [0036] (iv) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof.

[0037] In some embodiments, the method further comprises the following step: [0038] (v) using the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof as an indication for the potency of the mixture of DNA and RNA to induce the biological activity in a biological system.

[0039] In some embodiments, the DNA is present in the form of a vector, e.g., a vector comprising DNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity. In some embodiments, the vector is a DNA vector.

[0040] In some aspects, the invention provides a method for analyzing RNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprising the following steps: [0041] (i) providing the RNA; [0042] (ii) transfecting cells in vitro with the RNA; [0043] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; and [0044] (iv) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof.

[0045] In some embodiments, the method further comprises the following step: [0046] (v) using the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof as an indication for the potency of the RNA to induce the biological activity in a biological system.

[0047] In some embodiments, the cells mimic nucleic acid uptake mechanisms (such as RNA and/or DNA uptake mechanisms) of biological systems.

[0048] In some embodiments, the biological system is present in a human patient.

[0049] In some embodiments, the biological system comprises antigen presenting cells, preferably dendritic cells.

[0050] In some embodiments, the dendritic cells comprise immature dendritic cells.

[0051] In some embodiments, the cells are characterized by a macropinocytosis-mediated RNA uptake mechanism. In these embodiments, it is preferred that the nucleic acid (such as RNA and/or DNA) is formulated as lipoplex particles.

[0052] In some embodiments, the cells are characterized by an endosomal-mediated RNA uptake mechanism. In these embodiments, it is preferred that the nucleic acid (such as RNA and/or DNA) is formulated as lipid nanoparticles.
[0053] In some embodiments, the cells are cells from an animal cell line, in particular those which take up nucleic acid products, e.g., RNA-LPX, DNA-LPX, or RNA-LNP, using the same mechanism as cells of a recipient (i.e., the target cells of the recipient which are to take up the nucleic acid products, such as dendritic cells (DCs)) and which are suitable for routine testing in a QC-environment (such as GMP QC-environment).
[0054] In some embodiments, the cells are Chinese hamster ovary (CHO) cells. In some embodiments, the cells are selected from K562, HepG2, HEK293T, RAW, and C2C12 cells, such as from K562, HEK293T, RAW, and C2C12 cells.
[0055] In some embodiments, the fragment of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity is specific for the expressed amino acid sequence.
[0056] In some embodiments, the fragment of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity is not comprised by the amino acid sequence of a peptide or polypeptide having biological activity.
[0057] In some embodiments, the fragment of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprises an amino acid sequence enhancing antigen processing and/or presentation or a fragment thereof.
[0058] In some embodiments, the amino acid sequence enhancing antigen processing and/or presentation comprises an amino acid sequence corresponding to the transmembrane and cytoplasmic domain of a MHC molecule, preferably a MHC class I molecule.
[0059] In some embodiments, the amino acid sequence enhancing antigen processing and/or presentation comprises the amino acid sequence of SEQ ID NO: 2, or 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: 2.
[0060] In some embodiments, the fragment of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprises an amino acid sequence which breaks immunological tolerance or a fragment thereof.
[0061] In some embodiments, the amino acid sequence which breaks immunological tolerance comprises helper epitopes, preferably tetanus toxoid-derived helper epitopes.
[0062] In some embodiments, the amino acid sequence which breaks immunological tolerance comprises the amino acid sequence of SEQ ID NO: 3, or 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.
[0063] In some embodiments, the fragment of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprises an amino acid sequence producing bioluminescence.
[0064] In some embodiments, the amino acid sequence producing bioluminescence produces fluorescence.
[0065] In some the amino acid sequence producing bioluminescence is selected from the group consisting of Green Fluorescent Protein (GFP), Yellow Fluorescent Protein (YFP), Red Fluorescent Protein (RFP), Blue Fluorescent Protein (EBFP), Cyan Fluorescent Protein (ECFP), their variants (such as enhanced GFP (EGFP), Superfolder GFP (sfGFP), and luciferase.
[0066] In some embodiments, the method described herein comprises lysing the cells prior to determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof.
[0067] In some embodiments, the method described herein further comprises processing the cell lysate prior to determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof.
[0068] In some embodiments, processing the cell lysate comprises one or more selected from denaturation, reduction, protease digestion (such as digestion using trypsin, Glu-C, LysN, Lys-C, Asp-N chymotrypsin, or a mixture of any two or more of these proteases), alkylation, drying, reconstitution, and desalting, such as from tryptic digestion, alkylation and desalting.
[0069] In some embodiments, the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is determined using mass spectroscopy.
[0070] In some embodiments, the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is determined using liquid chromatography-mass spectrometry (LC-MS).
[0071] In some embodiments, the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is determined using targeted LC-MS.
[0072] In some embodiments, the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is determined using one or more amino acid sequences expressed by the cells as reference for quantification.
[0073] In some embodiments, the one or more amino acid sequences expressed by the cells comprise one or more amino acid sequences of housekeeping proteins.
[0074] In some embodiments, the potency of the nucleic acid (such as RNA and/or DNA) to induce the biological activity in a biological system comprises the therapeutic potency of the nucleic acid (such as RNA and/or DNA).
[0075] In some embodiments, the nucleic acid (such as RNA and/or DNA) has sufficient potency to induce the biological

activity in a biological system such as therapeutic potency if the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is above a pre-determined cut-off.
[0076] In some embodiments, the nucleic acid (such as RNA and/or DNA) does not have sufficient potency to induce the biological activity in a biological system such as therapeutic potency if the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is below a pre-determined cut-off.
[0077] In some embodiments, the pre-determined cut-off is determined using nucleic acid (such as RNA and/or DNA) known to have acceptable potency to induce the biological activity in a biological system such as therapeutic potency.
[0078] In some embodiments, the nucleic acid (such as RNA and/or DNA) used to determine the pre-determined cut-off and the nucleic acid (such as RNA and/or DNA) to be analyzed have the same chemical composition.
[0079] In some embodiments, the method described herein is for analyzing different batches of the same nucleic acid (such as RNA and/or DNA).
[0080] In some embodiments, nucleic acid (such as RNA and/or DNA) or nucleic acid batches (such as RNA and/or DNA batches) having sufficient potency to induce the biological activity in a biological system such as therapeutic potency are used or are to be used for therapy and/or nucleic acid (such as RNA and/or DNA) or nucleic acid batches (such as RNA and/or DNA batches) not having sufficient potency to induce the biological activity in a biological system such as therapeutic potency are not used or are not to be used for therapy.
[0081] In some embodiments, the potency of the nucleic acid (such as RNA and/or DNA) to induce the biological activity in a biological system such as therapeutic potency of the nucleic acid (such as RNA and/or DNA) reflects the quality such as the therapeutic quality of the nucleic acid (such as RNA and/or DNA).
[0082] In some embodiments, the quality of the nucleic acid (such as RNA and/or DNA) reflects whether and/or to what extent the nucleic acid (such as RNA and/or DNA) was exposed to detrimental conditions.
[0083] In some embodiments, the detrimental conditions comprise heat.
[0084] Using a living, cellular system instead of a cell-free system (such as a reticulocyte lysate) may provide the advantage that the potency assay provided herein is capable of indicating, whether the potency of the nucleic acid (such as RNA and/or DNA) to induce the biological activity in a biological system (such as therapeutic potency of the nucleic acid) reflects one or more parameters of the nucleic acid (or of the formulation/composition comprising the nucleic acid, such as RNA-LPX etc.), whereas a potency assay based on a cell-free system is not capable of providing such an indication.
[0085] Thus, in some embodiments, the potency of the nucleic acid (such as RNA and/or DNA) to induce the biological activity in a biological system such as therapeutic potency of the nucleic acid (such as RNA and/or DNA) reflects one or more parameters selected from the group consisting of particle parameters, formulation/composition parameters, and nucleic acid (such as RNA and/or DNA) parameters.
[0086] In some embodiments, particle parameters include size, surface charge, lipid quality (e.g., degradation), particle structure (e.g., lamellarity), and intracellular nucleic acid (such as RNA and/or DNA) release, N/P ratio, concentration of free nucleic acid (such as RNA and/or DNA), and concentration of accessible nucleic acid (such as RNA and/or DNA).
[0087] In some embodiments, formulation/composition parameters include osmolality, pH, quality of formulation/composition components other than the nucleic acid (e.g., quality of buffer components), and concentration of endotoxin.
[0088] In some embodiments, nucleic acid (such as RNA and/or DNA) parameters include concentration, sequence correctness (e.g., frameshift or premature stop), integrity, and concentration of endotoxin.
[0089] In some embodiments, RNA parameters include RNA concentration, sequence correctness (e.g., frameshift or premature stop), RNA integrity, concentration of endotoxin, capping, polyA sequence, concentration of dsRNA, and UTR (5′ and/or 3′). In particular, it is noted in this respect that the potency assay provided herein is capable of indicating, whether the potency of the RNA to induce the biological activity in a biological system (such as therapeutic potency of the RNA) reflects capping (because only capped RNA is translated into the encoded peptide or polypeptide), whereas a potency assay based on a cell-free system is not capable of providing such an indication (because in such a cell-free system also uncapped RNA is translated into the encoded peptide).
[0090] In some embodiments, the potency of the nucleic acid (such as RNA and/or DNA) to induce the biological activity in a biological system such as therapeutic potency of the nucleic acid (such as RNA and/or DNA) reflects the integrity of the nucleic acid (such as RNA and/or DNA).
[0091] In some embodiments, the nucleic acid is RNA and the potency of the RNAto induce the biological activity in a biological system such as therapeutic potency of the RNA reflects the capping of the RNA.
[0092] In some embodiments, the method described herein is for analyzing the potency of the nucleic acid (such as RNA and/or DNA) to induce the biological activity in a biological system.
[0093] In some embodiments, the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is indicative for the potency of the nucleic acid (such as RNA and/or DNA) to induce the biological activity in a biological system.
[0094] In some embodiments, the method described herein is for analyzing whether the quality and/or quantity of the nucleic acid (such as RNA and/or DNA) is sufficient to induce the biological activity in a biological system.
[0095] In some embodiments, the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is indicative for whether the quality and/or quantity of the

nucleic acid (such as RNA and/or DNA) is sufficient to induce the biological activity in a biological system.
[0096] In some embodiments, the biological activity comprises an ability to elicit a specific response in a disease-relevant system.
[0097] In some embodiments, the specific response comprises or is an immune response.
[0098] In some embodiments, the immune response comprises a T cell response.
[0099] In some embodiments, the peptide or polypeptide having biological activity is selected from the group consisting of vaccines (e.g., antigens, epitopes), proteins for replacement therapy, antibodies, antibody-like molecules, and cytokines.
[0100] In some embodiments, the nucleic acid is RNA.
[0101] In some embodiments, the RNA is single stranded RNA.
[0102] In some embodiments, the RNA is mRNA.
[0103] In some embodiments, the RNA is generated by RNA in vitro transcription.
[0104] In some embodiments, the RNA comprises a 5′ cap structure.
[0105] In some embodiments, the RNA does not comprise modified ribonucleotides.
[0106] In some embodiments, the RNA comprises modified ribonucleotides. In some embodiments, the modified ribonucleotides comprise modified uridines. In some embodiments, the modified uridines comprise N1-methyl-pseudouridine.
[0107] In some embodiments, the nucleic acid is DNA.
[0108] In some embodiments, the DNA is present in the form of a vector.
[0109] In some embodiments, the vector comprises DNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity.
[0110] In some embodiments, the vector is a DNA vector.
[0111] In some embodiments, the nucleic acid is a mixture of RNA and DNA.
[0112] In some embodiments, the RNA in the mixture is single stranded RNA.
[0113] In some embodiments, the RNA in the mixture is mRNA.
[0114] In some embodiments, the RNA in the mixture is generated by RNA in vitro transcription.
[0115] In some embodiments, the RNA in the mixture comprises a 5′ cap structure.
[0116] In some embodiments, the RNA in the mixture does not comprise modified ribonucleotides.
[0117] In some embodiments, the RNA in the mixture comprises modified ribonucleotides. In some embodiments, the modified ribonucleotides comprise modified uridines. In some embodiments, the modified uridines comprise N1-methyl-pseudouridine.
[0118] In some embodiments, the DNA in the mixture is present in the form of a vector.
[0119] In some embodiments, the vector in the mixture comprises DNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity.
[0120] In some embodiments, the vector in the mixture is a DNA vector.
[0121] In some embodiments, the nucleic acid (such as RNA and/or DNA) is formulated with a delivery vehicle.
[0122] In some embodiments, the nucleic acid (such as RNA and/or DNA) is formulated with one or more compounds complexing the nucleic acid (such as RNA and/or DNA).
[0123] In some embodiments, the nucleic acid (such as RNA and/or DNA) is formulated as particles.
[0124] In some embodiments, the nucleic acid (such as RNA and/or DNA) is formulated as lipoplex particles. In these embodiments, it is preferred that the cells are characterized by a macropinocytosis-mediated RNA uptake mechanism.
[0125] In some embodiments, the nucleic acid (such as RNA and/or DNA) is formulated as lipid nanoparticles.
[0126] In some embodiments, the peptide or polypeptide having biological activity is a vaccine.
[0127] In some embodiments, the vaccine is a T cell vaccine.
[0128] In some embodiments, the nucleic acid (such as RNA and/or DNA) comprises a mixture of different nucleic acids (such as RNAs and/or DNAs, e.g., two or more RNAs, two or more DNAs, or one or more RNAs and one or more DNAs), wherein each nucleic acid (such as RNA and/or DNA) encodes an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity.
[0129] In some embodiments, the mixture of different nucleic acids (such as RNAs and/or DNAs, e.g., two or more RNAs, two or more DNAs, or one or more RNAs and one or more DNAs) comprises nucleic acids (such as RNAs and/or DNAs, e.g., two or more RNAs, two or more DNAs, or one or more RNAs and one or more DNAs) encoding different amino acid sequences comprising the amino acid sequence of a peptide or polypeptide having biological activity.
[0130] In some embodiments, the different amino acid sequences comprise the amino acid sequence of different peptides or polypeptides having biological activity.
[0131] In some embodiments, the different peptides or polypeptides having biological activity comprise different antigens.
[0132] In some embodiments, the nucleic acid (such as RNA and/or DNA) comprises a mixture of different nucleic acids (such as RNAs and/or DNAs, e.g., two or more RNAs, two or more DNAs, or one or more RNAs and one or more DNAs) encoding amino acid sequences comprising the amino acid sequence of different antigens.
[0133] In some aspects, the invention provides a method for analyzing the potency of nucleic acid encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity to induce the biological activity in a biological system comprising the following steps: [0134] (i) providing the nucleic acid; [0135] (ii) transfecting cells of an animal cell line, such as Chinese hamster ovary (CHO) cells, in vitro with the nucleic acid; [0136] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide

having biological activity; [0137] (iv) lysing the cells; [0138] (v) processing the cell lysate; and [0139] (vi) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof using mass spectroscopy, wherein one or more amino acid sequences expressed by the cells of the animal cell line, such as the CHO cells, are used as reference for quantification.
[0140] Embodiments of this method are as described herein.
[0141] In some aspects, the invention provides a method for analyzing the potency of DNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity to induce the biological activity in a biological system comprising the following steps: [0142] (i) providing the DNA; [0143] (ii) transfecting cells of an animal cell line, such as Chinese hamster ovary (CHO) cells, in vitro with the DNA; [0144] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; [0145] (iv) lysing the cells; [0146] (v) processing the cell lysate; and [0147] (vi) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof using mass spectroscopy, wherein one or more amino acid sequences expressed by the cells of the animal cell line, such as the CHO cells, are used as reference for quantification.
[0148] Embodiments of this method are as described herein.
[0149] In some aspects, the invention provides a method for analyzing the potency of a mixture of RNA and DNA each encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity to induce the biological activity in a biological system comprising the following steps: [0150] (i) providing the mixture of RNA and DNA; [0151] (ii) transfecting cells of an animal cell line, such as Chinese hamster ovary (CHO) cells, in vitro with the mixture of RNA and DNA; [0152] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; [0153] (iv) lysing the cells; [0154] (v) processing the cell lysate; and [0155] (vi) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof using mass spectroscopy, wherein one or more amino acid sequences expressed by the cells of the animal cell line, such as the CHO cells, are used as reference for quantification.
[0156] Embodiments of this method are as described herein.
[0157] In some aspects, the invention provides a method for analyzing the potency of RNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity to induce the biological activity in a biological system comprising the following steps: [0158] (i) providing the RNA; [0159] (ii) transfecting cells of an animal cell line, such as Chinese hamster ovary (CHO) cells, in vitro with the RNA; [0160] (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; [0161] (iv) lysing the cells; [0162] (v) processing the cell lysate; and [0163] (vi) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof using mass spectroscopy, wherein one or more amino acid sequences expressed by the cells of the animal cell line, such as the CHO cells, are used as reference for quantification.
[0164] Embodiments of this method are as described herein.
[0165] In some embodiments, the RNA described herein is single-stranded RNA that may be translated into the respective protein upon entering cells, e.g., cells used in the assays described herein and cells of a recipient. In addition to wildtype or codon-optimized sequences encoding the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity, e.g., a pharmaceutically active peptide or polypeptide such as 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.sub.2.sup.7,2′-OGppSpG) or m.sub.2.sup.7,3′-OGppp(m.sub.1.sup.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 alpha-globin mRNA, optionally with an optimized ‘Kozak sequence’ to increase translational efficiency 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) 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 may be used. This poly(A)-tail sequence was designed to enhance RNA stability and translational efficiency.
[0166] The amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity, e.g., a pharmaceutically active peptide or polypeptide such as antigen sequence, may comprise amino acid sequences other than the amino acid sequence of a peptide or polypeptide having biological activity. Such other amino acid sequences may support the function or activity of the peptide or polypeptide having biological activity. In some embodiments, such other amino acid sequences comprise an amino acid sequence enhancing antigen processing and/or presentation. Alternatively, or additionally, such other amino acid sequences comprise an amino acid sequence which breaks immunological tolerance. Alternatively, or additionally, such other amino acid sequences comprise an amino acid sequence which produces bioluminescence. Such other amino acid sequences may be useful for determining the amount of the amino acid sequence

comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof in the assays described herein. In particular, such other amino acid sequences may be useful for quantification by LC-MS/MS analysis.

[0167] The nucleic acids (such as RNA and/or DNA) described herein may be complexed with polymers, proteins and/or lipids, preferably lipids, to generate nucleic acid-particles for administration. If a combination of different nucleic acids is used, the nucleic acids may be complexed together or complexed separately.

Description

BRIEF DESCRIPTION OF THE FIGURES

[0168] FIG. **1**: Macropinocytosis-driven dose-dependent uptake of Cy5-labeled RNA-LPX in CHO cells. CHO cells were or were not pretreated with the selective macropinocytosis-inhibitor rottlerin and incubated with 0.2, 1.5 or 3 µg Cy5-labeled Luciferase-RNA-LPX. After washing cells were fixed and Cy5-signals were analyzed by fluorescence microscopy (left panel show representative images). Scale bar: 50 µm. (right) Quantification of Cy5-positive cells (n=>20.000 cells per dose and per treatment were analyzed). Error bars represent SEM. Abbreviations: CHO=Chinese hamster ovary cells; LPX=RNA-lipoplex; SEM=standard error of the mean.

[0169] FIG. **2**: Localization of RNA-LPX-encoded antigen in DCs or CHO cells. The cells were incubated with tumor antigen MAGE-A3-RNA-LPX (for construct see FIG. **12**) for 24 h, subsequently stained under native conditions with a MAGE A3-specific antibody (green) and analyzed by structured illumination microscopy. Small pictures represent orthogonal views. In C) the cell membranes were stained in parallel (red). The graph shows a line scan (white dotted bar) from the merge image. Bars represent 20 µm (A/B) or 25 µm (C). Abbreviations: CHO=Chinese hamster ovary; DC=dendritic cells; MAGE A3=melanoma-associated antigen 3.

[0170] FIG. **3**: Dose Response Curve for eGFP-LPX-lipofected DCs (left) and CHO cells (right). Cells were co-incubated with titrated amounts of RNA-LPX encoding the reporter eGFP. The eGFP fluorescence signals were analyzed and the average relative fluorescence units (RFU) were calculated. Error bars represent the standard deviation (n=3). Abbreviations: CHO=Chinese hamster ovary; DC=dendritic cells; eGFP=enhanced green fluorescent protein; LPX=lipoplex.

[0171] FIG. **4**: Dose Response Curve for DCs (left) or CHO Cells (right) After Lipofection with Unstressed (Control) or Stressed eGFP-RNA-LPX. The RNA-LPX were exposed to thermal stress before they were used for lipofection. Afterwards, the cells were co-incubated with titrated amount of the stressed or unstressed RNA-LPX encoding the reporter eGFP (eGFP-LPX). The eGFP fluorescence signals were analyzed and the average RFU is calculated. Error bars represent the standard deviation (n=3). Abbreviations: CHO=Chinese hamster ovary; DC=dendritic cells; eGFP=enhanced green fluorescent protein; RFU=relative fluorescence unit; RNA LPX=ribonucleic acid lipoplex.

[0172] FIG. **5**: Influence of DP Variation on the Potency Assay. A) Four heat-degraded RNAs resulting in different RNA-integrities (95%; 80%; 74% and 40%) were used for the manufacturing of the RNA-LPX and B) RNAs with (capped, black line) or without 5′cap (uncapped, blue line) were used for the manufacturing of the RNA-LPX. On the left side the fluorescence signal is plotted as a function of each drug product dose. The different colors represent the measurements with the respective DP manufactured with the indicates heat-degraded RNAs. On the right side the corresponding AUC-values were plotted. C) In vitro translation of capped (left) and uncapped RNA (right) using rabbit reticulocyte lysate to verify functionality of the used RNAs. Error bars represent the standard deviation (n=3).

[0173] FIG. **6**: Proof-of-concept. A) Triplicate measurement of cells after 24 h of lipofection with RNA-lipoplexes, encoding for eGFP. Three different RNA-lipoplex doses were used for transfection. B) Amino acid sequence of the eGFP-construct. In red the specific peptides which were identified in the cell lysates by MS/MS (n=3). C) Correlation of the eGFP fluorescence signal (in blue) in the lysates in comparison to the eGFP MS peptide-signal intensities (in red).

[0174] FIG. **7**: Proof-of-concept with representative iNeST RNA-LPX. Normalized LC-MS (MS/MS) MITD-peptide quantification of cells after 24 h of lipofection with iNeST design space RNA-lipoplexes. Error bars represent the standard deviation (n=6).

[0175] FIG. **8**: Dose Response Curve for iNeST-lipoplexed-lipofected CHO Cells. CHO cells were co-incubated with titrated amounts of RNA-lipoplex encoding for different neoantigens (iNeST5 or 6) and the normalized MITD-peptide quantification (MS/MS) were shown for iNeST construct 5 (left) and 6 (right). Abbreviations: CHO=Chinese hamster ovary; LPX=lipoplex.

[0176] FIG. **9**: Impact of Accelerated Temperature Stress on LC-MS Potency. All six iNeST RNA-LPX (iNeST1 to 6) were heat-stressed (2 days, 40° C.) and the resulting samples were applied at three different dose levels to cells and analyzed by the LC-MS-potency assay. A dose-response curve was plotted against the normalized MITD peptide quantity for the non-stressed samples (blue) and the stressed samples (orange) in parallel.

[0177] FIG. **10**: Influence of Two Time Points of the Accelerated Stress Condition. All six iNeST RNA-LPX (iNeST1 to 6) were heat-stressed for two (left graph) or ten days (right graph) at 40° C. The RNA integrity of the non-stressed control samples (gray bars) and the stressed samples (orange bars) are analyzed (dotted yellow line marked the RNA integrity specification limit). In parallel all samples were applied to cells and analyzed by the LC-MS-potency approach. The respective MITD-peptide quantity is normalized against the signal from the non-stressed MITD-peptide quantity (blue bars).

[0178] FIG. **11**: Measurement of Different Readouts Show a RNA-LPX Dose Response Curve Relation. A) CHO cells were lipofected with a tumor antigen MAGE A3 RNA-LPX in a dose-dependent manner. The cells were analyzed for MAGE A3 translation by the intended LC-MS potency approach (LC-MS, orange dots). The cells were analyzed in parallel for the surface-localization (see FIG. **2**) by native antibody-staining and quantitative HT-microscopy (AB, blue dots). The cells were also analyzed for their ability to activate T-cells by a specialized Jurkat NFAT assay (green dots). CHO cells were electroporated with HLA-A*0101 RNA, transfected with titrated amounts of MAGE-A3-encoding RNA-LPX and evaluated for their capability to activate Jurkat T cells expressing an HLA-A*0101-restricted MAGE-A3-TCR.Cognate activation of MAGE-A3-TCR-transfected Jurkat cells based on NFAT-driven expression of luciferase as reporter. B) Same experiment as in A, the CHO cells were lipofected with RNA-LPX encoding the antigen (E7). The CHO cells were electroporated with HLA-DQBA1*0102 and DQB1*0501 RNA, transfected with titrated amounts of HPV-E7 encoding RNA-LPX and evaluated for their capability to activate Jurkat T cells expressing an HLA-DQBA1*0102/DQB1*0501-restricted HPV-E7-TCR. Transfected CHO cells were subjected to targeted LC-MS assay to measure levels of translated HPV-E7 protein and in parallel for surface-localization via HT-microscopy. Cognate activation of HPV E7 TCR-transfected Jurkat cells based on NFAT-driven expression of luciferase as reporter. Error bars represent standard deviation (n=3).
[0179] FIG. **12**: General structure of RNAs coding for BNT111 (NY-ESO-1, tyrosinase, MAGE-A3, and TPTE) and BNT113 (HPV-E6 and HPV-E7) tumor antigens. Schematic illustration of RNAs' general structure with 5′-cap, 5′- and 3′-untranslated regions, coding sequences with intrinsic secretory signal peptide (if applicable), and poly(A)-tail. Please note that the individual elements are not drawn exactly true to scale compared to their respective sequence lengths. ORF=open reading frame; UTR=Untranslated region; sec=secretory signal peptide.
[0180] FIG. **13**: Schematic illustration of the general structure of the iNeST RNA drug substances with constant 5′-cap [beta-S-ARCA (D1)), 5′- and 3′-UTRs (hAg-Kozak and FI, respectively], N- and C-terminal fusion tags (sec2.0 and MITD, respectively), and poly(A) tail (A120), as well as patient-specific sequences encoding the neoepitopes (neo1 to 10) fused by GS-rich linkers. Abbreviations: GS—glycine and serine; MITD—major histocompatibility complex class I trafficking domain; sec—secretory signal peptide; UTR—untranslated region.
[0181] FIG. **14**: Overview of the Analyzed Tumor Antigen Constructs (#1:MAGE-A3, 2 #:Tyrosinase and 3 #NY-ESO, see also FIG. **12**). The C-terminal constant part of the constructs is highlighted (peptide #1; peptide #2 and MITD).
[0182] FIG. **15**: Results of the MS-Based Relative Quantification of Three (A-C) tumor antigens. For each construct six doses were applied in duplicates (n=2). Error bars represent the standard deviation.
[0183] FIG. **16**: Cells were lipofected with DP (RNA-LPX) encoding for MAGE A3 (FIG. **16**A) or TPTE (FIG. **16**B) in a dose-dependent manner. Analytical readouts (LC-MS; Jurkat NFAT (with HepG2 or CHO cells) and quantitative immunofluorescence microscopy (IF)) were done in parallel and the results are linearly fitted and compared. Each DP which was accelerated heat stressed at two conditions (3 days (squares) and 10 days (circles)) and the non-stressed DP (triangles) was applied on the cells in a dose-dependent manner. Error bars represent standard deviation.
[0184] FIG. **17**: Evaluation and comparison of the different analytical results (LC-MS vs Jurkat and RNA integrity (CE) vs LC-MS) of the non-stressed (right value), 3 days (middle value) and 10 days (left value) stressed DP. The slopes of the linear fits for the different analytical measurements were plotted on a graph and fitted by a linear regression.
[0185] FIG. **18**: Evaluation and comparison of the LC-MS (right bars in the diagrams) and Jurkat assay (left bars in the diagrams) results. The data points of the dose-response curve were analyzed with the relative potency GMP-software PLA. The different stressed samples are compared to the non-stressed samples (set as 100% potency) and the respective potency is calculated using the PLA software.
[0186] FIG. **19**: Cells were lipofected with DP (RNA-LPX) encoding for TYR (FIG. **19**A) or NY-ESO (FIG. **19**B) in a dose-dependent manner. Analytical readouts (LC-MS; Jurkat NFAT (with HepG2)) were done in parallel and the results are linearly fitted and compared. Each DP which was accelerated heat stressed at two conditions (3 days (squares) and 10 days (circles)) and in addition the non-stressed DP (triangles) was applied on the cells in a dose-dependent manner. Error bars represent standard deviation.
[0187] FIG. **20**: Evaluation and comparison of the different analytical results (LC-MS vs Jurkat and RNA integrity (CE) vs LC-MS) of the non-stressed (right value), 3 days (middle value) and 10 days (left value) stressed DP. The slopes of the linear fits for the different analytical measurements were plotted on a graph and fitted by a linear regression.
[0188] FIG. **21**: Evaluation and comparison of the LC-MS (right bars in the diagrams) and Jurkat assay (left bars in the diagrams) results. The data points of the dose-response curve were analyzed with the relative potency GMP-software PLA. The different stressed samples are compared to the non-stressed samples (set as 100% potency) and the respective potency is calculated using the PLA software
[0189] FIG. **22**: Cells were lipofected with DP (RNA-LPX) encoding for E6 (FIG. **22**A) or E7 (FIG. **22**B) in a dose-dependent manner. Analytical readouts (LC-MS; Jurkat NFAT (with HepG2 or CHO) and quantitative immunofluorescence microscopy (IF)) were done in parallel and the results are linearly fitted and compared. Each DP which was accelerated heat stressed at two conditions (3 days (squares) and 10 days (circles)) and in addition the non-stressed DP (triangles) was applied on the cells in a dose-dependent manner. Error bars represent standard deviation.
[0190] FIG. **23**: Evaluation and comparison of the different analytical results (LC-MS vs Jurkat and RNA integrity (CE) vs LC-MS) of the non-stressed (right value), 3 days (middle value) and 10 days (left value) stressed DP. The slopes of the linear fits for the different analytical measurements were plotted on a graph and fitted by a linear regression.
[0191] FIG. **24**: Evaluation and comparison of the LC-MS (right bars in the diagrams) and Jurkat assay (left bars in the diagrams) results. The data points of the dose-response curve were analyzed with the relative potency GMP-software PLA.

The different stressed samples are compared to the non-stressed samples (set as 100% potency) and the respective potency is calculated using the PLA software.

[0192] FIG. **25**: CHO cells were transfected with ATM or CTM DP (LNP). The DP contains RNA encoding T cell string fusion constructs. The DP's were applied in a dose-dependent manner (0.75 to 9 µg, also untranslated cells (UT) were analyzed as control) on the cells. After RNA uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. Two T cell specific peptides were analyzed by LC-MS/MS (PRM) and the normalized expression values were plotted.

[0193] FIG. **26**: CHO cells were transfected with DP ATM (LNP) with uRNA or mod RNA. Each DP contains 4 RNAs encoding for 4 fusion proteins composed of 8 TB antigens (Ag85A+Hrp1; ESAT6+RpfD; M72+VapB47; RpfA+HbhA). The DP was applied in a dose-dependent manner (150 to 2400 ng) on the cells. After uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. TB-antigen specific peptides were analyzed by PRM. Error bars represent standard deviation.

[0194] FIG. **27**: CHO cells were transfected with LNP control containing RNA encoding for the reporter luciferase, the luciferase being a secreted version. The DP was applied in one dose in triplicates on the cells. After uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. Luciferase specific peptides were analyzed by PRM. Error bars represent standard deviation.

[0195] FIG. **28**: Duplicate fluorescence measurement of CHO cells after 24 h of lipofection with DNA-lipofectamine 2000 (Lipofectamine 2000 (Fisher Scientific GmbH, 11668030), LOT: 2418953), the DNA encoding for eGFP. Four different DNA-lipofectamine doses (225 to 1800 ng) were applied onto the cells. After RNA uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. The eGFP fluorescence signal (fluorescence intensity) of the cells was correlated in comparison to the eGFP MS/MS peptide-signal intensities (MS signal) derived from two different GFP-specific-peptides (Top1 and Top2).

[0196] FIG. **29**: Cells were lipofected with DP (RNA-LPX) encoding for five different tumor antigens in a dose-dependent manner (eleven different dosages per antigen). After 24 h incubation time the cells were harvested and subjected to LC-MS analysis. The Peptide 3 (of the p2p16 domain) is quantified via the LC-MS approach and the normalized expression is plotted against the applied dosages.

[0197] FIG. **30**: Cells were lipofected with DP (RNA-LPX) encoding for one tumor antigen in a dose-dependent manner (eleven different dosages per antigen). After 24 h incubation time the cells were harvested and subjected to LC-MS analysis. The Peptide 3 (of the p2p16 domain) is quantified via the LC-MS approach and the normalized expression is plotted against the applied dosages.

DESCRIPTION OF THE SEQUENCES

[0198] The following table provides a listing of certain sequences referenced herein.

TABLE-US-00001 TABLE 1 DESCRIPTION OF THE SEQUENCES SEQ ID NO: Description SEQUENCE
Sec/MITD 1 Sec (amino MRVMAPRTLILLLSGALALTETWAGS acid) 2 MITD (amino
IVGIVAGLAVLAVVVIGAVVATVMCRRKSSGGKGGSYSQAASSDSAQGSDVSLTA acid) P2P16 3 P2P16 (amino
KKQYIKANSKFIGITELKKLGGGKRGGGKKMTNSVDDALINSTKIYSYFPSVISKVNQGAQGKKL acid) GS
Linker 4 GS Linker 1 GGSGGGGSGG 5 GS Linker 2 GSSGGGGSPGGGSS 5′-UTR (hAg-Kozak) 6 5′-
UTR AACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACC 3′-UTR (FI element) 7 3′-
UTR
CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUCCCCCGACCUCGGGU
CCCAGGUAUGCUCCCACCUCCACCUGCCCCACUCACCACCUCUGCUAGUUCCAGACACCUCCCAAGCACGCAGCA
UGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAUAAACGAA
AGUUUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC A30L70 8 A30L70
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA Helper epitopes 9 P2 QYIKANSKFIGITEL 10
P16 MTNSVDDALINSTKIYSYFPSVISKVNQGAQG

DETAILED DESCRIPTION OF THE INVENTION

[0199] Although the present disclosure is further described in more 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.

[0200] In the following, the elements of the present disclosure will be described in more detail. 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 preferred embodiments should not be construed to limit the present disclosure to only the explicitly described embodiments. This description should be understood to support and encompass embodiments which combine the explicitly described embodiments with any number of the disclosed and/or preferred elements. Furthermore, any permutations and combinations of all described elements in this application should be considered disclosed by the description of the present application unless the context indicates otherwise.

[0201] The practice of the present disclosure will employ, unless otherwise indicated, conventional chemistry, biochemistry, cell biology, immunology, and recombinant DNA techniques which are explained in the literature in the field.

[0202] Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated feature, element, member, integer or step or group of features, elements, members, integers or steps but not the exclusion of any other feature, element, member, integer or step or group of features, elements, members, integers or steps. The term "consisting essentially of" limits the scope of a claim or disclosure to the specified features, elements, members, integers, or steps and those that do not materially affect the basic and novel characteristic(s) of the claim or disclosure. The term "consisting of" limits the scope of a claim or disclosure to the specified features, elements, members, integers, or steps. The term "comprising" encompasses the term "consisting essentially of" which, in turn, encompasses the term "consisting of". Thus, at each occurrence in the present application, the term "comprising" may be replaced with the term "consisting essentially of" or "consisting of". Likewise, at each occurrence in the present application, the term "consisting essentially of" may be replaced with the term "consisting of".
[0203] The terms "a", "an" and "the" and similar references used in the context of describing the present 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 the context.
[0204] All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by the context.
[0205] The use of any and all examples, or exemplary language (e.g., "such as"), provided herein is intended merely to better illustrate the present disclosure and does not pose a limitation on the scope of the present disclosure otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the present disclosure.
[0206] The term "optional" or "optionally" as used herein means that the subsequently described event, circumstance or condition may or may not occur, and that the description includes instances where said event, circumstance, or condition occurs and instances in which it does not occur.
[0207] Where used herein, "and/or" is to be taken as specific disclosure of each of the two specified features or components with or without the other. For example, "X and/or Y" is to be taken as specific disclosure of each of (i) X, (ii) Y, and (iii) X and Y, just as if each is set out individually herein.
[0208] In the context of the present disclosure, the term "about" denotes an interval of accuracy that the person of ordinary skill will understand to still ensure the technical effect of the feature in question. The term typically indicates deviation from the indicated numerical value by ±10%, ±5%, ±4%, ±3%, ±2%, ±1%, ±0.9%, ±0.8%, ±0.7%, ±0.6%, ±0.5%, ±0.4%, ±0.3%, ±0.2%, ±0.1%, ±0.05%, and for example ±0.01%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±10%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±5%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±4%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±3%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±2%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±1%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.9%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.8%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.7%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.6%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.5%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.4%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.3%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.2%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.1%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.05%. In some embodiments, "about" indicates deviation from the indicated numerical value by ±0.01%. As will be appreciated by the person of ordinary skill, the specific such deviation for a numerical value for a given technical effect will depend on the nature of the technical effect. For example, a natural or biological technical effect may generally have a larger such deviation than one for a man-made or engineering technical effect.
[0209] 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 were individually recited herein.
[0210] 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 invention is not entitled to antedate such disclosure by virtue of prior invention.
Definitions
[0211] In the following, definitions will be provided which apply to all aspects of the present disclosure.
[0212] The following terms have the following meanings unless otherwise indicated. Any undefined terms have their art recognized meanings.
[0213] The "therapeutic potential" or "potency" of nucleic acid (such as RNA and/or DNA) refers to the therapeutic quality of the nucleic acid, the ability of the nucleic acid to provide a therapeutic benefit when administered to a subject. In particular embodiments, the therapeutic potential of nucleic acid can be measured, determined, identified, quantified, confirmed and/or validated by expression, in particular strong expression, e.g., expression above a threshold, of the peptide or polypeptide encoded by the nucleic acid that indicates the therapeutic potential of the nucleic acid. In one embodiment,

therapeutic potential refers to an ability of a nucleic acid (such as an RNA and/or DNA) to express a pharmaceutically active peptide or polypeptide in vivo said pharmaceutically active peptide or polypeptide exerting its pharmaceutical, e.g., therapeutic, effect.

[0214] In some embodiments, nucleic acid (such as RNA and/or DNA) that shows strong expression, e.g., expression above a threshold, has "sufficient therapeutic potential". The therapeutic potential of the nucleic acid is sufficient if the nucleic acid has the ability in vivo to express an encoded pharmaceutically active peptide or polypeptide such that that a meaningful pharmaceutical, e.g., therapeutic, effect is achieved.

[0215] As used herein, phrases such as "determining the amount" or "determining expression" or similar phrases with reference to an amino acid sequence (peptide or polypeptide) refer to determining the quantity or presence of an amino acid sequence.

[0216] Terms such as "reduce" or "inhibit" as used herein means the ability to cause an overall decrease, for example, of about 5% or greater, about 10% or greater, about 15% or greater, about 20% or greater, about 25% or greater, about 30% or greater, about 40% or greater, about 50% or greater, or about 75% or greater, in the level. The term "inhibit" or similar phrases includes a complete or essentially complete inhibition, i.e. a reduction to zero or essentially to zero.

[0217] The term "enhance" as used herein means the ability to cause an overall increase, or enhancement, for example, by at least about 5% or greater, about 10% or greater, about 15% or greater, about 20% or greater, about 25% or greater, about 30% or greater, about 40% or greater, about 50% or greater, about 75% or greater, or about 100% or greater in the level.

[0218] "Physiological pH" as used herein refers to a pH of about 7.4. In some embodiments, physiological pH is from 7.3 to 7.5. In some embodiments, physiological pH is from 7.35 to 7.45. In some embodiments, physiological pH is 7.3, 7.35, 7.4, 7.45, or 7.5.

[0219] As used in the present disclosure, "% w/v" refers to weight by volume percent, which is a unit of concentration measuring the amount of solute in grams (g) expressed as a percent of the total volume of solution in milliliters (mL).

[0220] As used in the present disclosure, "% by weight" refers to weight percent, which is a unit of concentration measuring the amount of a substance in grams (g) expressed as a percent of the total weight of the total composition in grams (g).

[0221] As used in the present disclosure, "mol %" is defined as the ratio of the number of moles of one component to the total number of moles of all components, multiplied by 100.

[0222] As used in the present disclosure, "mol % of the total lipid" is defined as the ratio of the number of moles of one lipid component to the total number of moles of all lipids, multiplied by 100. In this context, in some embodiments, the term "total lipid" includes lipids and lipid-like material.

[0223] The term "ionic strength" refers to the mathematical relationship between the number of different kinds of ionic species in a particular solution and their respective charges. Thus, ionic strength I is represented mathematically by the formula:

[00001] $I = \frac{1}{2}$ .Math. $\underset{i}{\text{.Math.}}$ z_i^2 .Math. c_i

in which c is the molar concentration of a particular ionic species and z the absolute value of its charge. The sum Σ is taken over all the different kinds of ions (i) in solution.

[0224] According to the disclosure, the term "ionic strength" in some embodiments relates to the presence of monovalent ions. Regarding the presence of divalent ions, in particular divalent cations, their concentration or effective concentration (presence of free ions) due to the presence of chelating agents is, in some embodiments, sufficiently low so as to prevent degradation of the nucleic acid. In some embodiments, the concentration or effective concentration of divalent ions is below the catalytic level for hydrolysis of the phosphodiester bonds between nucleotides such as RNA nucleotides. In some embodiments, the concentration of free divalent ions is 20 μM or less. In some embodiments, there are no or essentially no free divalent ions. "Osmolality" refers to the concentration of a particular solute expressed as the number of osmoles of solute per kilogram of solvent.

[0225] The term "lyophilizing" or "lyophilization" refers to the freeze-drying of a substance by freezing it and then reducing the surrounding pressure (e.g., below 15 Pa, such as below 10 Pa, below 5 Pa, or 1 Pa or less) to allow the frozen medium in the substance to sublimate directly from the solid phase to the gas phase. Thus, the terms "lyophilizing" and "freeze-drying" are used herein interchangeably.

[0226] The term "spray-drying" refers to spray-drying a substance by mixing (heated) gas with a fluid that is atomized (sprayed) within a vessel (spray dryer), where the solvent from the formed droplets evaporates, leading to a dry powder.

[0227] The term "reconstitute" relates to adding a solvent such as water to a dried product to return it to a liquid state such as its original liquid state.

[0228] The term "recombinant" in the context of the present disclosure means "made through genetic engineering". In one embodiment, a "recombinant object" in the context of the present disclosure is not occurring naturally.

[0229] 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. The term "found in nature" means "present in nature" and includes known objects as well as objects that have not yet been discovered and/or isolated from nature, but that may be discovered and/or isolated in the future from a natural source.

[0230] As used herein, the terms "room temperature" and "ambient temperature" are used interchangeably herein and refer to temperatures from at least about 15° C., e.g., from about 15° C. to about 35° C., from about 15° C. to about 30° C., from about 15° C. to about 25° C., or from about 17° C. to about 22° C. Such temperatures will include 15° C., 16° C., 17° C.,

18° C., 19° C., 20° C., 21° C. and 22° C. In some embodiments, the temperature is from 15° C. to about 25° C. In some embodiments, the temperature is from 17° C. to about 25° C. In some embodiments, the temperature is about 15° C. In some embodiments, the temperature is about 16° C. In some embodiments, the temperature is about 17° C. In some embodiments, the temperature is about 18° C. In some embodiments, the temperature is about 19° C. In some embodiments, the temperature is about 20° C. In some embodiments, the temperature is about 21° C. In some embodiments, the temperature is about 22° C.

[0231] The term “EDTA” refers to ethylenediaminetetraacetic acid disodium salt. All concentrations are given with respect to the EDTA disodium salt.

[0232] The term “cryoprotectant” relates to a substance that is added to a formulation in order to protect the active ingredients during the freezing stages.

[0233] The term “lyoprotectant” relates to a substance that is added to a formulation in order to protect the active ingredients during the drying stages.

[0234] According to the present disclosure, the term “peptide” 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 “polypeptide” refers to large peptides, in particular peptides having at least about 151 amino acids. “Peptides” and “polypeptides” are both protein molecules.

[0235] The term “biological activity” means the response of a biological system to a molecule. Such biological systems may be, for example, a cell or an organism. In some embodiments, such response is therapeutically or pharmaceutically useful. In some embodiments, a biological activity comprises a pharmaceutical activity.

[0236] The term “biological system”, as used herein, refers to any system of interacting or potentially interacting biological constituents whose behavior can be characterized in whole or part by one or more biological processes or mechanisms. A biological system can include, for example, an individual cell, a collection of cells such as a cell culture, an organ, a tissue, and a multi-cellular organism such as an individual or subject, e.g., a human patient.

[0237] In some embodiments, a biological system is present in or is an individual or subject and a biological activity in such biological system is an activity which is therapeutically or pharmaceutically useful, i.e., the biological activity results in or contributes to a therapeutically or pharmaceutically useful effect.

[0238] According to various embodiments of the present disclosure, a nucleic acid (such as RNA and/or DNA) encoding a peptide or polypeptide is taken up by or introduced, i.e. transfected or transduced, into a cell which cell may be present in vitro or in a subject, resulting in expression of said peptide or polypeptide. The cell may, e.g., express the encoded peptide or polypeptide intracellularly (e.g. in the cytoplasm and/or in the nucleus), may secrete the encoded peptide or polypeptide, and/or may express it on the surface.

[0239] According to the present disclosure, terms such as “nucleic acid expressing” and “nucleic acid encoding” or similar terms are used interchangeably herein and with respect to a particular peptide or polypeptide mean that the nucleic acid, if present in the appropriate environment, e.g. within a cell, can be expressed to produce said peptide or polypeptide.

[0240] The term “portion” refers to a fraction. With respect to a particular structure such as an amino acid sequence or protein the term “portion” thereof may designate a continuous or a discontinuous fraction of said structure.

[0241] The terms “part” and “fragment” are used interchangeably herein and refer to a continuous element. For example, a part of a structure such as an amino acid sequence or protein refers to a continuous element of said structure. When used in context of a composition, the term “part” means a portion of the composition. For example, a part of a composition may be any portion from 0.1% to 99.9% (such as 0.1%, 0.5%, 1%, 5%, 10%, 50%, 90%, or 99%) of said composition. “Fragment”, with reference to an amino acid sequence (peptide or polypeptide), 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 comprises, e.g., at least 6, in particular at least 8, at least 10, 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. A fragment of an amino acid sequence comprises, e.g., a sequence of up to 8, in particular up to 10, up to 12, up to 15, up to 20, up to 30 or up to 55, consecutive amino acids of the amino acid sequence. “Variant,” as used herein and with reference to an amino acid sequence (peptide or polypeptide), is meant an amino acid sequence that differs from a parent amino acid sequence by virtue of at least one amino acid (e.g., a different amino acid, or a modification of the same amino acid). 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. In some embodiments, the variant amino acid sequence has at least one amino acid difference as compared to the parent amino acid sequence, e.g., from 1 to about 20 amino acid differences, such as from 1 to about 10 or from 1 to about 5 amino acid differences compared to the parent.

[0242] 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 polypeptide has an amino acid sequence that has not been intentionally modified.

[0243] For the purposes of the present disclosure, “variants” of an amino acid sequence (peptide or polypeptide) may

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, post-translationally modified variants, conformations, 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.

[0244] 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.

[0245] 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.

[0246] 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 peptides or polypeptides and/or to replacing amino acids with other ones having similar properties. In some embodiments, amino acid changes in peptide and polypeptide 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 some embodiments, conservative amino acid substitutions include substitutions within the following groups: [0247] glycine, alanine; [0248] valine, isoleucine, leucine; [0249] aspartic acid, glutamic acid; [0250] asparagine, glutamine; [0251] serine, threonine; [0252] lysine, arginine; and [0253] phenylalanine, tyrosine.

[0254] In some embodiments the degree of similarity, such as 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%. In some embodiments, the degree of similarity or identity is given 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, e.g., 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, such as sequence identity, can be done with art known tools, such as 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.

[0255] “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.

[0256] The terms “% identical” and “% 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.
[0257] 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.
[0258] 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 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.
[0259] 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, e.g., at least 95%, at least 98 or at least 99% identity of the amino acid residues.
[0260] 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 polypeptides having substitutions, additions, insertions or deletions, is described in detail in Molecular Cloning: A Laboratory Manual, 4.sup.th Edition, M. R. Green and J. Sambrook eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2012, for example. Furthermore, the peptides, polypeptides 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.
[0261] In some embodiments, a fragment or variant of an amino acid sequence (peptide or polypeptide) is 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 some embodiments, 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., function of the functional fragment or 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, function of the functional fragment or functional variant may be enhanced compared to the parent molecule or sequence.
[0262] An amino acid sequence (peptide or polypeptide) “derived from” a designated amino acid sequence (peptide or polypeptide) refers to the origin of the first amino acid sequence. In some embodiments, 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.
[0263] In some embodiments, “isolated” means removed (e.g., purified) from the natural state or from an artificial composition, such as a composition from a production process. For example, a nucleic acid, peptide or polypeptide naturally present in a living animal is not “isolated”, but the same nucleic acid, peptide or polypeptide partially or completely separated from the coexisting materials of its natural state is “isolated”. An isolated nucleic acid, peptide or polypeptide can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.
[0264] The term “transfection” relates to the introduction of nucleic acids, in particular RNA, into a cell. For purposes of the present disclosure, 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, or the cell may be in vitro, e.g., outside of a patient. Thus, according to the present disclosure, 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 the body of a patient. According to the disclosure, 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, for example. Generally, nucleic acid encoding antigen is transiently transfected into cells. RNA can be transfected into cells to transiently express its coded protein.
[0265] Cells which are useful for transfection in the methods described herein include, but are not limited to, cells from an animal cell line, such as Chinese hamster ovary (CHO), K562, HepG2, HEK293T, RAW, and C2C12 cells. In some

embodiments, the cells are CHO, K562, HEK293T, RAW, and C2C12 cells. In some embodiments, the cells are Chinese hamster ovary (CHO) cells.
[0266] The disclosure includes analogs of a peptide or polypeptide. According to the present disclosure, an analog of a peptide or polypeptide is a modified form of said peptide or polypeptide from which it has been derived and has at least one functional property of said peptide or polypeptide. E.g., a pharmacological active analog of a peptide or polypeptide has at least one of the pharmacological activities of the peptide or polypeptide from which the analog has been derived. Such modifications include any chemical modification and comprise single or multiple substitutions, deletions and/or additions of any molecules associated with the peptide or polypeptide, such as carbohydrates, lipids and/or peptides or polypeptides. In some embodiments, “analogs” of peptides or polypeptides include those modified forms resulting from glycosylation, acetylation, phosphorylation, amidation, palmitoylation, myristoylation, isoprenylation, lipidation, alkylation, derivatization, introduction of protective/blocking groups, proteolytic cleavage or binding to an antibody or to another cellular ligand. The term “analog” also extends to all functional chemical equivalents of said peptides and polypeptides.
[0267] 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.
[0268] As used herein “endogenous” refers to any material from or produced inside an organism, cell, tissue or system.
[0269] As used herein, the term “exogenous” refers to any material introduced from or produced outside an organism, cell, tissue or system.
[0270] The term “expression” as used herein is defined as the transcription and/or translation of a particular nucleotide sequence.
[0271] 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 (especially mRNA). Subsequently, the RNA may be translated into peptide or polypeptide.
[0272] 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 polypeptide.
[0273] Prodrugs of a particular compound described herein are those compounds that upon administration to an individual undergo chemical conversion under physiological conditions to provide the particular compound. Additionally, prodrugs can be converted to the particular compound by chemical or biochemical methods in an ex vivo environment. For example, prodrugs can be slowly converted to the particular compound when, for example, placed in a transdermal patch reservoir with a suitable enzyme or chemical reagent. Exemplary prodrugs are esters (using an alcohol or a carboxy group contained in the particular compound) or amides (using an amino or a carboxy group contained in the particular compound) which are hydrolyzable in vivo. Specifically, any amino group which is contained in the particular compound and which bears at least one hydrogen atom can be converted into a prodrug form. Typical N-prodrug forms include carbamates, Mannich bases, enamines, and enaminones.
[0274] In the present specification, a structural formula of a compound may represent a certain isomer of said compound. It is to be understood, however, that the present invention includes all isomers such as geometrical isomers, optical isomers based on an asymmetrical carbon, stereoisomers, tautomers and the like which occur structurally and isomer mixtures and is not limited to the description of the formula.
[0275] “Isomers” are compounds having the same molecular formula but differ in structure (“structural isomers”) or in the geometrical (spatial) positioning of the functional groups and/or atoms (“stereoisomers”). “Enantiomers” are a pair of stereoisomers which are non-superimposable mirror-images of each other. A “racemic mixture” or “racemate” contains a pair of enantiomers in equal amounts and is denoted by the prefix (±). “Diastereomers” are stereoisomers which are non-superimposable and which are not mirror-images of each other. “Tautomers” are structural isomers of the same chemical substance that spontaneously and reversibly interconvert into each other, even when pure, due to the migration of individual atoms or groups of atoms; i.e., the tautomers are in a dynamic chemical equilibrium with each other. An example of tautomers are the isomers of the keto-enol-tautomerism. “Conformers” are stereoisomers that can be interconverted just by rotations about formally single bonds, and include—in particular—those leading to different 3-dimensional forms of (hetero)cyclic rings, such as chair, half-chair, boat, and twist-boat forms of cyclohexane.
[0276] The term “average diameter” refers to the mean hydrodynamic diameter of particles as measured by dynamic light scattering (DLS) with data analysis using the so-called cumulant algorithm, which provides as results the so-called Z.sub.average with the dimension of a length, and the polydispersity index (PDI), 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.sub.average.
[0277] In some embodiments, the “polydispersity index” is may be 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.
[0278] The “radius of gyration” (abbreviated herein as R.sub.g) of a particle about an axis of rotation is the radial distance of a point from the axis of rotation at which, if the whole mass of the particle is assumed to be concentrated, its moment of inertia about the given axis would be the same as with its actual distribution of mass. Mathematically, R.sub.g is the root mean square distance of the particle's components from either its center of mass or a given axis. For example, for a macromolecule composed of n mass elements, of masses m.sub.i (i=1, 2, 3, . . . , n), located at fixed distances s.sub.i from the center of mass, R.sub.g is the square-root of the mass average of s.sub.i.sup.2 over all mass elements and can be

calculated as follows:

[00002] $R_g = \left(\sum_{i=1}^{n} m_i \, s_i^2 \right) / \left(\sum_{i=1}^{n} m_i \right)^{1/2}$

[0279] The radius of gyration can be determined or calculated experimentally, e.g., by using light scattering. In particular, for small scattering vectors a the structure function S is defined as follows:

[00003] $S(\vec{q}) \approx N \left(1 - \frac{q^2 R_g^2}{3}\right)$

wherein N is the number of components (Guinier's law).

[0280] The “hydrodynamic radius” (which is sometimes called “Stokes radius” or “Stokes-Einstein radius”) of a particle is the radius of a hypothetical hard sphere that diffuses at the same rate as said particle. The hydrodynamic radius is related to the mobility of the particle, taking into account not only size but also solvent effects. For example, a smaller charged particle with stronger hydration may have a greater hydrodynamic radius than a larger charged particle with weaker hydration. This is because the smaller particle drags a greater number ofwater molecules with it as it moves through the solution. Since the actual dimensions of the particle in a solvent are not directly measurable, the hydrodynamic radius may be defined by the Stokes-Einstein equation:

[00004] $R_h = \frac{k_B T}{6 \pi \eta D}$

wherein k.sub.B is the Boltzmann constant; T is the temperature; n is the viscosity of the solvent; and D is the diffusion coefficient. The diffusion coefficient can be determined experimentally, e.g., by using dynamic light scattering (DLS). Thus, one procedure to determine the hydrodynamic radius of a particle or a population of particles (such as the hydrodynamic radius of particles contained in a sample or control composition as disclosed herein or the hydrodynamic radius of a particle peak obtained from subjecting such a sample or control composition to field-flow fractionation) is to measure the DLS signal of said particle or population of particles (such as DLS signal of particles contained in a sample or control composition as disclosed herein or the DLS signal of a particle peak obtained from subjecting such a sample or control composition to field-flow fractionation).

[0281] The expression “light scattering” as used herein refers to the physical process where light is forced to deviate from a straight trajectory by one or more paths due to localized non-uniformities in the medium through which the light passes.

[0282] The term “UV” means ultraviolet and designates a band of the electromagnetic spectrum with a wavelength from 10 nm to 400 nm, i.e., shorter than that of visible light but longer than X-rays. The expression “multi-angle light scattering” or “MALS” as used herein relates to a technique for measuring the light scattered by a sample into a plurality of angles. “Multi-angle” means in this respect that scattered light can be detected at different discrete angles as measured, for example, by a single detector moved over a range including the specific angles selected or an array of detectors fixed at specific angular locations. In certain embodiments, the light source used in MALS is a laser source (MALLS: multi-angle laser light scattering). Based on the MALS signal of a composition comprising particles and by using an appropriate formalism (e.g., Zimm plot, Berry plot, or Debye plot), it is possible to determine the radius of gyration (R.sub.g) and, thus, the size of said particles. Preferably, the Zimm plot is a graphical presentation using the following equation:

[00005] $\frac{R}{K^* c} = M_w P(\quad) - 2A_2 c M_w^2 P^2(\quad)$

wherein c is the mass concentration of the particles in the solvent (g/mL); A.sub.2 is the second virial coefficient (mol.Math.mL/g.sup.2); P(ϑ) is a form factor relating to the dependence of scattered light intensity on angle; R.sub.ϑ is the excess Rayleigh ratio (cm.sup.−1); and K* is an optical constant that is equal to 4π.sup.2η.sub.o(dn/dc).sup.2λ.sub.0.sup.−4N.sub.A.sup.−1, where η.sub.o is the refractive index of the solvent at the incident radiation (vacuum) wavelength, λ.sub.0 is the incident radiation (vacuum) wavelength (nm), N.sub.A is Avogadro's number (mol.sup.−1), and dn/dc is the differential refractive index increment (mL/g) (cf., e.g., Buchholz et al. (Electrophoresis 22 (2001), 4118-4128); B. H. Zimm (J. Chem. Phys. 13 (1945), 141; P. Debye (J. Appl. Phys. 15 (1944): 338; and W. Burchard (Anal. Chem. 75 (2003), 4279-4291). Preferably, the Berry plot is calculated the following term:

[00006] $\sqrt{\frac{R}{K^* c}}$

wherein c, R.sub.ϑ and K* are as defined above. Preferably, the Debye plot is calculated the following term:

[00007] $\frac{K^* c}{R}$

wherein c, R.sub.ϑ and K* are as defined above.

[0283] The expression “dynamic light scattering” or “DLS” as used herein refers to a technique to determine the size and size distribution profile of particles, in particular with respect to the hydrodynamic radius of the particles. A monochromatic light source, usually a laser, is shot through a polarizer and into a sample. The scattered light then goes through a second polarizer where it is detected and the resulting image is projected onto a screen. The particles in the solution are being hit with the light and diffract the light in all directions. The diffracted light from the particles can either interfere constructively (light regions) or destructively (dark regions). This process is repeated at short time intervals and the resulting set of speckle patterns are analyzed by an autocorrelator that compares the intensity of light at each spot over time.

[0284] The expression “static light scattering” or “SLS” as used herein refers to a technique to determine the size and size distribution profile of particles, in particular with respect to the radius of gyration of the particles, and/or the molar mass of particles. A high-intensity monochromatic light, usually a laser, is launched in a solution containing the particles. One or many detectors are used to measure the scattering intensity at one or many angles. The angular dependence is needed to obtain accurate measurements of both molar mass and size for all macromolecules of radius. Hence simultaneous measurements at several angles relative to the direction of incident light, known as multi-angle light scattering (MALS) or

multi-angle laser light scattering (MALLS), is generally regarded as the standard implementation of static light scattering.

Nucleic Acids

[0285] The term “nucleic acid” comprises deoxyribonucleic acid (DNA), ribonucleic acid (RNA), combinations thereof, and modified forms thereof. The term comprises genomic DNA, cDNA, mRNA, recombinantly produced and chemically synthesized molecules. In some embodiments, a nucleic acid is DNA. In some embodiments, a nucleic acid is RNA. In some embodiments, a nucleic acid is a mixture of DNA and RNA. In some embodiments, a nucleic acid is DNA. A nucleic acid may be present as a single-stranded or double-stranded and linear or covalently circularly closed molecule. A nucleic acid can be isolated. The term “isolated nucleic acid” means, according to the present disclosure, that the nucleic acid (i) was amplified in vitro, for example via polymerase chain reaction (PCR) for DNA or in vitro transcription (using, e.g., an RNA polymerase) for RNA, (ii) was produced recombinantly by cloning, (iii) was purified, for example, by cleavage and separation by gel electrophoresis, or (iv) was synthesized, for example, by chemical synthesis. The term “nucleoside” (abbreviated herein as “N”) relates to compounds which can be thought of as nucleotides without a phosphate group. While a nucleoside is a nucleobase linked to a sugar (e.g., ribose or deoxyribose), a nucleotide is composed of a nucleoside and one or more phosphate groups. Examples of nucleosides include cytidine, uridine, pseudouridine, adenosine, and guanosine.

[0286] The five standard nucleosides which usually make up naturally occurring nucleic acids are uridine, adenosine, thymidine, cytidine and guanosine. The five nucleosides are commonly abbreviated to their one letter codes U, A, T, C and G, respectively. However, thymidine is more commonly written as “dT” (“d” represents “deoxy”) as it contains a 2′-deoxyribofuranose moiety rather than the ribofuranose ring found in uridine. This is because thymidine is found in deoxyribonucleic acid (DNA) and not ribonucleic acid (RNA). Conversely, uridine is found in RNA and not DNA. The remaining three nucleosides may be found in both RNA and DNA. In RNA, they would be represented as A, C and G, whereas in DNA they would be represented as dA, dC and dG.

[0287] A modified purine (A or G) or pyrimidine (C, T, or U) base moiety is preferably modified by one or more alkyl groups, more preferably one or more C.sub.1-4 alkyl groups, even more preferably one or more methyl groups. Particular examples of modified purine or pyrimidine base moieties include N.sup.7-alkyl-guanine, N.sup.6-alkyl-adenine, 5-alkyl-cytosine, 5-alkyl-uracil, and N(1)-alkyl-uracil, such as N.sup.7—C.sub.1-4 alkyl-guanine, N.sup.6—C.sub.1-4 alkyl-adenine, 5-C.sub.1-4 alkyl-cytosine, 5-C.sub.1-4 alkyl-uracil, and N(1)-C.sub.1-4 alkyl-uracil, preferably N.sup.7-methyl-guanine, N.sup.6-methyl-adenine, 5-methyl-cytosine, 5-methyl-uracil, and N(1)-methyl-uracil.

[0288] Herein, the term “DNA” relates to a nucleic acid molecule which includes deoxyribonucleotide residues. In preferred embodiments, the DNA contains all or a majority of deoxyribonucleotide residues. As used herein, “deoxyribonucleotide” refers to a nucleotide which lacks a hydroxyl group at the 2′-position of a β-D-ribofuranosyl group. DNA encompasses without limitation, double stranded DNA, single stranded DNA, isolated DNA such as partially purified DNA, essentially pure DNA, synthetic DNA, recombinantly produced DNA, as well as modified DNA that differs from naturally occurring DNA 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 DNA nucleotides or to the end(s) of DNA. It is also contemplated herein that nucleotides in DNA may be non-standard nucleotides, such as chemically synthesized nucleotides or ribonucleotides. For the present disclosure, these altered DNAs are considered analogs of naturally-occurring DNA. A molecule contains “a majority of deoxyribonucleotide residues” if the content of deoxyribonucleotide residues in the molecule is more than 50% (such as at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%), based on the total number of nucleotide residues in the molecule. The total number of nucleotide residues in a molecule is the sum of all nucleotide residues (irrespective of whether the nucleotide residues are standard (i.e., naturally occurring) nucleotide residues or analogs thereof).

[0289] DNA may be recombinant DNA and may be obtained by cloning of a nucleic acid, in particular cDNA. The cDNA may be obtained by reverse transcription of RNA.

[0290] 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 herein that nucleotides in RNA may be non-standard nucleotides, such as chemically synthesized nucleotides or deoxynucleotides. For the present disclosure, these altered/modified nucleotides can be referred to as analogs of naturally occurring nucleotides, and the corresponding RNAs containing such altered/modified nucleotides (i.e., altered/modified RNAs) can be referred to as analogs of naturally occurring RNAs. A molecule contains “a majority of ribonucleotide residues” if the content of ribonucleotide residues in the molecule is more than 50% (such as at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%), based on the total number of nucleotide residues in the molecule. The total number of nucleotide residues in a molecule is the sum of all nucleotide residues (irrespective of whether the nucleotide residues are standard (i.e., naturally occurring) nucleotide residues or analogs thereof).

[0291] “RNA” includes mRNA, tRNA, ribosomal RNA (rRNA), small nuclear RNA (snRNA), self-amplifying RNA

(saRNA), single-stranded RNA (ssRNA), dsRNA, inhibitory RNA (such as antisense ssRNA, small interfering RNA (siRNA), or microRNA (miRNA)), activating RNA (such as small activating RNA) and immunostimulatory RNA (isRNA). In some embodiments, "RNA" refers to mRNA.

[0292] The term "in vitro transcription" or "IVT" as used herein means that the transcription (i.e., the generation of RNA) is conducted in a cell-free manner. I.e., IVT does not use living/cultured cells but rather the transcription machinery extracted from cells (e.g., cell lysates or the isolated components thereof, including an RNA polymerase (preferably T7, T3 or SP6 polymerase)).

mRNA

[0293] According to the present disclosure, the term "mRNA" means "messenger-RNA" and relates to a "transcript" which may be generated by using a DNA template and may encode a peptide or polypeptide. Typically, an mRNA comprises a 5′-UTR, a peptide/polypeptide coding region, and a 3′-UTR. In the context of the present disclosure, mRNA may be generated by in vitro transcription (IVT) from a DNA template. As set forth above, the in vitro transcription methodology is known to the skilled person, and a variety of in vitro transcription kits is commercially available. mRNA is single-stranded but may contain self-complementary sequences that allow parts of the mRNA to fold and pair with itself to form double helices.

[0294] According to the present disclosure, "dsRNA" means double-stranded RNA and is RNA with two partially or completely complementary strands.

[0295] In preferred embodiments of the present disclosure, the mRNA relates to an RNA transcript which encodes a peptide or polypeptide.

[0296] In some embodiments, the mRNA which preferably encodes a peptide or polypeptide has a length of at least 45 nucleotides (such as at least 60, at least 90, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1,000, at least 1,500, at least 2,000, at least 2,500, at least 3,000, at least 3,500, at least 4,000, at least 4,500, at least 5,000, at least 6,000, at least 7,000, at least 8,000, at least 9,000 nucleotides), preferably up to 15,000, such as up to 14,000, up to 13,000, up to 12,000 nucleotides, up to 11,000 nucleotides or up to 10,000 nucleotides.

[0297] As established in the art, mRNA generally contains a 5′ untranslated region (5′-UTR), a peptide/polypeptide coding region and a 3′ untranslated region (3′-UTR). In some embodiments, the mRNA is produced by in vitro transcription or chemical synthesis. In some embodiments, the mRNA is produced by in vitro transcription using a DNA template. The in vitro transcription methodology is known to the skilled person; cf., e.g., Molecular Cloning: A Laboratory Manual, 4.sup.th Edition, M. R. Green and J. Sambrook eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2012. Furthermore, a variety of in vitro transcription kits is commercially available, e.g., from Thermo Fisher Scientific (such as TranscriptAid™ T7 kit, MEGAscript® T7 kit, MAXlscript®), New England BioLabs Inc. (such as HiScribe™ T7 kit, HiScribe™ T7 ARCA mRNA kit), Promega (such as RiboMAX™, HeLaScribe®, Riboprobe® systems), Jena Bioscience (such as SP6 or T7 transcription kits), and Epicentre (such as AmpliScribe™). For providing modified mRNA, correspondingly modified nucleotides, such as modified naturally occurring nucleotides, non-naturally occurring nucleotides and/or modified non-naturally occurring nucleotides, can be incorporated during synthesis (preferably in vitro transcription), or modifications can be effected in and/or added to the mRNA after transcription.

[0298] In some embodiments, mRNA is in vitro transcribed mRNA (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. Particular examples of RNA polymerases are the T7, T3, and SP6 RNA polymerases. Preferably, the in vitro transcription 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.

[0299] In some embodiments of the present disclosure, the mRNA is "replicon mRNA" or simply a "replicon", in particular "self-replicating mRNA" or "self-amplifying mRNA". In certain embodiments, the replicon or self-replicating mRNA is derived from or comprises elements derived from an 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 Jose 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 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.

[0300] In some embodiments of the present disclosure, the mRNA contains one or more modifications, e.g., in order to increase its stability and/or increase translation efficiency and/or decrease immunogenicity and/or decrease cytotoxicity. For example, in order to increase expression of the mRNA, it may be modified within the coding region, i.e., the sequence encoding the expressed peptide or polypeptide, preferably without altering the sequence of the expressed peptide or polypeptide. Such modifications are described, for example, in WO 2007/036366 and PCT/EP2019/056502, and include the following: a 5′-cap structure; an extension or truncation of the naturally occurring poly(A) tail; an alteration of the 5′- and/or 3′-untranslated regions (UTR) such as introduction of a UTR which is not related to the coding region of said RNA; the replacement of one or more naturally occurring nucleotides with synthetic nucleotides; and codon optimization (e.g., to alter, preferably increase, the GC content of the RNA).

[0301] In some embodiments, the mRNA comprises a 5′-cap structure. In some embodiments, the mRNA does not have uncapped 5′-triphosphates. In some embodiments, the mRNA may comprise a conventional 5′-cap and/or a 5′-cap analog. The term “conventional 5′-cap” refers to a cap structure found on the 5′-end of an mRNA molecule and generally consists of a guanosine 5′-triphosphate (Gppp) which is connected via its triphosphate moiety to the 5′-end of the next nucleotide of the mRNA (i.e., the guanosine is connected via a 5′ to 5′ triphosphate linkage to the rest of the mRNA). The guanosine may be methylated at position N.sup.7 (resulting in the cap structure m.sup.7Gppp). The term “5′-cap analog” includes a 5′-cap which is based on a conventional 5′-cap but which has been modified at either the 2′- or 3′-position of the m.sup.7guanosine structure in order to avoid an integration of the 5′-cap analog in the reverse orientation (such 5′-cap analogs are also called anti-reverse cap analogs (ARCAs)). Particularly preferred 5′-cap analogs are those having one or more substitutions at the bridging and non-bridging oxygen in the phosphate bridge, such as phosphorothioate modified 5′-cap analogs at the β-phosphate (such as m.sub.2.sup.7,2′OG(5′)ppSp(5′)G (referred to as beta-S-ARCA or β-S-ARCA)), as described in PCT/EP2019/056502. Providing an mRNA with a 5′-cap structure as described herein may be achieved by in vitro transcription of a DNA template in presence of a corresponding 5′-cap compound, wherein said 5′-cap structure is co-transcriptionally incorporated into the generated mRNA strand, or the mRNA may be generated, for example, by in vitro transcription, and the 5′-cap structure may be attached to the mRNA post-transcriptionally using capping enzymes, for example, capping enzymes of vaccinia virus.

[0302] In some embodiments, the mRNA comprises a 5′-cap structure selected from the group consisting of m.sub.2.sup.7,2′OG(5′)ppSp(5′)G (in particular its D1 diastereomer), m.sub.2.sup.7,3′OG(5′)ppp(5′)G, and m.sub.2.sup.7,3′-OGppp(m.sub.1.sup.2′-O)ApG.

[0303] In some embodiments, the mRNA comprises a cap0, cap1, or cap2, preferably cap1 or cap2. According to the present disclosure, the term “cap0” means the structure “m.sup.7GpppN”, wherein N is any nucleoside bearing an OH moiety at position 2′. According to the present disclosure, the term “cap1” means the structure “m.sup.7GpppNm”, wherein Nm is any nucleoside bearing an OCH.sub.3 moiety at position 2′. According to the present disclosure, the term “cap2” means the structure “m.sup.7GpppNmNm”, wherein each Nm is independently any nucleoside bearing an OCH.sub.3 moiety at position 2′.

[0304] The D1 diastereomer of beta-S-ARCA (β-S-ARCA) has the following structure:

##STR00001##

[0305] 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. The HPLC preferably is an analytical HPLC.

[0306] In some embodiments, a Supelcosil LC-18-T RP column, preferably of the format: 5 μm, 4.6×250 mm is used for separation, whereby a flow rate of 1.3 ml/min can be applied. In some embodiments, a gradient of methanol in ammonium acetate, for example, a 0-25% linear gradient of methanol in 0.05 M ammonium acetate, pH=5.9, within 15 min is used. UV-detection (VWD) can be performed at 260 nm and fluorescence detection (FLD) can be performed with excitation at 280 nm and detection at 337 nm.

[0307] The 5′-cap analog m.sub.2.sup.7,3′-OGppp(m.sub.1.sup.2′-O)ApG (also referred to as m.sub.2.sup.7,3′OG(5′)ppp(5′)m.sup.2′-OApG) which is a building block of a cap1 has the following structure:

##STR00002##

[0308] An exemplary cap0 mRNA comprising β-S-ARCA and mRNA has the following structure:

##STR00003##

[0309] An exemplary cap0 mRNA comprising m.sub.2.sup.7,3′OG(5′)ppp(5′)G and mRNA has the following structure:

##STR00004##

[0310] An exemplary cap1 mRNA comprising m.sub.2.sup.7,3′-OGppp(m.sub.1.sup.2′-O)ApG and mRNA has the following structure:

##STR00005##

[0311] As used herein, the term “poly-A tail” or “poly-A sequence” refers to an uninterrupted or interrupted sequence of adenylate residues which is typically located at the 3′-end of an mRNA molecule. Poly-A tails or poly-A sequences are known to those of skill in the art and may follow the 3′-UTR in the mRNAs described herein. An uninterrupted poly-A tail

is characterized by consecutive adenylate residues. In nature, an uninterrupted poly-A tail is typical. mRNAs disclosed herein can have a poly-A tail attached to the free 3′-end of the mRNA by a template-independent RNA polymerase after transcription or a poly-A tail encoded by DNA and transcribed by a template-dependent RNA polymerase.

[0312] It has been demonstrated that a poly-A tail of about 120 A nucleotides has a beneficial influence on the levels of mRNA 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 tail (Holtkamp et al., 2006, Blood, vol. 108, pp. 4009-4017).

[0313] The poly-A tail may be of any length. In some embodiments, a poly-A tail 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 tail, 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 tail 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 tail, i.e., 100% by number of nucleotides in the poly-A tail, are A nucleotides. The term "A nucleotide" or "A" refers to adenylate.

[0314] In some embodiments, a poly-A tail 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 tail (coding strand) is referred to as poly(A) cassette.

[0315] 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 disclosure. 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 tail contained in an mRNA molecule described herein 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.

[0316] In some embodiments, no nucleotides other than A nucleotides flank a poly-A tail at its 3′-end, i.e., the poly-A tail is not masked or followed at its 3′-end by a nucleotide other than A.

[0317] In some embodiments, a poly-A tail 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 tail 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 tail 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 tail comprises the poly-A tail shown in SEQ ID NO: 8. In some embodiments, the poly-A tail comprises at least 100 nucleotides. In some embodiments, the poly-A tail comprises about 150 nucleotides. In some embodiments, the poly-A tail comprises about 120 nucleotides.

[0318] In some embodiments, mRNA used in 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 generally 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. Incorporation of a 3′-UTR into the 3′-non translated region of an RNA (preferably mRNA) molecule can result in an enhancement in translation efficiency. A synergistic effect may be achieved by incorporating two or more of such 3′-UTRs (which are preferably arranged in a head-to-tail orientation; cf., e.g., Holtkamp et al., Blood 108, 4009-4017 (2006)). The 3′-UTRs may be autologous or heterologous to the RNA (e.g., mRNA) into which they are introduced. In certain embodiments, the 3′-UTR is derived from a globin gene or mRNA, such as a gene or mRNA of alpha2-globin, alpha1-globin, or beta-globin, e.g., beta-globin, e.g., human beta-globin. For example, the RNA (e.g., mRNA) may be modified by the replacement of the existing 3′-UTR with or the insertion of one or more, e.g., two copies of a 3′-UTR derived from a globin gene, such as alpha2-globin, alpha1-globin, beta-globin, e.g., beta-globin, e.g., human beta-globin.

[0319] A particularly preferred 5′-UTR comprises the nucleotide sequence of SEQ ID NO: 6. A particularly preferred 3′-UTR comprises the nucleotide sequence of SEQ ID NO: 7.

[0320] In some embodiments, RNA comprises a 5′-UTR comprising the nucleotide sequence of SEQ ID NO: 6, 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: 6.

[0321] In some embodiments, RNA comprises a 3′-UTR comprising the nucleotide sequence of SEQ ID NO: 7, 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: 7.

[0322] The mRNA may have modified ribonucleotides in order to increase its stability and/or decrease immunogenicity and/or decrease cytotoxicity. For example, in some embodiments, uridine in the mRNA described herein is replaced (partially or completely, preferably completely) by a modified nucleoside. In some embodiments, the modified nucleoside is a modified uridine.

[0323] In some embodiments, the modified uridine replacing uridine is selected from the group consisting of pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), 5-methyl-uridine (m5U), and combinations thereof.

[0324] In some embodiments, the modified nucleoside replacing (partially or completely, preferably completely) uridine in the mRNA may be any one or more of 3-methyl-uridine (m3U), 5-methoxy-uridine (mo5U), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s2U), 4-thio-uridine (s4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-oxyacetic acid (cmo5U), uridine 5-oxyacetic acid methyl ester (mcmo5U), 5-carboxymethyl-uridine (cm5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm5U), 5-carboxyhydroxymethyl-uridine methyl ester (mchm5U), 5-methoxycarbonylmethyl-uridine (mcm5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm5s2U), 5-aminomethyl-2-thio-uridine (nm5s2U), 5-methylaminomethyl-uridine (mnm5U), 1-ethyl-pseudouridine, 5-methylaminomethyl-2-thio-uridine (mnm5s2U), 5-methylaminomethyl-2-seleno-uridine (mnm5se2U), 5-carbamoylmethyl-uridine (ncm5U), 5-carboxymethylaminomethyl-uridine (cmnm5U), 5-carboxymethylaminomethyl-2-thio-uridine (cmnm5s2U), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (τm5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine(τm5s2U), 1-taurinomethyl-4-thio-pseudouridine), 5-methyl-2-thio-uridine (m5s2U), 1-methyl-4-thio-pseudouridine (m1s4ψ), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine (m3ψ), 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 (m5D), 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 (acp3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp3ψ), 5-(isopentenylaminomethyl)uridine (inm5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm5s2U), α-thio-uridine, 2′-O-methyl-uridine (Um), 5,2′-O-dimethyl-uridine (m5Um), 2′-O-methyl-pseudouridine (ψm), 2-thio-2′-O-methyl-uridine (s2Um), 5-methoxycarbonylmethyl-2′-O-methyl-uridine (mcm5Um), 5-carbamoylmethyl-2′-O-methyl-uridine (ncm5Um), 5-carboxymethylaminomethyl-2′-O-methyl-uridine (cmnm5Um), 3,2′-O-dimethyl-uridine (m3Um), 5-(isopentenylaminomethyl)-2′-O-methyl-uridine (inm5Um), 1-thio-uridine, deoxythymidine, 2′-F-ara-uridine, 2′-F-uridine, 2′-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, 5-[3-(1-E-propenylamino)uridine, or any other modified uridine known in the art.

[0325] An RNA (preferably mRNA) which is modified by pseudouridine (replacing partially or completely, preferably completely, uridine) is referred to herein as “Ψ-modified”, whereas the term “m1Ψ-modified” means that the RNA (preferably mRNA) contains N(1)-methylpseudouridine (replacing partially or completely, preferably completely, uridine). Furthermore, the term “m5U-modified” means that the RNA (preferably mRNA) contains 5-methyluridine (replacing partially or completely, preferably completely, uridine). Such Ψ- or m1Ψ- or m5U-modified RNAs usually exhibit decreased immunogenicity compared to their unmodified forms and, thus, are preferred in applications where the induction of an immune response is to be avoided or minimized. In some embodiments, the RNA (preferably mRNA) contains N(1)-methylpseudouridine replacing completely uridine

[0326] The codons of the mRNA used in the present disclosure may further be optimized, e.g., to increase the GC content of the RNA and/or to replace codons which are rare in the cell (or subject) in which the peptide or polypeptide of interest is to be expressed by codons which are synonymous frequent codons in said cell (or subject). In some embodiments, the amino acid sequence encoded by the mRNA used in the present disclosure 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 some embodiments, the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

[0327] 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 disclosure, coding regions may be codon-optimized for optimal expression in a subject to be treated using the mRNA 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 mRNA may be modified such that codons for which frequently occurring tRNAs are available are inserted in place of “rare codons”.

[0328] In some embodiments, the guanosine/cytosine (G/C) content of the coding region of the mRNA 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 mRNA is preferably not modified compared to the amino acid sequence encoded by the wild type RNA. This modification of the mRNA 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 favorable

codons for the stability can be determined (so-called alternative codon usage). Depending on the amino acid to be encoded by the mRNA, there are various possibilities for modification of the mRNA 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.
[0329] In various embodiments, the G/C content of the coding region of the mRNA 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.
[0330] A combination of the above described modifications, i.e., incorporation of a 5′-cap structure, incorporation of a poly-A sequence, unmasking of a poly-A sequence, alteration of the 5′- and/or 3′-UTR (such as incorporation of one or more 3′-UTRs), replacing one or more naturally occurring nucleotides with synthetic nucleotides (e.g., 5-methylcytidine for cytidine and/or pseudouridine (Ψ) or N(1)-methylpseudouridine (m1Ψ) or 5-methyluridine (m5U) for uridine), and codon optimization, has a synergistic influence on the stability of RNA (preferably mRNA) and increase in translation efficiency. Thus, in some embodiments, the mRNA used in the present disclosure contains a combination of at least two, at least three, at least four or all five of the above-mentioned modifications, i.e., (i) incorporation of a 5′-cap structure, (ii) incorporation of a poly-A sequence, unmasking of a poly-A sequence; (iii) alteration of the 5′- and/or 3′-UTR (such as incorporation of one or more 3′-UTRs); (iv) replacing one or more naturally occurring nucleotides with synthetic nucleotides (e.g., 5-methylcytidine for cytidine and/or pseudouridine (Ψ) or N(1)-methylpseudouridine (m1Ψ) or 5-methyluridine (m5U) for uridine), and (v) codon optimization. Some aspects of the disclosure involve the targeted delivery of the mRNA disclosed herein to certain cells or tissues. In some embodiments, 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 mRNA administered is mRNA encoding an antigen or epitope for inducing an immune response. In some embodiments, the target cell is a spleen cell. In some embodiments, the target cell is an antigen presenting cell such as a professional antigen presenting cell in the spleen. In some embodiments, the target cell is a dendritic cell in the spleen.
[0331] 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 naive lymphocytes and initiate an adaptive immune response.
[0332] Lipid-based mRNA delivery systems have an inherent preference to the liver. Liver accumulation is caused by the discontinuous nature of the hepatic vasculature or the lipid metabolism (liposomes and lipid or cholesterol conjugates). In some embodiments, the target organ is liver and the target tissue is liver tissue. The delivery to such target tissue is preferred, in particular, if presence of mRNA or of the encoded peptide or polypeptide in this organ or tissue is desired and/or if it is desired to express large amounts of the encoded peptide or polypeptide and/or if systemic presence of the encoded peptide or polypeptide, in particular in significant amounts, is desired or required.
[0333] In some embodiments, after administration of the mRNA particles described herein, at least a portion of the mRNA is delivered to a target cell or target organ. In some embodiments, at least a portion of the mRNA is delivered to the cytosol of the target cell. In some embodiments, the mRNA is mRNA encoding a peptide or polypeptide and the mRNA is translated by the target cell to produce the peptide or polypeptide. In some embodiments, the target cell is a cell in the liver. In some embodiments, the target cell is a muscle cell. In some embodiments, the target cell is an endothelial cell. In some embodiments the target cell is a tumor cell or a cell in the tumor microenvironment. In some embodiments, the target cell is a blood cell. In some embodiments, the target cell is a cell in the lymph nodes. In some embodiments, the target cell is a cell in the lung. In some embodiments, the target cell is a blood cell. In some embodiments, the target cell is a cell in the skin. In some embodiments, the target cell is a spleen cell. In some embodiments, the target cell is an antigen presenting cell such as a professional antigen presenting cell in the spleen. In some embodiments, the target cell is a dendritic cell in the spleen. In some embodiments, the target cell is a T cell. In some embodiments, the target cell is a B cell. In some embodiments, the target cell is a NK cell. In some embodiments, the target cell is a monocyte. Thus, RNA particles described herein may be used for delivering mRNA to such target cell.
Pharmaceutically Active Peptides or Polypeptides
[0334] “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 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.
[0335] In some embodiments, nucleic acid such as mRNA used in the present disclosure comprises a nucleic acid sequence encoding one or more peptides or polypeptides, preferably a pharmaceutically active peptide or polypeptide.
[0336] In a preferred embodiment, nucleic acid such as mRNA used in the present disclosure comprises a nucleic acid sequence encoding a peptide or polypeptide, preferably a pharmaceutically active peptide or polypeptide, and is capable of

expressing said peptide or polypeptide, in particular if transferred into a cell or subject. Thus, in some embodiments, the nucleic acid used in the present disclosure contains a coding region (open reading frame (ORF)) encoding a peptide or polypeptide, e.g., encoding a pharmaceutically active peptide or polypeptide. In this respect, an “open reading frame” or “ORF” is a continuous stretch of codons beginning with a start codon and ending with a stop codon. Such nucleic acid encoding a pharmaceutically active peptide or polypeptide is also referred to herein as “pharmaceutically active nucleic acid”. In particular, such mRNA encoding a pharmaceutically active peptide or polypeptide is also referred to herein as “pharmaceutically active mRNA”.

[0337] According to the present disclosure, the term “pharmaceutically active peptide or polypeptide” means a peptide or polypeptide that can be used in the treatment of an individual where the expression of a peptide or polypeptide would be of benefit, e.g., in ameliorating the symptoms of a disease. Preferably, a pharmaceutically active peptide or polypeptide 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. In some embodiments, a pharmaceutically active peptide or polypeptide has a positive or advantageous effect on the condition or disease state of an individual when administered to the individual in a therapeutically effective amount. A pharmaceutically active peptide or polypeptide may have prophylactic properties and may be used to delay the onset of a disease or to lessen the severity of such disease. The term “pharmaceutically active peptide or polypeptide” includes entire peptides or polypeptides, and can also refer to pharmaceutically active fragments thereof. It can also include pharmaceutically active variants and/or analogs of a peptide or polypeptide.

[0338] Specific examples of pharmaceutically active peptides and polypeptides include, but are not limited to, cytokines, hormones, adhesion molecules, immunoglobulins, immunologically active compounds, growth factors, protease inhibitors, enzymes, receptors, apoptosis regulators, transcription factors, tumor suppressor proteins, structural proteins, reprogramming factors, genomic engineering proteins, and blood proteins.

[0339] The term “cytokines” relates to proteins which have a molecular weight of about 5 to 60 kDa and which participate in cell signaling (e.g., paracrine, endocrine, and/or autocrine signaling). In particular, when released, cytokines exert an effect on the behavior of cells around the place of their release. Examples of cytokines include lymphokines, interleukins, chemokines, interferons, and tumor necrosis factors (TNFs). According to the present disclosure, cytokines do not include hormones or growth factors. Cytokines differ from hormones in that (i) they usually act at much more variable concentrations than hormones and (ii) generally are made by a broad range of cells (nearly all nucleated cells can produce cytokines). Interferons are usually characterized by antiviral, antiproliferative and immunomodulatory activities. Interferons are proteins that alter and regulate the transcription of genes within a cell by binding to interferon receptors on the regulated cell's surface, thereby preventing viral replication within the cells. The interferons can be grouped into two types. IFN-gamma is the sole type II interferon; all others are type I interferons. Particular examples of cytokines include erythropoietin (EPO), colony stimulating factor (CSF), granulocyte colony stimulating factor (G-CSF), granulocyte-macrophage colony stimulating factor (GM-CSF), tumor necrosis factor (TNF), bone morphogenetic protein (BMP), interferon alfa (IFNα), interferon beta (IFNβ), interferon gamma (INFγ), interleukin 2 (IL-2), interleukin 4 (IL-4), interleukin 10 (IL-10), interleukin 11 (IL-11), interleukin 12 (IL-12), interleukin 15 (IL-15), and interleukin 21 (IL-21), as well as variants and derivatives thereof.

[0340] In some embodiments, a pharmaceutically active peptide or polypeptide comprises a replacement protein. In these embodiments, the present disclosure provides a method for treatment of a subject having a disorder requiring protein replacement (e.g., protein deficiency disorders) comprising administering to the subject nucleic acid as described herein encoding a replacement protein. The term “protein replacement” refers to the introduction of a protein (including functional variants thereof) into a subject having a deficiency in such protein. The term also refers to the introduction of a protein into a subject otherwise requiring or benefiting from providing a protein, e.g., suffering from protein insufficiency. The term “disorder characterized by a protein deficiency” refers to any disorder that presents with a pathology caused by absent or insufficient amounts of a protein. This term encompasses protein folding disorders, i.e., conformational disorders, that result in a biologically inactive protein product. Protein insufficiency can be involved in infectious diseases, immunosuppression, organ failure, glandular problems, radiation illness, nutritional deficiency, poisoning, or other environmental or external insults.

[0341] The term “hormones” relates to a class of signaling molecules produced by glands, wherein signaling usually includes the following steps: (i) synthesis of a hormone in a particular tissue; (ii) storage and secretion; (iii) transport of the hormone to its target; (iv) binding of the hormone by a receptor; (v) relay and amplification of the signal; and (vi) breakdown of the hormone. Hormones differ from cytokines in that (1) hormones usually act in less variable concentrations and (2) generally are made by specific kinds of cells. In some embodiments, a “hormone” is a peptide or polypeptide hormone, such as insulin, vasopressin, prolactin, adrenocorticotropic hormone (ACTH), thyroid hormone, growth hormones (such as human grown hormone or bovine somatotropin), oxytocin, atrial-natriuretic peptide (ANP), glucagon, somatostatin, cholecystokinin, gastrin, and leptins.

[0342] The term “adhesion molecules” relates to proteins which are located on the surface of a cell and which are involved in binding of the cell with other cells or with the extracellular matrix (ECM). Adhesion molecules are typically transmembrane receptors and can be classified as calcium-independent (e.g., integrins, immunoglobulin superfamily, lymphocyte homing receptors) and calcium-dependent (cadherins and selectins). Particular examples of adhesion molecules are integrins, lymphocyte homing receptors, selectins (e.g., P-selectin), and addressins.

[0343] Integrins are also involved in signal transduction. In particular, upon ligand binding, integrins modulate cell signaling pathways, e.g., pathways of transmembrane protein kinases such as receptor tyrosine kinases (RTK). Such

regulation can lead to cellular growth, division, survival, or differentiation or to apoptosis. Particular examples of integrins include: α.sub.1β.sub.1, α.sub.2β.sub.1, α.sub.3β.sub.1, α.sub.4β.sub.1, α.sub.5β.sub.1, α.sub.6β.sub.1, α.sub.7β.sub.1, α.sub.Lβ.sub.2, α.sub.Mβ.sub.2, α.sub.IIbβ.sub.3, α.sub.Vβ.sub.1, α.sub.Vβ.sub.3, α.sub.Vβ.sub.5, α.sub.Vβ.sub.6, α.sub.Vβ.sub.8, and α.sub.6β.sub.4.

[0344] The term “immunoglobulins” or “immunoglobulin superfamily” refers to molecules which are involved in the recognition, binding, and/or adhesion processes of cells. Molecules belonging to this superfamily share the feature that they contain a region known as immunoglobulin domain or fold. Members of the immunoglobulin superfamily include antibodies (e.g., IgG), T cell receptors (TCRs), major histocompatibility complex (MHC) molecules, co-receptors (e.g., CD4, CD8, CD19), antigen receptor accessory molecules (e.g., CD-3γ, CD3-δ, CD-3ε, CD79a, CD79b), co-stimulatory or inhibitory molecules (e.g., CD28, CD80, CD86), and other.

[0345] The term “immunologically active compound” relates to any compound altering an immune response, e.g., by inducing and/or suppressing maturation of immune cells, inducing and/or suppressing cytokine biosynthesis, and/or altering humoral immunity by stimulating antibody production by B cells. Immunologically active compounds possess potent immunostimulating activity including, but not limited to, antiviral and antitumor activity, and can also down-regulate other aspects of the immune response, for example shifting the immune response away from a TH2 immune response, which is useful for treating a wide range of TH2 mediated diseases. Immunologically active compounds can be useful as vaccine adjuvants. Particular examples of immunologically active compounds include interleukins, colony stimulating factor (CSF), granulocyte colony stimulating factor (G-CSF), granulocyte-macrophage colony stimulating factor (GM-CSF), erythropoietin, tumor necrosis factor (TNF), interferons, integrins, addressins, selectins, homing receptors, and antigens, in particular tumor-associated antigens, pathogen-associated antigens (such as bacterial, parasitic, or viral antigens), allergens, and autoantigens. An immunologically active compound may be a vaccine antigen, i.e., an antigen whose inoculation into a subject induces an immune response.

[0346] An “antigen” according to the present disclosure covers any substance that will elicit an immune response and/or any substance against which an immune response or an immune mechanism such as a cellular response and/or humoral response is directed. This also includes situations wherein the antigen is processed into antigen peptides and an immune response or an immune mechanism is directed against one or more antigen peptides, in particular if presented in the context of MHC molecules. In particular, an “antigen” relates to any substance, such as a peptide or polypeptide, that reacts specifically with antibodies or T-lymphocytes (T-cells). The term “antigen” may comprise a molecule that comprises at least one epitope, such as a T cell epitope. In some embodiments, an antigen is a molecule which, optionally after processing, induces an immune reaction, which may be specific for the antigen (including cells expressing the antigen). In some embodiments, an antigen is a disease-associated antigen, such as a tumor antigen, a viral antigen, or a bacterial antigen, or an epitope derived from such antigen.

[0347] The term “autoantigen” or “self-antigen” refers to an antigen which originates from within the body of a subject (i.e., the autoantigen can also be called “autologous antigen”) and which produces an abnormally vigorous immune response against this normal part of the body. Such vigorous immune reactions against autoantigens may be the cause of “autoimmune diseases”. According to the present disclosure, any suitable antigen may be used, which is a candidate for an immune response, wherein the immune response may be both a humoral as well as a cellular immune response. In the context of some embodiments of the present disclosure, the antigen is presented by a cell, such as by an antigen presenting cell, in the context of MHC molecules, which results in an immune response against the antigen. An antigen may be a product which corresponds to or is derived from a naturally occurring antigen. Such naturally occurring antigens may include or may be derived from allergens, viruses, bacteria, fungi, parasites and other infectious agents and pathogens or an antigen may also be a tumor antigen. According to the present disclosure, an antigen may correspond to a naturally occurring product, for example, a viral protein, or a part thereof.

[0348] The term “disease-associated antigen” is used in its broadest sense to refer to any antigen associated with a disease. A disease-associated antigen is a molecule which contains epitopes that will stimulate a host's immune system to make a cellular antigen-specific immune response and/or a humoral antibody response against the disease. Disease-associated antigens include pathogen-associated antigens, i.e., antigens which are associated with infection by microbes, typically microbial antigens (such as bacterial or viral antigens), or antigens associated with cancer, typically tumors, such as tumor antigens.

[0349] In some embodiments, the antigen is a tumor antigen, i.e., a part of a tumor cell, in particular those which primarily occur intracellularly or as surface antigens of tumor cells. In another embodiment, the antigen is a pathogen-associated antigen, i.e., an antigen derived from a pathogen, e.g., from a virus, bacterium, unicellular organism, or parasite, for example a viral antigen such as viral ribonucleoprotein or coat protein. In some embodiments, the antigen should be presented by MHC molecules which results in modulation, in particular activation of cells of the immune system, such as CD4+ and CD8+ lymphocytes, in particular via the modulation of the activity of a T-cell receptor.

[0350] The term “tumor antigen” refers to a constituent of cancer cells which may be derived from the cytoplasm, the cell surface or the cell nucleus. In particular, it refers to those antigens which are produced intracellularly or as surface antigens on tumor cells. For example, tumor antigens include the carcinoembryonal antigen, α1-fetoprotein, isoferritin, and fetal sulphoglycoprotein, α2-H-ferroprotein and γ-fetoprotein, as well as various virus tumor antigens. According to some embodiments of the present disclosure, a tumor antigen comprises any antigen which is characteristic for tumors or cancers as well as for tumor or cancer cells with respect to type and/or expression level.

[0351] 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 a viral ribonucleoprotein or an envelope protein.
[0352] The term “bacterial antigen” refers to any bacterial component having antigenic properties, i.e. being able to provoke an immune response in an individual. The bacterial antigen may be derived from the cell wall or cytoplasm membrane of the bacterium.
[0353] The term “epitope” refers to an antigenic determinant in a molecule such as an antigen, i.e., to a part in or fragment of the molecule that is recognized by the immune system, for example, that is recognized by antibodies, T cells or B cells, in particular when presented in the context of MHC molecules. An epitope of a protein may comprises a continuous or discontinuous portion of said protein and, e.g., may be between about 5 and about 100, between about 5 and about 50, between about 8 and about 30, or about 10 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 some embodiments, the epitope in the context of the present disclosure is a T cell epitope.
[0354] Terms such as “epitope”, “fragment of an antigen”, “immunogenic peptide” and “antigen peptide” are used interchangeably herein and, e.g., may relate to an incomplete representation of an antigen which is, e.g., capable of eliciting an immune response against the antigen or a cell expressing or comprising and presenting the antigen. In some embodiments, the terms relate to an immunogenic portion of an antigen. In some embodiments, it is a portion of an antigen that is recognized (i.e., specifically bound) by a T cell receptor, in particular if presented in the context of MHC molecules. Certain preferred immunogenic portions bind to an MHC class I or class II molecule. 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, between about 8 and about 30, or between about 8 and about 25 amino acids in length, for example, the epitope may be 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids in length. In some embodiments, an epitope is between about 10 and about 25 amino acids in length. The term “epitope” includes T cell epitopes.
[0355] 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 about 10 amino acids long although longer or shorter peptides may be effective. In the case of class 11 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.
[0356] The peptide and polypeptide antigen can be 2 to 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.
[0357] The peptide or polypeptide antigen can be any peptide or polypeptide that can induce or increase the ability of the immune system to develop antibodies and T cell responses to the peptide or polypeptide.
[0358] In some embodiments, vaccine antigen, i.e., an antigen whose inoculation into a subject induces an immune response, is recognized by an immune effector cell. In some embodiments, 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 disclosure, the vaccine antigen may be, e.g., presented or present on the surface of a cell, such as an antigen presenting cell. In some embodiments, an antigen is presented by a diseased cell (such as tumor cell or an infected cell). In some embodiments, an antigen receptor is a TCR which binds to an epitope of an antigen presented in the context of MHC. In some embodiments, 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 some embodiments, 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 release cytotoxic factors, e.g., perforins and granzymes.
[0359] According to some embodiments, an amino acid sequence enhancing antigen processing and/or presentation is fused, either directly or through a linker, to an antigenic peptide or polypeptide.
[0360] Accordingly, in some embodiments, the nucleic acid (such RNA and/or DNA) described herein comprises at least one coding region encoding an antigenic peptide or polypeptide and an amino acid sequence enhancing antigen processing and/or presentation.
[0361] In some embodiments, antigen for vaccination which may be administered in the form of nucleic acid coding therefor comprises a naturally occurring antigen or a fragment such as an epitope thereof.
[0362] Such amino acid sequences enhancing antigen processing and/or presentation are preferably located at the C-terminus of the antigenic peptide or polypeptide (and optionally at the C-terminus of an amino acid sequence which breaks immunological tolerance), without being limited thereto. Amino acid sequences enhancing antigen processing and/or

presentation as defined herein preferably improve antigen processing and presentation. In one embodiment, the amino acid sequence enhancing antigen processing and/or presentation as defined herein includes, without being limited thereto, sequences derived from the human MHC class I complex (HLA-B51, haplotype A2, B27/B51, Cw2/Cw3), in particular a sequence comprising the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof. Besides improving antigen processing and presentation such amino acid sequence enhancing antigen processing and/or presentation may also be used for determining expression of an amino acid sequence in the processes described herein.

[0363] In one embodiment, an amino acid sequence enhancing antigen processing and/or presentation comprises the amino acid sequence of SEQ ID NO: 2, 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: 2, or a functional fragment of the amino acid sequence of SEQ ID NO: 2, 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: 2. In one embodiment, an amino acid sequence enhancing antigen processing and/or presentation comprises the amino acid sequence of SEQ ID NO: 2.

[0364] Accordingly, in particularly preferred embodiments, the RNA described herein comprises at least one coding region encoding an antigenic peptide or polypeptide and an amino acid sequence enhancing antigen processing and/or presentation, said amino acid sequence enhancing antigen processing and/or presentation preferably being fused to the antigenic peptide or polypeptide, more preferably to the C-terminus of the antigenic peptide or polypeptide as described herein.

[0365] Furthermore, a secretory sequence, e.g., a sequence comprising the amino acid sequence of SEQ ID NO: 1, may be fused to the N-terminus of the antigenic peptide or polypeptide.

[0366] Amino acid sequences derived from tetanus toxoid of *Clostridium tetani* may be employed to overcome self-tolerance mechanisms in order to efficiently mount an immune response to self-antigens by providing T-cell help during priming.

[0367] It is known that tetanus toxoid heavy chain includes epitopes that can bind promiscuously to MHC class II alleles and induce CD4.sup.+ memory T cells in almost all tetanus vaccinated individuals. In addition, the combination of tetanus toxoid (TT) helper epitopes with tumor-associated antigens is known to improve the immune stimulation compared to application of tumor-associated antigen alone by providing CD4.sup.+-mediated T-cell help during priming. To reduce the risk of stimulating CD8.sup.+ T cells with the tetanus sequences which might compete with the intended induction of tumor antigen-specific T-cell response, not the whole fragment C of tetanus toxoid is used as it is known to contain CD8.sup.+ T-cell epitopes. Two peptide sequences containing promiscuously binding helper epitopes were selected alternatively to ensure binding to as many MHC class II alleles as possible. Based on the data of the ex vivo studies the well-known epitopes p2 (QYIKANSKFIGITEL; TT.sub.830-844) and p16 (MTNSVDDALINSTKIYSYFPSVISKVNQGAQG; TT.sub.578-609) were selected. The p2 epitope was already used for peptide vaccination in clinical trials to boost anti-melanoma activity.

[0368] Non-clinical data showed that RNA vaccines encoding both a tumor antigen plus promiscuously binding tetanus toxoid sequences lead to enhanced CD8.sup.+ T-cell responses directed against the tumor antigen and improved break of tolerance. Immunomonitoring data from patients vaccinated with vaccines including those sequences fused in frame with the tumor antigen-specific sequences reveal that the tetanus sequences chosen are able to induce tetanus-specific T-cell responses in almost all patients.

[0369] According to some embodiments, an amino acid sequence which breaks immunological tolerance is fused, either directly or through a linker, e.g., a linker having the amino acid sequence according to SEQ ID NO: 4, to the antigenic peptide or polypeptide.

[0370] Such amino acid sequences which break immunological tolerance are preferably located at the C-terminus of the antigenic peptide or polypeptide (and optionally at the N-terminus of the amino acid sequence enhancing antigen processing and/or presentation, wherein the amino acid sequence which breaks immunological tolerance and the amino acid sequence enhancing antigen processing and/or presentation may be fused either directly or through a linker, e.g., a linker having the amino acid sequence according to SEQ ID NO: 5), without being limited thereto. Amino acid sequences which break immunological tolerance as defined herein preferably improve T cell responses. In one embodiment, the amino acid sequence which breaks immunological tolerance as defined herein includes, without being limited thereto, sequences derived from tetanus toxoid-derived helper sequences p2 and p16 (P2P16), in particular a sequence comprising the amino acid sequence of SEQ ID NO: 3 or a functional variant thereof.

[0371] In some embodiments, an amino acid sequence which breaks immunological tolerance comprises the amino acid sequence of SEQ ID NO: 3, 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 a functional fragment of the amino acid sequence of SEQ ID NO: 3, 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: 3. In one embodiment, an amino acid sequence which breaks immunological tolerance comprises the amino acid sequence of SEQ ID NO: 3.

[0372] According to some embodiments, an amino acid sequence which produces bioluminescence is fused, either directly or through a linker, e.g., a linker having the amino acid sequence according to SEQ ID NO: 4, to the antigenic peptide or polypeptide.

[0373] Such amino acid sequences which produces bioluminescence are preferably located at the C-terminus of the antigenic peptide or polypeptide (and optionally at the N-terminus of (i) the amino acid sequence enhancing antigen processing and/or presentation or (ii) the amino acid sequence which breaks immunological tolerance, wherein the amino

acid sequence which produces bioluminescence and (i) the amino acid sequence enhancing antigen processing and/or presentation or (ii) the amino acid sequence which breaks immunological tolerance may be fused either directly or through a linker, e.g., a linker having the amino acid sequence according to SEQ ID NO: 5), without being limited thereto. Amino acid sequences which produce bioluminescence as defined herein preferably improve the determination of the amount of the antigenic peptide or polypeptide. In some embodiments, the amino acid sequence which produces bioluminescence as defined herein produces fluorescence. In some embodiments, the amino acid sequence which produces bioluminescence as defined herein includes, without being limited thereto, sequences derived from Green Fluorescent Protein (GFP), Yellow Fluorescent Protein (YFP), Red Fluorescent Protein (RFP), Blue Fluorescent Protein (EBFP), Cyan Fluorescent Protein (ECFP), their variants (such as enhanced GFP (EGFP), Superfolder GFP (sfGFP), and luciferase.

[0374] In the following, embodiments of vaccine RNAs are described, wherein certain terms used when describing elements thereof have the following meanings:

[0375] hAg-Kozak: 5′-UTR sequence of the human alpha-globin mRNA with an optimized ‘Kozak sequence’ to increase translational efficiency.

[0376] sec/MITD: Fusion-protein tags derived from the sequence encoding the human MHC class I complex (HLA-B51, haplotype A2, B27/B51, Cw2/Cw3), which have been shown to improve antigen processing and presentation. Sec corresponds to the 78 bp fragment coding for the secretory signal peptide, which guides translocation of the nascent polypeptide chain into the endoplasmatic reticulum. MITD corresponds to the transmembrane and cytoplasmic domain of the MHC class I molecule, also called MHC class I trafficking domain.

[0377] Antigen: Sequences encoding the respective antigen/epitope.

[0378] 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.

[0379] P2P16: Sequence coding for tetanus toxoid-derived helper epitopes to break immunological tolerance.

[0380] 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.

[0381] 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.

[0382] In one embodiment, vaccine RNA described herein has the structure: [0383] beta-S-ARCA(D1)-hAg-Kozak-sec-GS(1)-Antigen-GS(2)-P2P16-GS(3)-MITD-FI-A30L70

[0384] In one embodiment, vaccine antigen described herein has the structure: [0385] sec-GS(1)-Antigen-GS(2)-P2P16-GS(3)-MITD

[0386] In one embodiment, hAg-Kozak comprises the nucleotide sequence of SEQ ID NO: 6. In one embodiment, sec comprises the amino acid sequence of SEQ ID NO: 1. In one embodiment, P2P16 comprises the amino acid sequence of SEQ ID NO: 3. In one embodiment, MITD comprises the amino acid sequence of SEQ ID NO: 2. In one embodiment, GS(1) comprises the amino acid sequence of SEQ ID NO: 4. In one embodiment, GS(2) comprises the amino acid sequence of SEQ ID NO: 4. In one embodiment, GS(3) comprises the amino acid sequence of SEQ ID NO: 5. In one embodiment, FI comprises the nucleotide sequence of SEQ ID NO: 7. In one embodiment, A30L70 comprises the nucleotide sequence of SEQ ID NO: 8.

[0387] In some embodiments, an antigen receptor is an antibody or B cell receptor which binds to an epitope in an antigen. In some embodiments, an antibody or B cell receptor binds to native epitopes of an antigen.

[0388] 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 may be, e.g., at least 4, at least 8, pat least 12, or 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 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.

[0389] The term “extracellular portion” or “exodomain” in the context of the present disclosure 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. In some embodiments, the term refers to one or more extracellular loops or domains or a fragment thereof.

[0390] 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, in particular when presented on the surface of antigen presenting cells or diseased cells such as cancer cells in the context of MHC molecules and preferably exerts effector functions of T cells. T cells are considered to be specific for antigen if the cells kill target cells expressing an

antigen. T cell specificity may be evaluated using any of a variety of standard techniques, for example, within a chromium release assay or proliferation assay. Alternatively, synthesis of lymphokines (such as interferon-γ) can be measured.
[0391] The term “target” shall mean an agent such as a cell or tissue which is a target for an immune response such as a cellular immune response. Targets include cells that present an antigen or an antigen epitope, i.e., a peptide fragment derived from an antigen. In some embodiments, the target cell is a cell expressing an antigen and presenting said antigen with class I MHC. “Antigen processing” refers to the degradation of an antigen into processing products which are fragments of said antigen (e.g., the degradation of a polypeptide 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.
[0392] By “antigen-responsive CTL” is meant a CD8.sup.+ T-cell that is responsive to an antigen or a peptide derived from said antigen, which is presented with class I MHC on the surface of antigen presenting cells.
[0393] According to the disclosure, CTL responsiveness may include sustained calcium flux, cell division, production of cytokines such as IFN-γ and TNF-α, up-regulation of activation markers such as CD44 and CD69, and specific cytolytic killing of tumor antigen expressing target cells. CTL responsiveness may also be determined using an artificial reporter that accurately indicates CTL responsiveness.
[0394] “Activation” or “stimulation”, as used herein, refers to the state of a cell that has been sufficiently stimulated to induce detectable cellular proliferation, such as an immune effector cell such as T cell. 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.
[0395] 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.
[0396] The term “expansion” refers to a process wherein a specific entity is multiplied. In some embodiments, the term is 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. In some embodiments, expansion leads to differentiation of the immune effector cells.
[0397] The terms “immune response” and “immune reaction” are used herein interchangeably in their conventional meaning and refer to an integrated bodily response to an antigen and may refer to a cellular immune response, a humoral immune response, or both. According to the disclosure, the term “immune response to” or “immune response against” with respect to an agent such as an antigen, cell or tissue, relates to an immune response such as a cellular response directed against the agent. An immune response may comprise one or more reactions selected from the group consisting of developing antibodies against one or more antigens and expansion of antigen-specific T-lymphocytes, such as CD4.sup.+ and CD8.sup.+ T-lymphocytes, e.g. CD8.sup.+ T-lymphocytes, which may be detected in various proliferation or cytokine production tests in vitro.
[0398] The terms “inducing an immune response” and “eliciting an immune response” and similar terms in the context of the present disclosure refer to the induction of an immune response, such as the induction of a cellular immune response, a humoral immune response, or both. The immune response may be protective/preventive/prophylactic and/or therapeutic. The immune response may be directed against any immunogen or antigen or antigen peptide, such as against a tumor-associated antigen or a pathogen-associated antigen (e.g., an antigen of a virus (such as influenza virus (A, B, or C), CMV or RSV)). “Inducing” in this context may mean that there was no immune response against a particular antigen or pathogen before induction, but it may also mean that there was a certain level of immune response against a particular antigen or pathogen before induction and after induction said immune response is enhanced. Thus, “inducing the immune response” in this context also includes “enhancing the immune response”. In some embodiments, after inducing an immune response in an individual, said individual is protected from developing a disease such as an infectious disease or a cancerous disease or the disease condition is ameliorated by inducing an immune response.
[0399] The terms “cellular immune response”, “cellular response”, “cell-mediated immunity” or similar terms are meant to include a cellular response directed to cells characterized by expression of an antigen and/or presentation of an antigen with class I or class II MHC. The cellular response relates to cells called T cells or T lymphocytes which act as either “helpers” or “killers”. The helper T cells (also termed CD4.sup.+ T cells) play a central role by regulating the immune response and the killer cells (also termed cytotoxic T cells, cytolytic T cells, CD8.sup.+ T cells or CTLs) kill cells such as diseased cells.
[0400] The term “humoral immune response” refers to a process in living organisms wherein antibodies are produced in response to agents and organisms, which they ultimately neutralize and/or eliminate. The specificity of the antibody response is mediated by T and/or B cells through membrane-associated receptors that bind antigen of a single specificity. Following binding of an appropriate antigen and receipt of various other activating signals, B lymphocytes divide, which produces memory B cells as well as antibody secreting plasma cell clones, each producing antibodies that recognize the identical antigenic epitope as was recognized by its antigen receptor. Memory B lymphocytes remain dormant until they are subsequently activated by their specific antigen. These lymphocytes provide the cellular basis of memory and the resulting escalation in antibody response when re-exposed to a specific antigen.
[0401] The term “antibody” as used herein, refers to an immunoglobulin molecule, which is able to specifically bind to an epitope on an antigen. In particular, the term “antibody” refers to a glycoprotein comprising at least two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds. The term “antibody” includes monoclonal antibodies, recombinant antibodies, human antibodies, humanized antibodies, chimeric antibodies and combinations of any of the foregoing. Each heavy chain is comprised of a heavy chain variable region (VH) and a heavy chain constant region (CH).

Each light chain is comprised of a light chain variable region (VL) and a light chain constant region (CL). The variable regions and constant regions are also referred to herein as variable domains and constant domains, respectively. The VH and VL regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDRs), interspersed with regions that are more conserved, termed framework regions (FRs). 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 CDRs of a VH are termed HCDR1, HCDR2 and HCDR3, the CDRs of a VL are termed LCDR1, LCDR2 and LCDR3. The variable regions of the heavy and light chains contain a binding domain that interacts with an antigen. The constant regions of an antibody comprise the heavy chain constant region (CH) and the light chain constant region (CL), wherein CH can be further subdivided into constant domain CH1, a hinge region, and constant domains CH2 and CH3 (arranged from amino-terminus to carboxy-terminus in the following order: CH1, CH2, CH3). 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 (C1q) of the classical complement system. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoactive portions of intact immunoglobulins. Antibodies are typically tetramers of immunoglobulin molecules. Antibodies may exist in a variety of forms including, for example, polyclonal antibodies, monoclonal antibodies, Fv, Fab and F(ab).sub.2, as well as single chain antibodies and humanized antibodies.

[0402] The term “immunoglobulin” relates to proteins of the immunoglobulin superfamily, such as to antigen receptors such as antibodies or the B cell receptor (BCR). The immunoglobulins are characterized by a structural domain, i.e., the immunoglobulin domain, having a characteristic immunoglobulin (Ig) fold. The term encompasses membrane bound immunoglobulins as well as soluble immunoglobulins. Membrane bound immunoglobulins are also termed surface immunoglobulins or membrane immunoglobulins, which are generally part of the BCR. Soluble immunoglobulins are generally termed antibodies. Immunoglobulins generally comprise several chains, typically two identical heavy chains and two identical light chains which are linked via disulfide bonds. These chains are primarily composed of immunoglobulin domains, such as the V.sub.L (variable light chain) domain, C.sub.L (constant light chain) domain, V.sub.H (variable heavy chain) domain, and the C.sub.H (constant heavy chain) domains C.sub.H1, C.sub.H2, C.sub.H3, and C.sub.H4. There are five types of mammalian immunoglobulin heavy chains, i.e., α, δ, ε, γ, and μ which account for the different classes of antibodies, i.e., IgA, IgD, IgE, IgG, and IgM. As opposed to the heavy chains of soluble immunoglobulins, the heavy chains of membrane or surface immunoglobulins comprise a transmembrane domain and a short cytoplasmic domain at their carboxy-terminus. In mammals there are two types of light chains, i.e., lambda and kappa. The immunoglobulin chains comprise a variable region and a constant region. The constant region is essentially conserved within the different isotypes of the immunoglobulins, wherein the variable part is highly divers and accounts for antigen recognition.

[0403] The terms “vaccination” and “immunization” describe the process of treating an individual for therapeutic or prophylactic reasons and relate to the procedure of administering one or more immunogen(s) or antigen(s) or derivatives thereof, in particular in the form of RNA (especially mRNA) coding therefor, as described herein to an individual and stimulating an immune response against said one or more immunogen(s) or antigen(s) or cells characterized by presentation of said one or more immunogen(s) or antigen(s).

[0404] By “cell characterized by presentation of an antigen” or “cell presenting an antigen” or “MHC molecules which present an antigen on the surface of an antigen presenting cell” or similar expressions is meant a cell such as a diseased cell, in particular a tumor cell or an infected cell, or an antigen presenting cell presenting the antigen or an antigen peptide, either directly or following processing, in the context of MHC molecules, such as MHC class I and/or MHC class II molecules. In some embodiments, the MHC molecules are MHC class I molecules.

[0405] The term “allergen” refers to a kind of antigen which originates from outside the body of a subject (i.e., the allergen can also be called “heterologous antigen”) and which produces an abnormally vigorous immune response in which the immune system of the subject fights off a perceived threat that would otherwise be harmless to the subject. “Allergies” are the diseases caused by such vigorous immune reactions against allergens. An allergen usually is an antigen which is able to stimulate a type-I hypersensitivity reaction in atopic individuals through immunoglobulin E (IgE) responses. Particular examples of allergens include allergens derived from peanut proteins (e.g., Ara h 2.02), ovalbumin, grass pollen proteins (e.g., Phl p 5), and proteins of dust mites (e.g., Der p 2).

[0406] The term “growth factors” refers to molecules which are able to stimulate cellular growth, proliferation, healing, and/or cellular differentiation. Typically, growth factors act as signaling molecules between cells. The term “growth factors” include particular cytokines and hormones which bind to specific receptors on the surface of their target cells. Examples of growth factors include bone morphogenetic proteins (BMPs), fibroblast growth factors (FGFs), vascular endothelial growth factors (VEGFs), such as VEGFA, epidermal growth factor (EGF), insulin-like growth factor, ephrins, macrophage colony-stimulating factor, granulocyte colony-stimulating factor, granulocyte macrophage colony-stimulating factor, neuregulins, neurotrophins (e.g., brain-derived neurotrophic factor (BDNF), nerve growth factor (NGF)), placental growth factor (PGF), platelet-derived growth factor (PDGF), renalase (RNLS) (anti-apoptotic survival factor), T-cell growth factor (TCGF), thrombopoietin (TPO), transforming growth factors (transforming growth factor alpha (TGF-α), transforming growth factor beta (TGF-β)), and tumor necrosis factor-alpha (TNF-α). In some embodiments, a “growth factor” is a peptide or polypeptide growth factor.

[0407] The term “protease inhibitors” refers to molecules, in particular peptides or polypeptides, which inhibit the function of proteases. Protease inhibitors can be classified by the protease which is inhibited (e.g., aspartic protease inhibitors) or by their mechanism of action (e.g., suicide inhibitors, such as serpins). Particular examples of protease inhibitors include

serpins, such as alpha 1-antitrypsin, aprotinin, and bestatin.
[0408] The term “enzymes” refers to macromolecular biological catalysts which accelerate chemical reactions. Like any catalyst, enzymes are not consumed in the reaction they catalyze and do not alter the equilibrium of said reaction. Unlike many other catalysts, enzymes are much more specific. In some embodiments, an enzyme is essential for homeostasis of a subject, e.g., any malfunction (in particular, decreased activity which may be caused by any of mutation, deletion or decreased production) of the enzyme results in a disease. Examples of enzymes include herpes simplex virus type 1 thymidine kinase (HSV1-TK), hexosaminidase, phenylalanine hydroxylase, pseudocholinesterase, and lactase.
[0409] The term “receptors” refers to protein molecules which receive signals (in particular chemical signals called ligands) from outside a cell. The binding of a signal (e.g., ligand) to a receptor causes some kind of response of the cell, e.g., the intracellular activation of a kinase. Receptors include transmembrane receptors (such as ion channel-linked (ionotropic) receptors, G protein-linked (metabotropic) receptors, and enzyme-linked receptors) and intracellular receptors (such as cytoplasmic receptors and nuclear receptors). Particular examples of receptors include steroid hormone receptors, growth factor receptors, and peptide receptors (i.e., receptors whose ligands are peptides), such as P-selectin glycoprotein ligand-1 (PSGL-1). The term “growth factor receptors” refers to receptors which bind to growth factors.
[0410] The term “apoptosis regulators” refers to molecules, in particular peptides or polypeptides, which modulate apoptosis, i.e., which either activate or inhibit apoptosis. Apoptosis regulators can be grouped into two broad classes: those which modulate mitochondrial function and those which regulate caspases. The first class includes proteins (e.g., BCL-2, BCL-xL) which act to preserve mitochondrial integrity by preventing loss of mitochondrial membrane potential and/or release of pro-apoptotic proteins such as cytochrome C into the cytosol. Also to this first class belong proapoptotic proteins (e.g., BAX, BAK, BIM) which promote release of cytochrome C. The second class includes proteins such as the inhibitors of apoptosis proteins (e.g., XIAP) or FLIP which block the activation of caspases.
[0411] The term “transcription factors” relates to proteins which regulate the rate of transcription of genetic information from DNA to messenger RNA, in particular by binding to a specific DNA sequence. Transcription factors may regulate cell division, cell growth, and cell death throughout life; cell migration and organization during embryonic development; and/or in response to signals from outside the cell, such as a hormone. Transcription factors contain at least one DNA-binding domain which binds to a specific DNA sequence, usually adjacent to the genes which are regulated by the transcription factors. Particular examples of transcription factors include MECP2, FOXP2, FOXP3, the STAT protein family, and the HOX protein family.
[0412] The term “tumor suppressor proteins” relates to molecules, in particular peptides or polypeptides, which protect a cell from one step on the path to cancer. Tumor-suppressor proteins (usually encoded by corresponding tumor-suppressor genes) exhibit a weakening or repressive effect on the regulation of the cell cycle and/or promote apoptosis. Their functions may be one or more of the following: repression of genes essential for the continuing of the cell cycle; coupling the cell cycle to DNA damage (as long as damaged DNA is present in a cell, no cell division should take place); initiation of apoptosis, if the damaged DNA cannot be repaired; metastasis suppression (e.g., preventing tumor cells from dispersing, blocking loss of contact inhibition, and inhibiting metastasis); and DNA repair. Particular examples of tumor-suppressor proteins include p53, phosphatase and tensin homolog (PTEN), SWI/SNF (SWltch/Sucrose Non-Fermentable), von Hippel-Lindau tumor suppressor (pVHL), adenomatous polyposis coli (APC), CD95, suppression of tumorigenicity 5 (ST5), suppression of tumorigenicity 5 (ST5), suppression of tumorigenicity 14 (ST14), and Yippee-like 3 (YPEL3).
[0413] The term “structural proteins” refers to proteins which confer stiffness and rigidity to otherwise-fluid biological components. Structural proteins are mostly fibrous (such as collagen and elastin) but may also be globular (such as actin and tubulin). Usually, globular proteins are soluble as monomers, but polymerize to form long, fibers which, for example, may make up the cytoskeleton. Other structural proteins are motor proteins (such as myosin, kinesin, and dynein) which are capable of generating mechanical forces, and surfactant proteins. Particular examples of structural proteins include collagen, surfactant protein A, surfactant protein B, surfactant protein C, surfactant protein D, elastin, tubulin, actin, and myosin.
[0414] The term “reprogramming factors” or “reprogramming transcription factors” relates to molecules, in particular peptides or polypeptides, which, when expressed in somatic cells optionally together with further agents such as further reprogramming factors, lead to reprogramming or de-differentiation of said somatic cells to cells having stem cell characteristics, in particular pluripotency. Particular examples of reprogramming factors include OCT4, SOX2, c-MYC, KLF4, LIN28, and NANOG.
[0415] The term “genomic engineering proteins” relates to proteins which are able to insert, delete or replace DNA in the genome of a subject. Particular examples of genomic engineering proteins include meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly spaced short palindromic repeat-CRISPR-associated protein 9 (CRISPR-Cas9).
[0416] The term “blood proteins” relates to peptides or polypeptides which are present in blood plasma of a subject, in particular blood plasma of a healthy subject. Blood proteins have diverse functions such as transport (e.g., albumin, transferrin), enzymatic activity (e.g., thrombin or ceruloplasmin), blood clotting (e.g., fibrinogen), defense against pathogens (e.g., complement components and immunoglobulins), protease inhibitors (e.g., alpha 1-antitrypsin), etc. Particular examples of blood proteins include thrombin, serum albumin, Factor VII, Factor VIII, insulin, Factor IX, Factor X, tissue plasminogen activator, protein C, von Willebrand factor, antithrombin III, glucocerebrosidase, erythropoietin, granulocyte colony stimulating factor (G-CSF), modified Factor VIII, and anticoagulants.
[0417] Thus, in some embodiments, the pharmaceutically active peptide or polypeptide is (i) a cytokine, preferably selected

from the group consisting of erythropoietin (EPO), interleukin 4 (IL-2), and interleukin 10 (IL-11), more preferably EPO; (ii) an adhesion molecule, in particular an integrin; (iii) an immunoglobulin, in particular an antibody; (iv) an immunologically active compound, in particular an antigen; (v) a hormone, in particular vasopressin, insulin or growth hormone; (vi) a growth factor, in particular VEGFA; (vii) a protease inhibitor, in particular alpha 1-antitrypsin; (viii) an enzyme, preferably selected from the group consisting of herpes simplex virus type 1 thymidine kinase (HSV1-TK), hexosaminidase, phenylalanine hydroxylase, pseudocholinesterase, pancreatic enzymes, and lactase; (ix) a receptor, in particular growth factor receptors; (x) an apoptosis regulator, in particular BAX; (xi) a transcription factor, in particular FOXP3; (xii) a tumor suppressor protein, in particular p53; (xiii) a structural protein, in particular surfactant protein B; (xiv) a reprogramming factor, e.g., selected from the group consisting of OCT4, SOX2, c-MYC, KLF4, LIN28 and NANOG; (xv) a genomic engineering protein, in particular clustered regularly spaced short palindromic repeat-CRISPR-associated protein 9 (CRISPR-Cas9); and (xvi) a blood protein, in particular fibrinogen.

[0418] In some embodiments, a pharmaceutically active peptide or polypeptide comprises one or more antigens or one or more epitopes, i.e., administration of the peptide or polypeptide to a subject elicits an immune response against the one or more antigens or one or more epitopes in a subject which may be therapeutic or partially or fully protective.

[0419] In some embodiments, the nucleic acid such as mRNA encodes at least one epitope.

[0420] In some embodiments, the epitope is derived from a tumor antigen. The tumor antigen may be a “standard” antigen, which is generally known to be expressed in various cancers. The tumor antigen may also be a “neo-antigen”, which is specific to an individual's tumor and has not been previously recognized by the immune system. A neo-antigen or neo-epitope may result from one or more cancer-specific mutations in the genome of cancer cells resulting in amino acid changes. Examples of tumor antigens include, without limitation, p53, ART-4, BAGE, beta-catenin/m, Bcr-abL CAMEL, CAP-1, CASP-8, CDC27/m, CDK4/m, CEA, the cell surface proteins of the claudin family, such as CLAUD FN-6, CLAUDIN-18.2 and CLAUDIN-12, c-MYC, CT, Cyp-B, DAM, ELF2M, ETV6-AML1, G250, GAGE, GnT-V, Gap 100, HAGE, HER-2/neu, HPV-E7, HPV-E6, HAST-2, hTERT (or hTRT), LAGE, LDLR/FUT, MAGE-A, preferably MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A 10, MAGE-A 11, or MAGE-A12, MAGE-B, MAGE-C, MART-1/Melan-A, MC1R, Myosin/m, MUC1, MUM-1, MUM-2, MUM-3, NA88-A, NF1 NY-ESO-1, NY-BR-1, p190 minor BCR-abL, Pml/RARa, PRAME, proteinase 3, PSA, PSM, RAGE, RU1 or RU2, SAGE, SART-1 or SART-3, SCGB3A2, SCP1, SCP2, SCP3, SSX, SURVIVIN, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2, TPTE, WT, and WT-1.

[0421] Cancer mutations vary with each individual. Thus, cancer mutations that encode novel epitopes (neo-epitopes) represent attractive targets in the development of vaccine compositions and immunotherapies. The efficacy of tumor immunotherapy relies on the selection of cancer-specific antigens and epitopes capable of inducing a potent immune response within a host. RNA can be used to deliver patient-specific tumor epitopes to a patient. Dendritic cells (DCs) residing in the spleen represent antigen-presenting cells of particular interest for RNA expression of immunogenic epitopes or antigens such as tumor epitopes. The use of multiple epitopes has been shown to promote therapeutic efficacy in tumor vaccine compositions. Rapid sequencing of the tumor mutanome may provide multiple epitopes for individualized vaccines which can be encoded by mRNA described herein, e.g., as a single polypeptide wherein the epitopes are optionally separated by linkers. In some embodiments of the present disclosure, the mRNA encodes at least one epitope, at least two epitopes, at least three epitopes, at least four epitopes, at least five epitopes, at least six epitopes, at least seven epitopes, at least eight epitopes, at least nine epitopes, or at least ten epitopes. Exemplary embodiments include mRNA that encodes at least five epitopes (termed a “pentatope”) and mRNA that encodes at least ten epitopes (termed a “decatope”).

[0422] In some embodiments, the antigen or epitope is derived from a pathogen-associated antigen, in particular from a viral antigen. In some embodiments, the antigen or epitope is derived from a SARS-CoV-2 S protein, an immunogenic variant thereof, or an immunogenic fragment of the SARS-CoV-2 S protein or the immunogenic variant thereof. Thus, in some embodiments, the mRNA used in the present disclosure encodes an amino acid sequence comprising a SARS-CoV-2 S protein, an immunogenic variant thereof, or an immunogenic fragment of the SARS-CoV-2 S protein or the immunogenic variant thereof.

[0423] In some embodiments of the present disclosure the antigen (such as a tumor antigen or 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).

[0424] In some embodiments, beta-S-ARCA(D1) is utilized as specific capping structure at the 5′-end of the mRNA. In some embodiments, m.sub.2.sup.7,3′-OGppp(m.sub.1.sup.2′-O) ApG is utilized as specific capping structure at the 5′-end of the mRNA. In some embodiments, the 5′-UTR sequence is derived from the human alpha-globin mRNA and optionally has an optimized ‘Kozak sequence’ to increase translational efficiency. In some embodiments, 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 some embodiments, 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 some embodiments, 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.

[0425] In some embodiments, mRNA encoding an antigen (such as a tumor antigen or a vaccine antigen) is expressed in cells of the subject treated to provide the antigen. In some embodiments, the mRNA is transiently expressed in cells of the subject. In some embodiments, the mRNA is in vitro transcribed. In some embodiments, expression of the antigen is at the cell surface. In some embodiments, the antigen is expressed and presented in the context of MHC. In some embodiments, expression of the antigen is into the extracellular space, i.e., the antigen is secreted.
[0426] The 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.
[0427] A peptide and polypeptide antigen which is provided to a subject according to the present disclosure by administering mRNA encoding a peptide and polypeptide antigen, wherein the antigen is a vaccine antigen, preferably results in the induction of an immune response, e.g., a humoral and/or cellular immune response in the subject being provided the peptide or polypeptide antigen. Said immune response is preferably directed against a target antigen. Thus, a vaccine antigen may comprise the target antigen, a variant thereof, or a fragment thereof. In some embodiments, such fragment or variant is immunologically equivalent to the target antigen. In the context of the present disclosure, 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, the 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. Thus, according to the present disclosure, 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" according to the disclosure 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.
[0428] 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.
[0429] In some embodiments, the mRNA used in the present disclosure is non-immunogenic. RNA encoding an immunostimulant may be administered according to the present disclosure to provide an adjuvant effect. The RNA encoding an immunostimulant may be standard RNA or non-immunogenic RNA.
[0430] The term "non-immunogenic RNA" (such as "non-immunogenic mRNA") 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 certain embodiments, non-immunogenic RNA, which is also termed modified RNA (mod RNA) herein, is rendered non-immunogenic by incorporating modified nucleosides suppressing RNA-mediated activation of innate immune receptors into the RNA and/or removing double-stranded RNA (dsRNA).
[0431] For rendering the non-immunogenic RNA (especially mRNA) 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 some embodiments, the modified nucleosides comprise a replacement of one or more uridines with a nucleoside comprising a modified nucleobase. In some embodiments, the modified nucleobase is a modified uracil. In some embodiments, the nucleoside comprising a modified nucleobase is selected from the group consisting of 3-methyl-uridine (m.sup.3U), 5-methoxy-uridine (mo.sup.5U), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s.sup.2U), 4-thio-uridine (s.sup.4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho.sup.5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-oxyacetic acid (cmo.sup.5U), uridine 5-oxyacetic acid methyl ester (mcmo.sup.5U), 5-carboxymethyl-uridine (cm.sup.5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm.sup.5U), 5-carboxyhydroxymethyl-uridine methyl ester (mchm.sup.5U), 5-methoxycarbonylmethyl-uridine (mcm.sup.5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm.sup.5s2U), 5-aminomethyl-2-thio-uridine (nm.sup.5s2U), 5-methylaminomethyl-uridine (mnm.sup.5U), 1-ethyl-pseudouridine, 5-methylaminomethyl-2-thio-uridine (mnm.sup.5s2U), 5-methylaminomethyl-2-seleno-uridine (mnm.sup.5se.sup.2U), 5-carbamoylmethyl-uridine (ncm.sup.5U), 5-carboxymethylaminomethyl-uridine (cmnm.sup.5U), 5-carboxymethylaminomethyl-2-thio-uridine (cmnm.sup.5s.sup.2U), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (τm.sup.5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine(τm5s2U), 1-taurinomethyl-4-thio-pseudouridine), 5-methyl-2-thio-uridine (m.sup.5s2U), 1-methyl-4-thio-pseudouridine (m.sup.1s.sup.4ψ), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine (m.sup.3ψ), 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.sup.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.sup.3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp.sup.3 ψ), 5-

(isopentenylaminomethyl)uridine (inm.sup.5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm.sup.5s.sup.2U), α-thio-uridine, 2′-O-methyl-uridine (Um), 5,2′-O-dimethyl-uridine (m.sup.5Um), 2′-O-methyl-pseudouridine (ψm), 2-thio-2′-O-methyl-uridine (s.sup.2Um), 5-methoxycarbonylmethyl-2′-O-methyl-uridine (mcm.sup.5Um), 5-carbamoylmethyl-2′-O-methyl-uridine (ncm.sup.5Um), 5-carboxymethylaminomethyl-2′-O-methyl-uridine (cmnm.sup.5Um), 3,2′-O-dimethyl-uridine (m.sup.3Um), 5-(isopentenylaminomethyl)-2′-O-methyl-uridine (inm.sup.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 certain embodiments, the nucleoside comprising a modified nucleobase is pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ) or 5-methyl-uridine (m5U), in particular N1-methyl-pseudouridine.

[0432] In some embodiments, 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.

[0433] 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 some embodiments, 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. Suitable methods for providing ssRNA are disclosed, for example, in WO 2017/182524.

[0434] 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.

[0435] In some embodiments, the removal of dsRNA (especially mRNA) 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 some embodiments, the non-immunogenic RNA (especially mRNA) is free or essentially free of dsRNA. In some embodiments, the non-immunogenic RNA (especially mRNA) 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 (especially mRNA) 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.).

[0436] In some embodiments, the non-immunogenic RNA (especially mRNA) is translated in a cell more efficiently than standard RNA with the same sequence. In some embodiments, translation is enhanced by a factor of 2-fold relative to its unmodified counterpart. In some embodiments, translation is enhanced by a 3-fold factor. In some embodiments, translation is enhanced by a 4-fold factor. In some embodiments, translation is enhanced by a 5-fold factor. In some embodiments, translation is enhanced by a 6-fold factor. In some embodiments, translation is enhanced by a 7-fold factor. In some embodiments, translation is enhanced by an 8-fold factor.

[0437] In some embodiments, translation is enhanced by a 9-fold factor. In some embodiments, translation is enhanced by a 10-fold factor. In some embodiments, translation is enhanced by a 15-fold factor. In some embodiments, translation is enhanced by a 20-fold factor. In some embodiments, translation is enhanced by a 50-fold factor. In some embodiments, translation is enhanced by a 100-fold factor. In some embodiments, translation is enhanced by a 200-fold factor. In some embodiments, translation is enhanced by a 500-fold factor. In some embodiments, translation is enhanced by a 1000-fold factor. In some embodiments, translation is enhanced by a 2000-fold factor. In some embodiments, the factor is 10-1000-fold. In some embodiments, the factor is 10-100-fold. In some embodiments, the factor is 10-200-fold. In some embodiments, the factor is 10-300-fold. In some embodiments, the factor is 10-500-fold. In some embodiments, the factor is 20-1000-fold. In some embodiments, the factor is 30-1000-fold. In some embodiments, the factor is 50-1000-fold. In some embodiments, the factor is 100-1000-fold. In some embodiments, the factor is 200-1000-fold. In some embodiments, translation is enhanced by any other significant amount or range of amounts.

[0438] In some embodiments, the non-immunogenic RNA (especially mRNA) exhibits significantly less innate immunogenicity than standard RNA with the same sequence. In some embodiments, the non-immunogenic RNA (especially mRNA) exhibits an innate immune response that is 2-fold less than its unmodified counterpart. In some embodiments, innate immunogenicity is reduced by a 3-fold factor. In some embodiments, innate immunogenicity is reduced by a 4-fold factor. In some embodiments, innate immunogenicity is reduced by a 5-fold factor. In some embodiments, innate immunogenicity is reduced by a 6-fold factor. In some embodiments, innate immunogenicity is reduced by a 7-fold factor. In some embodiments, innate immunogenicity is reduced by a 8-fold factor. In some embodiments, innate immunogenicity is reduced by a 9-fold factor. In some embodiments, innate immunogenicity is

reduced by a 10-fold factor. In some embodiments, innate immunogenicity is reduced by a 15-fold factor. In some embodiments, innate immunogenicity is reduced by a 20-fold factor. In some embodiments, innate immunogenicity is reduced by a 50-fold factor. In some embodiments, innate immunogenicity is reduced by a 100-fold factor. In some embodiments, innate immunogenicity is reduced by a 200-fold factor. In some embodiments, innate immunogenicity is reduced by a 500-fold factor. In some embodiments, innate immunogenicity is reduced by a 1000-fold factor. In some embodiments, innate immunogenicity is reduced by a 2000-fold factor.

[0439] The term “exhibits significantly less innate immunogenicity” refers to a detectable decrease in innate immunogenicity. In some embodiments, the term refers to a decrease such that an effective amount of the non-immunogenic RNA (especially mRNA) can be administered without triggering a detectable innate immune response. In some embodiments, the term refers to a decrease such that the non-immunogenic RNA (especially mRNA) 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 some embodiments, the decrease is such that the non-immunogenic RNA (especially mRNA) can be repeatedly administered without eliciting an innate immune response sufficient to eliminate detectable production of the protein encoded by the non-immunogenic RNA.

[0440] “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.

Particles

[0441] Nucleic acids (such as RNA and/or DNA, in particular mRNA) described herein may be present in particles comprising (i) the nucleic acid, and (ii) at least one cationic or cationically ionizable compound such as a polymer or lipid complexing the nucleic acid. 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.

[0442] Different types of RNA containing particles have been described previously to be suitable for delivery of RNA in particulate form (cf., e.g., Kaczmarek, J. C. et al., 2017, Genome Medicine 9, 60). For non-viral RNA delivery vehicles, nanoparticle encapsulation of RNA physically protects RNA from degradation and, depending on the specific chemistry, can aid in cellular uptake and endosomal escape.

[0443] In the context of the present disclosure, the term “particle” relates to a structured entity formed by molecules or molecule complexes, in particular particle forming compounds. In some embodiments, the particle contains an envelope (e.g., one or more layers or lamellas) made of one or more types of amphiphilic substances (e.g., amphiphilic lipids). In this context, the expression “amphiphilic substance” means that the substance possesses both hydrophilic and lipophilic properties. The envelope may also comprise additional substances (e.g., additional lipids) which do not have to be amphiphilic. Thus, the particle may be a monolamellar or multilamellar structure, wherein the substances constituting the one or more layers or lamellas comprise one or more types of amphiphilic substances (in particular selected from the group consisting of amphiphilic lipids) optionally in combination with additional substances (e.g., additional lipids) which do not have to be amphiphilic. In some embodiments, the term “particle” relates to a micro- or nano-sized structure, such as a micro- or nano-sized compact structure. According to the present disclosure, the term “particle” includes nanoparticles.

[0444] An “RNA particle” can be used to deliver RNA to a target site of interest (e.g., cell, tissue, organ, and the like). An RNA particle may be formed from lipids comprising at least one cationic or cationically ionizable lipid or lipid-like material. Without intending to be bound by any theory, it is believed that the cationic or cationically ionizable lipid or lipid-like material combines together with the RNA to form aggregates, and this aggregation results in colloidally stable particles.

[0445] Nucleic acid particles (such RNA particles, DNA particles or DNA/RNA particles) described herein include lipid nanoparticle (LNP)-based and lipoplex (LPX)-based formulations.

[0446] In general, a lipoplex (LPX) is obtainable from mixing two aqueous phases, namely a phase comprising nucleic acid (such as RNA and/or DNA) and a phase comprising a dispersion of lipids. In some embodiments, the lipid phase comprises liposomes.

[0447] In some embodiments, liposomes are self-closed unilamellar or multilamellar vesicular particles wherein the lamellae comprise lipid bilayers and the encapsulated lumen comprises an aqueous phase. A prerequisite for using liposomes for nanoparticle formation is that the lipids in the mixture as required are able to form lamellar (bilayer) phases in the applied aqueous environment.

[0448] In some embodiments, liposomes comprise unilamellar or multilamellar phospholipid bilayers enclosing an aqueous core (also referred to herein as an aqueous lumen). They may be prepared from materials possessing polar head (hydrophilic) groups and nonpolar tail (hydrophobic) groups. In some embodiments, cationic lipids employed in formulating liposomes designed for the delivery of nucleic acids are amphiphilic in nature and consist of a positively charged (cationic) amine head group linked to a hydrocarbon chain or cholesterol derivative via glycerol.

[0449] In some embodiments, lipoplexes are multilamellar liposome-based formulations that form upon electrostatic interaction of cationic liposomes with nucleic acids (such as RNAs and/or DNAs). In some embodiments, formed lipoplexes possess distinct internal arrangements of molecules that arise due to the transformation from liposomal structure into compact nucleic acid-lipoplexes (such as RNA- and/or DNA-lipoplexes). In some embodiments, these formulations are characterized by their poor encapsulation of the nucleic acid (such as RNA) and incomplete entrapment of the nucleic

acid (such as RNA).
[0450] In some embodiments, an LPX particle comprises an amphiphilic lipid, in particular cationic or cationically ionizable amphiphilic lipid, and nucleic acid (such as RNA and/or DNA, especially mRNA) as described herein. In some embodiments, electrostatic interactions between positively charged liposomes (made from one or more amphiphilic lipids, in particular cationic or cationically ionizable amphiphilic lipids) and negatively charged nucleic acid (especially mRNA) results in complexation and spontaneous formation of nucleic acid lipoplex particles. Positively charged liposomes may be generally synthesized using a cationic or cationically ionizable amphiphilic lipid, such as DOTMA and/or DODMA, and additional lipids, such as DOPE. In some embodiments, a nucleic acid (such as RNA and/or DNA, especially mRNA) lipoplex particle is a nanoparticle.
[0451] In general, a lipid nanoparticle (LNP) is obtainable from direct mixing of nucleic acid (such as RNA and/or DNA) in an aqueous phase with lipids in a phase comprising an organic solvent, such as ethanol. In that case, lipids or lipid mixtures can be used for particle formation, which do not form lamellar (bilayer) phases in water.
[0452] In some embodiments, LNPs comprise or consist of a cationic/ionizable lipid and helper lipids such as phospholipids, cholesterol, and/or polyethylene glycol (PEG) lipids. In some embodiments, in the nucleic acid LNPs (such as RNA LNPs, e.g., mRNA LNPs) described herein the nucleic acid (such as RNA, e.g., mRNA) is bound by ionizable lipid that occupies the central core of the LNP. In some embodiments, PEG lipid forms the surface of the LNP, along with phospholipids. In some embodiments, the surface comprises a bilayer. In some embodiments, cholesterol and ionizable lipid in charged and uncharged forms can be distributed throughout the LNP.
[0453] In some embodiments, nucleic acid (such as RNA and/or DNA, e.g., mRNA) may be noncovalently associated with a particle as described herein. In embodiments, the nucleic acid (such as RNA and/or DNA, especially mRNA) may be adhered to the outer surface of the particle (surface nucleic acid (such as surface RNA, especially surface mRNA)) and/or may be contained in the particle (encapsulated nucleic acid (such as encapsulated RNA, especially encapsulated mRNA)). In some embodiments, the particles (e.g., LNPs and LPXs) described herein have a size (such as a diameter) in the range of about 10 to about 2000 nm, such as at least about 15 nm (e.g., at least about 20 nm, at least about 25 nm, at least about 30 nm, at least about 35 nm, at least about 40 nm, at least about 45 nm, at least about 50 nm, at least about 55 nm, at least about 60 nm, at least about 65 nm, at least about 70 nm, at least about 75 nm, at least about 80 nm, at least about 85 nm, at least about 90 nm, at least about 95 nm, or at least about 100 nm) and/or at most 1900 nm (e.g., at most about 1900 nm, at most about 1800 nm, at most about 1700 nm, at most about 1600 nm, at most about 1500 nm, at most about 1400 nm, at most about 1300 nm, at most about 1200 nm, at most about 1100 nm, at most about 1000 nm, at most about 950 nm, at most about 900 nm, at most about 850 nm, at most about 800 nm, at most about 750 nm, at most about 700 nm, at most about 650 nm, at most about 600 nm, at most about 550 nm, or at most about 500 nm), such as in the range of about 20 to about 1500 nm, such as about 30 to about 1200 nm, about 40 to about 1100 nm, about 50 to about 1000 nm, about 60 to about 900 nm, about 70 to 800 nm, about 80 to 700 nm, about 90 to 600 nm, or about 50 to 500 nm or about 100 to 500 nm, such as in the range of 10 to 1000 nm, 15 to 500 nm, 20 to 450 nm, 25 to 400 nm, 30 to 350 nm, 40 to 300 nm, 50 to 250 nm, 60 to 200 nm, or 70 to 150 nm.
[0454] In some embodiments, the particles (e.g., LNPs and LPXs) described herein have an average diameter that in some embodiments ranges from about 50 nm to about 1000 nm, from about 50 nm to about 800 nm, from about 50 nm to about 700 nm, from about 50 nm to about 600 nm, from about 50 nm to about 500 nm, from about 50 nm to about 450 nm, from about 50 nm to about 400 nm, from about 50 nm to about 350 nm, from about 50 nm to about 300 nm, from about 50 nm to about 250 nm, from about 50 nm to about 200 nm, from about 100 nm to about 1000 nm, from about 100 nm to about 800 nm, from about 100 nm to about 700 nm, from about 100 nm to about 600 nm, from about 100 nm to about 500 nm, from about 100 nm to about 450 nm, from about 100 nm to about 400 nm, from about 100 nm to about 350 nm, from about 100 nm to about 300 nm, from about 100 nm to about 250 nm, from about 100 nm to about 200 nm, from about 150 nm to about 1000 nm, from about 150 nm to about 800 nm, from about 150 nm to about 700 nm, from about 150 nm to about 600 nm, from about 150 nm to about 500 nm, from about 150 nm to about 450 nm, from about 150 nm to about 400 nm, from about 150 nm to about 350 nm, from about 150 nm to about 300 nm, from about 150 nm to about 250 nm, from about 150 nm to about 200 nm, from about 200 nm to about 1000 nm, from about 200 nm to about 800 nm, from about 200 nm to about 700 nm, from about 200 nm to about 600 nm, from about 200 nm to about 500 nm, from about 200 nm to about 450 nm, from about 200 nm to about 400 nm, from about 200 nm to about 350 nm, from about 200 nm to about 300 nm, or from about 200 nm to about 250 nm.
[0455] In some embodiments, the particles described herein are nanoparticles. The term “nanoparticle” relates to a nano-sized particle comprising nucleic acid (especially mRNA) as described herein and at least one cationic or cationically ionizable lipid, wherein all three external dimensions of the particle are in the nanoscale, i.e., at least about 1 nm and below about 1000 nm. Preferably, the size of a particle is its diameter.
[0456] Nucleic acid particles described herein (especially mRNA particles) may exhibit a polydispersity index (PDI) less than about 0.5, less than about 0.4, less than about 0.3, less than about 0.2, less than about 0.1, or less than about 0.05. By way of example, the nucleic acid particles can exhibit a polydispersity index in a range of about 0.01 to about 0.4 or about 0.1 to about 0.3.
[0457] The N/P ratio gives the ratio of the nitrogen groups in the lipid to the number of phosphate groups in the nucleic acid. 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.
[0458] Nucleic acid particles (especially RNA particles such as mRNA 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 and mixing the colloid with nucleic acid to obtain nucleic acid particles.
[0459] 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.
[0460] For the preparation of colloids comprising at least one cationic or cationically ionizable lipid 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).
[0461] 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.
[0462] 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.
[0463] 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 nucleic acid (such as RNA and/or DNA, especially mRNA) lipoplex particles described herein are obtainable by adding nucleic acid (such as RNA and/or DNA, especially mRNA) to a colloidal liposome dispersion. Using the ethanol injection technique, such colloidal liposome dispersion is, in some embodiments, formed as follows: an ethanol solution comprising lipids, such as cationic or cationically ionizable lipids like DOTMA and/or DODMA and additional lipids, is injected into an aqueous solution under stirring. In some embodiments, the nucleic acid (such as RNA and/or DNA, especially mRNA) lipoplex particles described herein are obtainable without a step of extrusion.
[0464] 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.
[0465] Other methods having organic solvent free characteristics may also be used according to the present disclosure for preparing a colloid.
[0466] In some embodiments, LNPs comprise four components: ionizable cationic lipids, neutral lipids such as phospholipids, a steroid such as cholesterol, and a polymer conjugated lipid. In some embodiments, LNPs may be prepared by mixing lipids dissolved in ethanol rapidly with nucleic acid (such as RNA and/or DNA) in an aqueous buffer. While nucleic acid (such as RNA and/or DNA) particles described herein may comprise polymer conjugated lipids such as PEG lipids, provided herein are also nucleic acid (such as RNA and/or DNA) particles which do not comprise polymer conjugated lipids such as PEG lipids.
[0467] In some embodiments, the LNPs comprising nucleic acid (such as RNA and/or DNA) and at least one cationic or cationically ionizable lipid described herein are prepared by (a) preparing a nucleic acid (such as RNA and/or DNA) solution containing water and a buffering system; (b) preparing an ethanolic solution comprising the cationic or cationically ionizable lipid and, if present, one or more additional lipids; and (c) mixing the nucleic acid (such as RNA and/or DNA) solution prepared under (a) with the ethanolic solution prepared under (b), thereby preparing the formulation comprising LNPs. After step (c) one or more steps selected from diluting and filtrating, such as tangential flow filtrating, can follow.
[0468] In some embodiments, the LNPs comprising nucleic acid (such as RNA and/or DNA) and at least one cationic or cationically ionizable lipid described herein are prepared by (a′) preparing liposomes or a colloidal preparation of the cationic or cationically ionizable lipid and, if present, one or more additional lipids in an aqueous phase; and (b′) preparing a nucleic acid (such as RNA and/or DNA) solution containing water and a buffering system; and (c′) mixing the liposomes or colloidal preparation prepared under (a′) with the nucleic acid (such as RNA and/or DNA) solution prepared under (b′). After step (c′) one or more steps selected from diluting and filtrating, such as tangential flow filtrating, can follow.
[0469] The present disclosure describes particles comprising nucleic acid (such as RNA and/or DNA, especially mRNA) and at least one cationic or cationically ionizable lipid which associates with the nucleic acid (such as RNA and/or DNA) to form nucleic acid (such as RNA and/or DNA) particles and compositions comprising such particles. The nucleic acid (such as RNA and/or DNA) particles may comprise nucleic acid (such as RNA and/or DNA) 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.
[0470] Suitable cationic or cationically ionizable lipids 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.
[0471] In some embodiments, nucleic acid particles (such as RNA and/or DNA particles, especially mRNA particles) comprise more than one type of nucleic acid (such as RNA and/or DNA) molecules, where the molecular parameters of the

nucleic acid (such as RNA and/or DNA) molecules may be similar or different from each other, like with respect to molar mass or fundamental structural elements such as molecular architecture, capping (only RNA), coding regions or other features, In particulate formulation, it is possible that each nucleic acid (such as RNA and/or DNA) species is separately formulated as an individual particulate formulation. In that case, each individual particulate formulation will comprise one nucleic acid (such as RNA and/or DNA) 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 (such as RNA and/or DNA) species separately (typically each in the form of a nucleic acid (such as RNA and/or DNA)-containing solution) together with a particle-forming agent, thereby allowing the formation of particles. Respective particles will contain exclusively the specific nucleic acid (such as RNA and/or DNA) species that is being provided when the particles are formed (individual particulate formulations). In some embodiments, 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, 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 (such as RNA and/or DNA)-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 all nucleic acid (such as RNA and/or DNA) species of the pharmaceutical composition are formulated together as a combined particulate formulation. Such formulations are obtainable by providing a combined formulation (typically combined solution) of all nucleic acid (such as RNA and/or DNA) 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 comprise more than one nucleic acid (such as RNA and/or DNA) species. In a combined particulate composition different nucleic acid (such as RNA and/or DNA) species are typically present together in a single particle.

Polymers

[0472] 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.

[0473] 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.

[0474] 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.

[0475] 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.

[0476] In certain embodiments, polymer may be protamine or polyalkyleneimine.

[0477] 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.

[0478] 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.

[0479] 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.Math.10.sup.2 to 10.sup.7 Da, preferably 1000 to 10.sup.5 Da, more preferably 10000 to 40000 Da, more preferably 15000 to 30000 Da, even more preferably 20000 to 25000 Da.

[0480] Preferred according to the disclosure is linear polyalkyleneimine such as linear polyethyleneimine (PEI).

[0481] 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 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.
[0482] 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.
Lipids
[0483] 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 insoluble or poorly soluble in water, but soluble in many organic solvents. 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.
[0484] As used herein, the term “hydrophobic” refers to any a molecule, moiety or group which is substantially immiscible or insoluble in aqueous solution. The term hydrophobic group includes hydrocarbons having at least 6 carbon atoms. The hydrophobic group can have functional groups (e.g., ether, ester, halide, etc.) and atoms other than carbon and hydrogen as long as the group satisfies the condition of being substantially immiscible or insoluble in aqueous solution.
[0485] The term “hydrocarbon” includes alkyl, alkenyl, or alkynyl as defined herein. It should be appreciated that one or more of the hydrogen in alkyl, alkenyl, or alkynyl may be substituted with other atoms, e.g., halogen, oxygen or sulfur. Unless stated otherwise, hydrocarbon groups can also include a cyclic (alkyl, alkenyl or alkynyl) group or an aryl group, provided that the overall polarity of the hydrocarbon remains relatively nonpolar.
[0486] The term “alkyl” refers to a saturated linear or branched monovalent hydrocarbon moiety which may have six to thirty, typically six to twenty, often six to eighteen carbon atoms. Exemplary nonpolar alkyl groups include, but are not limited to, hexyl, decyl, dodecyl, tetradecyl, hexadecyl, octadecyl, and the like.
[0487] The term “alkenyl” refers to a linear or branched monovalent hydrocarbon moiety having at least one carbon carbon double bond in which the total carbon atoms may be six to thirty, typically six to twenty often six to eighteen.
[0488] The term “alkynyl” refers to a linear or branched monovalent hydrocarbon moiety having at least one carbon carbon triple bond in which the total carbon atoms may be six to thirty, typically six to twenty, often six to eighteen. Alkynyl groups can optionally have one or more carbon carbon double bonds.
[0489] 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.
[0490] The term “lipid-like material”, “lipid-like compound” or “lipid-like molecule” relates to substances, in particular amphiphilic substances, that structurally and/or functionally relate to lipids but may not be considered as 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. Examples of lipid-like compounds capable of spontaneous integration into cell membranes include functional lipid constructs such as synthetic function-spacer-lipid constructs (FSL), synthetic function-spacer-sterol constructs (FSS) as well as artificial amphipathic molecules. Lipids are generally cylindrical. The area occupied by the two alkyl chains is similar to the area occupied by the polar head group. Lipids have low solubility as monomers and tend to aggregate into planar bilayers that are water insoluble. Traditional surfactant monomers are generally cone shaped. The hydrophilic head groups tend to occupy more molecular space than the linear alkyl chains. In some embodiments, surfactants tend to aggregate into spherical or elliptoid micelles that are water soluble. While lipids also have the same general structure as surfactants—a polar hydrophilic head group and a nonpolar hydrophobic tail—lipids differ from surfactants in the shape of the monomers, in the type of aggregates formed in solution, and in the concentration range required for aggregation. 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.
[0491] 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 steroids, i.e., sterol-containing metabolites such as 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.

[0492] 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.

[0493] 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 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.

[0494] 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).

[0495] 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.

[0496] 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.

[0497] 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 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 ofglucosamine that is glycosylated with two 3-deoxy-D-manno-octulosonic acid (Kdo) residues.

[0498] 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.

[0499] 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/Cationically Ionizable Lipids

[0500] The nucleic acid particles (such RNA and/or DNA particles) described herein comprise at least one cationic or cationically ionizable lipid as particle forming agent. Cationic or cationically ionizable lipids contemplated for use herein include any cationic or cationically ionizable lipids (including lipid-like materials) which are able to electrostatically bind nucleic acid. In some embodiments, cationic or cationically ionizable lipids 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.

[0501] As used herein, a “cationic lipid” refers to a lipid or lipid-like material having a net positive charge. Cationic lipids 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.

[0502] In some embodiments, a cationic lipid 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.

[0503] As used herein, a “cationically ionizable lipid” refers to a lipid or lipid-like material which has a net positive charge or is neutral, i.e., which is not permanently cationic. Thus, depending on the pH of the composition in which the cationically ionizable lipid is solved, the cationically ionizable lipid is either positively charged or neutral. For purposes of the present disclosure, cationically ionizable lipids are covered by the term “cationic lipid” unless contradicted by the

circumstances.
[0504] In some embodiments, the cationic or cationically ionizable lipid comprises a head group which includes at least one nitrogen atom (N) which is positive charged or capable of being protonated, e.g., under physiological conditions.
[0505] Examples of cationic or cationically ionizable lipids include, but are not limited to N,N-dimethyl-2,3-dioleyloxypropylamine (DODMA), 1,2-dioleoyl-3-trimethylammonium propane (DOTAP); 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), 1,2-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-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.sub.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).
[0506] In some embodiments, the cationic or cationically ionizable lipid is DOTMA. In some embodiments, the cationic or cationically ionizable lipid is DODMA.
[0507] DOTMA is a cationic lipid with a quaternary amine headgroup. The structure of DOTMA may be represented as follows:
##STR00006##
[0508] DODMA is an ionizable cationic lipid with a tertiary amine headgroup. The structure of DODMA may be represented as follows:
##STR00007##
[0509] In some embodiments, the cationic or cationically ionizable lipid may comprise from about 10 mol % to about 95 mol %, from about 20 mol % to about 95 mol %, from about 20 mol % to about 90 mol %, from about 30 mol % to about 90 mol %, from about 40 mol % to about 90 mol %, or from about 40 mol % to about 80 mol % of the total lipid present in the particle.
Additional Lipids
[0510] Particles described herein may also comprise lipids (including lipid-like materials) other than cationic or cationically ionizable lipids (also collectively referred to herein as cationic lipids), i.e., non-cationic lipids (including non-cationic or 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. Optimizing the formulation of nucleic acid particles by addition of other hydrophobic moieties, such as cholesterol and lipids, in addition to a cationic or cationically ionizable lipid may enhance particle stability and efficacy of nucleic acid delivery.
[0511] One or more additional lipids may or may not affect the overall charge of the nucleic acid particles. In some embodiments, the or more additional lipids are 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.

[0512] In some embodiments, the nucleic acid particles (especially the particles comprising mRNA) described herein comprise a cationic or cationically ionizable lipid and one or more additional lipids.
[0513] Without wishing to be bound by theory, the amount of the cationic or cationically ionizable lipid compared to the amount of the one or more additional lipids 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 cationic or cationically ionizable lipid to the one or more additional lipids is from about 10:0 to about 1:9, about 4:1 to about 1:2, about 4:1 to about 1:1, about 3:1 to about 1:1, or about 3:1 to about 2:1.
[0514] In some embodiments, the one or more additional lipids comprised in the nucleic acid particles (especially in the particles comprising mRNA) described herein comprise one or more of the following: neutral lipids, steroids, and combinations thereof.
[0515] In some embodiments, the one or more additional lipids comprise a neutral lipid which is a phospholipid. In some embodiments, the phospholipid is selected from the group consisting of phosphatidylcholines, phosphatidylethanolamines, phosphatidylglycerols, phosphatidic acids, phosphatidylserines and sphingomyelins. 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-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), 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine (DOPG), 1,2-dipalmitoyl-sn-glycero-3-phospho-(1′-rac-glycerol) (DPPG), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE), N-palmitoyl-D-erythro-sphingosylphosphorylcholine (SM), and further phosphatidylethanolamine lipids with different hydrophobic chains. In some embodiments, the neutral lipid is selected from the group consisting of DSPC, DOPC, DMPC, DPPC, POPC, DOPE, DOPG, DPPG, POPE, DPPE, DMPE, DSPE, and SM. In some embodiments, the neutral lipid is selected from the group consisting of DSPC, DPPC, DMPC, DOPC, POPC, DOPE and SM. In some embodiments, the neutral lipid is DOPE.
[0516] In some embodiments, the additional lipid comprises one of the following: (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.
[0517] Thus, in some embodiments, the nucleic acid particles (especially the particles comprising mRNA) described herein comprise (1) a cationic or cationically ionizable lipid, and a phospholipid such as DOPE or (2) a cationic or cationically ionizable lipid and a phospholipid such as DOPE and cholesterol.
[0518] In some embodiments, the nucleic acid particles (especially the particles comprising mRNA) described herein comprise (1) DOTMA and DOPE, (2) DOTMA, DOPE and cholesterol, (3) DODMA and DOPE or (4) DODMA, DOPE and cholesterol.
[0519] DOPE is a neutral phospholipid. The structure of DOPE may be represented as follows:
##STR00008##
[0520] The structure of cholesterol may be represented as follows:
##STR00009##
[0521] In some embodiments, particles described herein do not 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.
[0522] In some embodiments, the additional 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 2 mol % to about 80 mol %, from about 5 mol % to about 80 mol %, from about 5 mol % to about 60 mol %, from about 5 mol % to about 50 mol %, from about 7.5 mol % to about 50 mol %, or from about 10 mol % to about 40 mol % of the total lipid present in the particle. In some embodiments, the additional lipid (e.g., one or more phospholipids and/or cholesterol) comprises about 10 mol %, about 15 mol %, or about 20 mol % of the total lipid present in the particle.
[0523] In some embodiments, the additional lipid comprises a mixture of: (i) a phospholipid such as DOPE; and (ii) cholesterol or a derivative thereof. In some embodiments, the molar ratio of the phospholipid such as DOPE to the cholesterol or a derivative thereof is from about 9:0 to about 1:10, about 2:1 to about 1:4, about 1:1 to about 1:4, or about 1:1 to about 1:3.

Polymer-Conjugated Lipids

[0524] In some embodiments, a particle may comprise at least one polymer-conjugated lipid. A polymer-conjugated lipid is typically a molecule comprising a lipid portion and a polymer portion conjugated thereto. In some embodiments, a polymer-conjugated lipid is a PEG-conjugated lipid, also referred to herein as pegylated lipid or PEG-lipid.
[0525] In some embodiments, a polymer-conjugated lipid is designed to sterically stabilize a lipid particle by forming a

protective hydrophilic layer that shields the hydrophobic lipid layer. In some embodiments, a polymer-conjugated lipid can reduce its association with serum proteins and/or the resulting uptake by the reticuloendothelial system when such lipid particles are administered in vivo.

[0526] Various PEG-conjugated lipids are known in the art and include, but are not limited to pegylated diacylglycerol (PEG-DAG) such as 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG-DMG), a pegylated phosphatidylethanoloamine (PEG-PE), a PEG succinate diacylglycerol (PEG-S-DAG) such as 4-O-(2′,3′-di(tetradecanoyloxy)propyl-1-O-(ω-methoxy(polyethoxy)ethyl)butanedioate (PEG-S-DMG), a pegylated ceramide (PEG-cer), or a PEG dialkoxypropylcarbamate such as ω-methoxy(polyethoxy)ethyl-N-(2,3-di(tetradecanoxy)propyl)carbamate or 2,3-di(tetradecanoxy)propyl-N-(ω methoxy(polyethoxy)ethyl)carbamate, and the like.

[0527] In some embodiments, a particle may comprise one or more PEG-conjugated lipids or pegylated lipids as described in WO 2017/075531 and WO 2018/081480, the entire contents of each of which are incorporated herein by reference for the purposes described herein.

Lipoplex Particles

[0528] In some embodiments of the present disclosure, the nucleic acid (such as RNA and/or DNA) described herein may be present in nucleic acid lipoplex particles (such as RNA and/or DNA lipoplex particles).

[0529] Lipoplexes (LPX) are electrostatic complexes which are generally formed by mixing preformed cationic lipid liposomes with anionic nucleic acid (such as RNA and/or DNA). Formed lipoplexes possess distinct internal arrangements of molecules that arise due to the transformation from liposomal structure into compact nucleic acid-lipoplexes (such as RNA- and/or DNA-lipoplexes). These formulations are generally characterized by their poor encapsulation of the nucleic acid and incomplete entrapment of the nucleic acid.

[0530] In certain embodiments, the nucleic acid lipoplex particles (such as RNA and/or DNA 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.

[0531] 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.

[0532] 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.

[0533] Nucleic acid lipoplex particles (such as RNA and/or DNA lipoplex particles) described herein have an average diameter that in some embodiments 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 nucleic acid lipoplex particles (such as RNA and/or DNA lipoplex particles) have an average diameter that ranges from about 250 nm to about 700 nm. In another embodiment, the nucleic acid lipoplex particles (such as RNA and/or DNA lipoplex particles) have an average diameter that ranges from about 300 nm to about 500 nm. In an exemplary embodiment, the nucleic acid lipoplex particles (such as RNA and/or DNA lipoplex particles) have an average diameter of about 400 nm.

[0534] The nucleic acid lipoplex particles (such as RNA and/or DNA lipoplex particles) and compositions comprising nucleic acid lipoplex particles (such as RNA and/or DNA lipoplex particles) described herein are useful for delivery of nucleic acid (such as RNA and/or DNA) to a target tissue after parenteral administration, in particular after intravenous administration.

[0535] 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, nucleic acid (such as RNA and/or DNA) lipoplex particles of the disclosure may be used for expressing nucleic acid (such as RNA and/or DNA) in the spleen. In an embodiment, after administration of the nucleic acid (such as RNA and/or DNA) lipoplex particles, no or essentially no nucleic acid (such as RNA) accumulation and/or nucleic acid (such as RNA) expression in the lung and/or liver occurs. In one embodiment, after administration of the nucleic acid (such as RNA and/or DNA) lipoplex particles, nucleic acid (such as RNA) accumulation and/or nucleic acid (such as RNA) expression in antigen presenting cells, such as professional antigen presenting cells in the spleen occurs. Thus, nucleic acid (such as RNA and/or DNA) lipoplex particles of the disclosure may be used for expressing nucleic acid (such as RNA and/or DNA), e.g., nucleic acid (such as RNA and/or DNA) encoding an antigen or at least one epitope, in such antigen presenting cells. In one embodiment, the antigen presenting cells are dendritic cells and/or macrophages.

[0536] The electric charge of the nucleic acid (such as RNA and/or DNA) 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 nucleic acid (such as 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 nucleic acid (such as RNA). The charge ratio of the positive charges present in the at least one

cationic lipid to the negative charges present in the nucleic acid (such as RNA) is calculated by the following equation: charge ratio=[(cationic lipid concentration (mol))*(the total number of positive charges in the cationic lipid)]/[(nucleic acid (such as RNA) concentration (mol))*(the total number of negative charges in nucleic acid (such as RNA))]. The concentration of nucleic acid (such as RNA) and the at least one cationic lipid amount can be determined using routine methods by one skilled in the art.

[0537] In one embodiment, at physiological pH the charge ratio of positive charges to negative charges in the nucleic acid (such as RNA and/or DNA) lipoplex particles is from 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 nucleic acid (such as RNA and/or DNA) lipoplex particles at physiological pH is 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.

Lipid Nanoparticles (LNPs)

[0538] In some embodiments, nucleic acid (such as RNA and/or DNA) described herein is present 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.

[0539] 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 PEG-lipid. LNPs may be prepared by mixing lipids dissolved in ethanol with nucleic acid in an aqueous buffer. In some embodiments, in the nucleic acid (such as RNA and/or DNA) LNPs described herein the nucleic acid (such as RNA and/or DNA, especially mRNA) is bound by ionizable lipid that occupies the central core of the LNP. PEG lipid forms the surface of the LNP, along with phospholipids. In some embodiments, the surface comprises a bilayer. In some embodiments, cholesterol and ionizable lipid in charged and uncharged forms can be distributed throughout the LNP.

[0540] In some embodiments, the LNP comprises one or more cationic lipids, and one or more stabilizing lipids. Stabilizing lipids include neutral lipids and pegylated lipids.

[0541] In some embodiments, the LNP comprises a cationic lipid, a neutral lipid, a steroid, a polymer-conjugated lipid; and the nucleic acid (such as RNA and/or DNA), encapsulated within or associated with the lipid nanoparticle.

[0542] In some embodiments, the LNP comprises from 40 to 55 mol percent, from 40 to 50 mol percent, from 41 to 50 mol percent, from 42 to 50 mol percent, from 43 to 50 mol percent, from 44 to 50 mol percent, from 45 to 50 mol percent, from 46 to 50 mol percent, or from 46 to 49 mol percent. In some embodiments, 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.

[0543] In some embodiments, 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.

[0544] In some embodiments, 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.

[0545] In some embodiments, the LNP comprises from 45 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 5 mol percent of a polymer-conjugated lipid; and the nucleic acid (such as RNA and/or DNA), encapsulated within or associated with the lipid nanoparticle.

[0546] In some embodiments, the mol percent is determined based on total mol of lipid present in the lipid nanoparticle. In some embodiments, the mol percent is determined based on total mol of cationic lipid, neutral lipid, steroid and polymer-conjugated lipid present in the lipid nanoparticle.

[0547] In some embodiments, 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 some embodiments, the neutral lipid is selected from the group consisting of DSPC, DPPC, DMPC, DOPC, POPC, DOPE and SM. In some embodiments, the neutral lipid is DSPC.

[0548] In some embodiments, the steroid is cholesterol.

[0549] In some embodiments, the polymer conjugated lipid is a pegylated lipid. In some embodiments, the pegylated lipid has the following structure:

##STR00010##

or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof, wherein: [0550] R.sup.12 and R.sup.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 some embodiments, R.sup.12 and R.sup.13 are each independently straight, saturated alkyl chains containing from 12 to 16 carbon atoms. In some embodiments, w has a mean value ranging from 40 to 55. In some embodiments, the average w is about 45. In some embodiments, R.sup.12 and R.sup.13 are each independently a straight, saturated alkyl chain containing about 14 carbon atoms, and w has a mean value of about 45.

[0551] In some embodiments, a pegylated lipid is or comprises 2-[(Polyethylene glycol)-2000]-N,N-ditetradecylacetamide.

[0552] In some embodiments, the cationic lipid component of the LNPs has the structure of Formula (III):

##STR00011##

or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein: [0553] one of L.sup.1 or L.sup.2 is —O(C═O)—, —(C═O)O—, —C(═O)—, —O—, —S(O).sub.x—, —S—S—, —C(═O)S—, SC(═O)—, —NR.sup.aC(═O)—, —C(═O)NR.sup.a—, NR.sup.aC(═O)NR.sup.a—, —OC(═O)NR.sup.a— or —NR.sup.aC(═O)O—, and the other of L.sup.1 or L.sup.2 is —O(C═O)—, —(C═O)O—, —C(═O)—, —O—, —S(O).sub.x—, —S—S—, —C(═O)S—, SC(═O)—, —NR.sup.aC(═O)—, —C(═O)NR.sup.a—, NR.sup.aC(═O)NR.sup.a—, —OC(═O)NR.sup.a—

or —NR.sup.aC(═O)O— or a direct bond; [0554] G.sup.1 and G.sup.2 are each independently unsubstituted C.sub.1-C.sub.12 alkylene or C.sub.1-C.sub.12 alkenylene; [0555] G.sup.3 is C.sub.1-C.sub.24 alkylene, C.sub.1-C.sub.24 alkenylene, C.sub.3-C.sub.8 cycloalkylene, C.sub.3-C.sub.8 cycloalkenylene; [0556] R.sup.a is H or C.sub.1-C.sub.12 alkyl; [0557] R.sup.1 and R.sup.2 are each independently C.sub.6-C.sub.24 alkyl or C.sub.6-C.sub.24 alkenyl; [0558] R.sup.3 is H, OR.sup.5, CN, —C(═O)OR.sup.4, —OC(═O)R.sup.4 or —NR.sup.5C(═O)R.sup.4; [0559] R.sup.4 is C.sub.1-C.sub.12 alkyl; [0560] R.sup.5 is H or C.sub.1-C.sub.6 alkyl; and [0561] x is 0, 1 or 2.

[0562] In some of the foregoing embodiments of Formula (III), the lipid has one of the following structures (IIIA) or (IIIB):

##STR00012##

wherein: [0563] A is a 3 to 8-membered cycloalkyl or cycloalkylene ring; [0564] R.sup.6 is, at each occurrence, independently H, OH or C.sub.1-C.sub.24 alkyl; [0565] n is an integer ranging from 1 to 15.

[0566] In some of the foregoing embodiments of Formula (III), the lipid has structure (IIIA), and in other embodiments, the lipid has structure (IIIB).

[0567] In other embodiments of Formula (III), the lipid has one of the following structures (IIIC) or (IIID):

##STR00013##

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

[0568] In any of the foregoing embodiments of Formula (III), one of L.sup.1 or L.sup.2 is —O(C═O)—. For example, in some embodiments each of L.sup.1 and L.sup.2 are —O(C═O)—. In some different embodiments of any of the foregoing, L.sup.1 and L.sup.2 are each independently —(C═O)O— or —O(C═O)—. For example, in some embodiments each of L.sup.1 and L.sup.2 is —(C═O)O—.

[0569] In some different embodiments of Formula (III), the lipid has one of the following structures (IIIE) or (IIIF):

##STR00014##

[0570] In some of the foregoing embodiments of Formula (III), the lipid has one of the following structures (IIIG), (IIIH), (IIII), or (IIIJ):

##STR00015##

[0571] 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.

[0572] 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.

[0573] In some of the foregoing embodiments of Formula (III), R.sup.6 is H. In other of the foregoing embodiments, R.sup.6 is C.sub.1-C.sub.24 alkyl. In other embodiments, R.sup.6 is OH.

[0574] In some embodiments of Formula (III), G.sup.3 is unsubstituted. In other embodiments, G3 is substituted. In various different embodiments, G.sup.3 is linear C.sub.1-C.sub.24 alkylene or linear C.sub.1-C.sub.24 alkenylene.

[0575] In some other foregoing embodiments of Formula (III), R.sup.1 or R.sup.2, or both, is C.sub.6-C.sub.24 alkenyl. For example, in some embodiments, R.sup.1 and R.sup.2 each, independently have the following structure:

##STR00016##

wherein: [0576] R.sup.7a and R.sup.7b are, at each occurrence, independently H or C.sub.1-C.sub.12 alkyl; and a is an integer from 2 to 12, [0577] wherein R.sup.7a, R.sup.7b and a are each selected such that R.sup.1 and R.sup.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.

[0578] In some of the foregoing embodiments of Formula (III), at least one occurrence of R.sup.7a is H. For example, in some embodiments, R.sup.7a is H at each occurrence. In other different embodiments of the foregoing, at least one occurrence of R.sup.7b is C.sub.1-C.sub.8 alkyl. For example, in some embodiments, C.sub.1-C.sub.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.sup.1 or R.sup.2, or both, has one of the following structures:

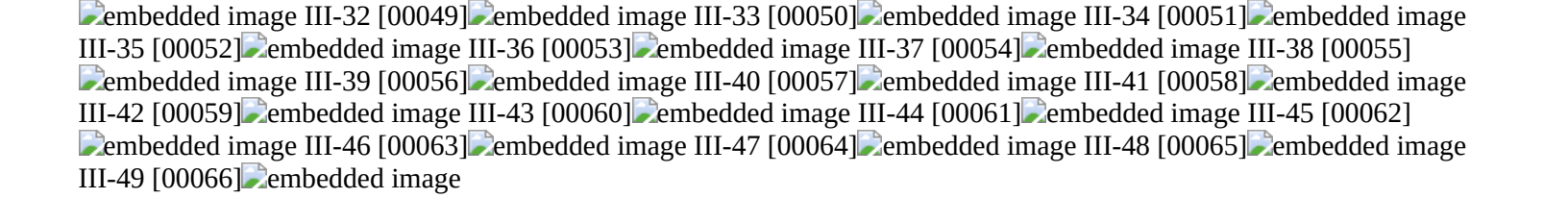
##STR00017##

[0579] In some of the foregoing embodiments of Formula (III), R.sup.3 is OH, CN, —C(═O)OR.sup.4, —OC(═O)R.sup.4 or —NHC(═O)R.sup.4. In some embodiments, R.sup.4 is methyl or ethyl.

[0580] In various different embodiments, the cationic lipid of Formula (III) has one of the structures set forth in the table below.

Representative Compounds of Formula (III).

TABLE-US-00002 No. Structure III-1 [00018]  III-2 [00019]  III-3 [00020]  III-4 [00021]  III-5 [00022]  III-6 [00023]  III-7 [00024]  III-8 [00025]  III-9 [00026]  III-10 [00027] III-11 [00028] III-12 [00029] III-13 [00030] III-14 [00031] III-15 [00032] III-16 [00033] III-17 [00034] III-18 [00035] III-19 [00036] III-20 [00037] III-21 [00038] III-22 [00039] III-23 [00040] III-24 [00041] III-25 [00042] III-26 [00043] III-27 [00044] III-28 [00045] III-29 [00046] III-30 [00047] III-31 [00048]


[0581] Various lipids (including, e.g., cationic lipids, neutral lipids, and polymer-conjugated lipids) are known in the art and can be used herein to form lipid nanoparticles, e.g., lipid nanoparticles targeting a specific cell type (e.g., liver cells). In some embodiments, a neutral lipid may be or comprise a phospholipid or derivative thereof (e.g., 1,2-Distearoyl-sn-glycero-3-phosphocholine (DPSC)) and/or cholesterol. In some embodiments, a polymer-conjugated lipid may be a PEG-conjugated lipid (e.g., 2-[(polyethylene glycol)-2000]—N,N-ditetradecylacetamide or a derivative thereof).
[0582] In some embodiments, the LNP comprises a lipid of Formula (III), nucleic acid (such as RNA and/or DNA), 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 ALC-0159.
##STR00067##
[0583] In some embodiments, the cationic lipid is present in the LNP in an amount from about 45 to about 50 mole percent. In some embodiments, the neutral lipid is present in the LNP in an amount from about 5 to about 15 mole percent. In some embodiments, the steroid is present in the LNP in an amount from about 35 to about 45 mole percent. In some embodiments, the pegylated lipid is present in the LNP in an amount from about 1 to about 5 mole percent.
[0584] In some embodiments, the LNP comprises a cationic lipid in an amount from about 45 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 5 mole percent.
[0585] 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 some embodiments, the N/P value is about 6.
Measuring Expression of Nucleic Acid in Transfected Cells
[0586] In order to quantify the amino acid sequence expressed in the cells transfected with nucleic acid encoding the amino acid sequence, one or more peptides of the nucleic acid-encoded sequence, e.g., unique MITD (MHC class I trafficking domain) peptide or a specific fragment of the peptide producing bioluminescence present e.g., at the C-terminal end of the encoded amino acid sequence, may be quantified from the total cell lysate using LC-MS/MS analysis.
Lysis of Cells
[0587] Any method which is suitable for lysing cells may be used in the assays described herein. In some embodiments, a buffer such as Tris/HCl buffer, e.g. having a pH of about 7.5 (e.g., adjusted with HCl), comprising a detergent such as a mild zwitterionic detergent, e.g., CHAPS (3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate) and/or CHAPSO (3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate) is used as lysis buffer. The lysis buffer may further comprise a chelating agent such as EDTA and/or one or more protease inhibitors.
[0588] The following table shows an example of a lysis buffer preparation and end component concentrations.
TABLE-US-00003 Ultrapure Water 9180 µL — 200 mM Tris/HCl pH 7.5 500 µL 10 mM 5M NaCl 300 µL 150 mM 0.5M EDTA 20 µL 1 mM CHAPS 100 mg 1% (w/v) Protease Inhibitor Cocktail 1 Tablet —
Quantification of Expression Product
[0589] Any method which is suitable for quantifying peptides and polypeptides may be used in the assays described herein. In some embodiments, Liquid Chromatography—Tandem Mass Spectrometry (LC-MS/MS) is used. In some embodiments, targeted LC-MS is used.
[0590] LC-MS/MS is a powerful analytical technique that combines the separating power of liquid chromatography with the highly sensitive and selective mass analysis capability of mass spectrometry. A sample solution containing analates of interest is pumped through a stationary phase (LC column) by a mobile phase flowing through at high pressure. Chemical interaction between the components of the sample, the stationary phase and the mobile phase affects different migration rates through the LC column affecting a separation. After elution from the LC column, the effluent is directed to the mass spectrometer. The mass spectrometer for an LC-MS/MS system has an ionization source where the LC column effluent is ionized creating charged particles. These charged particles then migrate under high vacuum through a series of mass analyzers by applying electromagnetic fields. The strength of this technique lies in the separation power of LC for a wide range of compounds combined with the capability of the MS to quantify compounds with a high degree of sensitivity and selectivity based on the unique mass/charge (m/z) transitions of each compound of interest.
[0591] Peptides monitored in the assays described herein may include: (1) peptides of the expressed nucleic acid, e.g., MITD7-28, Peptide 1 and/or Peptide 3 (both from tetanus toxin) and/or a fragment of the peptide producing bioluminescence, (2) Housekeeping peptides (used for the evaluation of suitability tests and for the calculation of the cell-based construct potency), e.g., CiRT-02, CiRT-06, CiRT-07, CiRT-11, CiRT-12 and/or CiRT-14, and optionally (3) Heavy-labelled, synthetic peptides (Used for retention time adjustment, digestion control, assessment of the chromatographic stability and for additional information of the sample preparation and MS performance).
[0592] In some embodiments, peptides of the expressed nucleic acid monitored in the assays described include one or more selected from the group consisting of MITD7-28 (GGSYSQAASSDSAQGSDVSLTA), Peptide 1 (FIGITELK), and Peptide 3 (IYSYFPSVISK).
[0593] In some embodiments, housekeeping peptides monitored in the assays described include one or more peptides

derived from proteins selected from the group consisting of Actin, 60S ribosomal protein L12, Heat shock protein SSA3, ADP/ATP translocase and/or 14-3-3 protein. In some embodiments, housekeeping peptides monitored in the assays described include one or more selected from the group consisting of CiRT-02 (AGFAGDDAPR; Actin), CiRT-06 (IGPLGLSPK; 60S ribosomal protein L12), CiRT-07 (TTPSYVAFTDTER; Heat shock protein SSA3), CiRT-11 (SYELPDGQVITIGNER; Actin), CiRT-12 (YFPTQALNFAFK; ADP/ATP translocase) and/or CiRT-14 (DSTLIMQLLR; 14-3-3 protein).

Compositions Comprising Nucleic Acid

[0594] A composition comprising one or more nucleic acids described herein, e.g., in the form of nucleic acid particles, may comprise salts, buffers, or other components as further described below.

[0595] In some embodiments, a salt for use in the compositions described herein comprises sodium chloride. Without wishing to be bound by theory, sodium chloride functions as an ionic osmolality agent for preconditioning nucleic acid (such as RNA and/or DNA) prior to mixing with lipids. In some embodiments, the compositions described herein may comprise alternative organic or inorganic salts. Alternative salts include, without limitation, potassium chloride, dipotassium phosphate, monopotassium phosphate, potassium acetate, potassium bicarbonate, potassium sulfate, disodium phosphate, monosodium phosphate, sodium acetate, sodium bicarbonate, sodium sulfate, lithium chloride, magnesium chloride, magnesium phosphate, calcium chloride, and sodium salts of ethylenediaminetetraacetic acid (EDTA).

[0596] Generally, compositions for storing nucleic acid (such as RNA and/or DNA) particles such as for freezing nucleic acid (such as RNA and/or DNA) particles comprise low sodium chloride concentrations, or comprises a low ionic strength. In some embodiments, the sodium chloride is at a concentration from 0 mM to about 50 mM, from 0 mM to about 40 mM, or from about 10 mM to about 50 mM.

[0597] According to the present disclosure, the nucleic acid (such as RNA and/or DNA) particle compositions described herein have a pH suitable for the stability of the nucleic acid (such as RNA and/or DNA) particles and, in particular, for the stability of the nucleic acid (such as RNA and/or DNA). Without wishing to be bound by theory, the use of a buffer system maintains the pH of the particle compositions described herein during manufacturing, storage and use of the compositions. In some embodiments of the present disclosure, the buffer system may comprise a solvent (in particular, water, such as deionized water, in particular water for injection) and a buffering substance. The buffering substance may be selected from 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES), 2-amino-2-(hydroxymethyl)propane-1,3-diol (Tris), acetate, and histidine. A preferred buffering substance is HEPES.

[0598] Compositions described herein may also comprise a cyroprotectant and/or a surfactant as stabilizer to avoid substantial loss of the product quality and, in particular, substantial loss of nucleic acid (such as RNA and/or DNA, especially mRNA) activity during storage, freezing, and/or lyophilization, for example to reduce or prevent aggregation, particle collapse, nucleic acid (such as RNA and/or DNA, especially mRNA) degradation and/or other types of damage.

[0599] In an embodiment, the cryoprotectant is a carbohydrate. The term “carbohydrate”, as used herein, refers to and encompasses monosaccharides, disaccharides, trisaccharides, oligosaccharides and polysaccharides.

[0600] In an embodiment, the cryoprotectant is a monosaccharide. The term “monosaccharide”, as used herein refers to a single carbohydrate unit (e.g., a simple sugar) that cannot be hydrolyzed to simpler carbohydrate units. Exemplary monosaccharide cryoprotectants include glucose, fructose, galactose, xylose, ribose and the like.

[0601] In an embodiment, the cryoprotectant is a disaccharide. The term “disaccharide”, as used herein refers to a compound or a chemical moiety formed by 2 monosaccharide units that are bonded together through a glycosidic linkage, for example through 1-4 linkages or 1-6 linkages. A disaccharide may be hydrolyzed into two monosaccharides. Exemplary disaccharide cryoprotectants include sucrose, trehalose, lactose, maltose and the like.

[0602] The term “trisaccharide” means three sugars linked together to form one molecule. Examples of a trisaccharides include raffinose and melezitose.

[0603] In an embodiment, the cryoprotectant is an oligosaccharide. The term “oligosaccharide”, as used herein refers to a compound or a chemical moiety formed by 3 to about 15, such as 3 to about 10 monosaccharide units that are bonded together through glycosidic linkages, for example through 1-4 linkages or 1-6 linkages, to form a linear, branched or cyclic structure. Exemplary oligosaccharide cryoprotectants include cyclodextrins, raffinose, melezitose, maltotriose, stachyose, acarbose, and the like. An oligosaccharide can be oxidized or reduced.

[0604] In an embodiment, the cryoprotectant is a cyclic oligosaccharide. The term “cyclic oligosaccharide”, as used herein refers to a compound or a chemical moiety formed by 3 to about 15, such as 6, 7, 8, 9, or 10 monosaccharide units that are bonded together through glycosidic linkages, for example through 1-4 linkages or 1-6 linkages, to form a cyclic structure. Exemplary cyclic oligosaccharide cryoprotectants include cyclic oligosaccharides that are discrete compounds, such as α cyclodextrin, β cyclodextrin, or γcyclodextrin.

[0605] Other exemplary cyclic oligosaccharide cryoprotectants include compounds which include a cyclodextrin moiety in a larger molecular structure, such as a polymer that contains a cyclic oligosaccharide moiety. A cyclic oligosaccharide can be oxidized or reduced, for example, oxidized to dicarbonyl forms. The term “cyclodextrin moiety”, as used herein refers to cyclodextrin (e.g., an α, β, or γcyclodextrin) radical that is incorporated into, or a part of, a larger molecular structure, such as a polymer. A cyclodextrin moiety can be bonded to one or more other moieties directly, or through an optional linker. A cyclodextrin moiety can be oxidized or reduced, for example, oxidized to dicarbonyl forms.

[0606] Carbohydrate cryoprotectants, e.g., cyclic oligosaccharide cryoprotectants, can be derivatized carbohydrates. For example, in an embodiment, the cryoprotectant is a derivatized cyclic oligosaccharide, e.g., a derivatized cyclodextrin, e.g., 2-hydroxypropyl-β-cyclodextrin, e.g., partially etherified cyclodextrins (e.g., partially etherified β cyclodextrins).

[0607] An exemplary cryoprotectant is a polysaccharide. The term “polysaccharide”, as used herein refers to a compound or a chemical moiety formed by at least 16 monosaccharide units that are bonded together through glycosidic linkages, for example through 1-4 linkages or 1-6 linkages, to form a linear, branched or cyclic structure, and includes polymers that comprise polysaccharides as part of their backbone structure. In backbones, the polysaccharide can be linear or cyclic.
[0608] Exemplary polysaccharide cryoprotectants include glycogen, amylase, cellulose, dextran, maltodextrin and the like.
[0609] In some embodiments, nucleic acid (such as RNA and/or DNA) particle compositions may include sucrose. Without wishing to be bound by theory, sucrose functions to promote cryoprotection of the compositions, thereby preventing nucleic acid (such as RNA and/or DNA, especially mRNA) particle aggregation and maintaining chemical and physical stability of the composition. In some embodiments, nucleic acid (such as RNA and/or DNA) particle compositions may include alternative cryoprotectants to sucrose. Alternative stabilizers include, without limitation, trehalose and glucose. In a specific embodiment, an alternative stabilizer to sucrose is trehalose or a mixture of sucrose and trehalose.
[0610] A preferred cryoprotectant is selected from the group consisting of sucrose, trehalose, glucose, and a combination thereof, such as a combination of sucrose and trehalose. In a preferred embodiment, the cryoprotectant is sucrose.
[0611] Some embodiments of the present disclosure contemplate the use of a chelating agent in a nucleic acid (such as RNA and/or DNA) composition described herein. Chelating agents refer to chemical compounds that are capable of forming at least two coordinate covalent bonds with a metal ion, thereby generating a stable, water-soluble complex. Without wishing to be bound by theory, chelating agents reduce the concentration of free divalent ions, which may otherwise induce accelerated nucleic acid (such as RNA and/or DNA) degradation in the present disclosure. Examples of suitable chelating agents include, without limitation, ethylenediaminetetraacetic acid (EDTA), a salt of EDTA, desferrioxamine B, deferoxamine, dithiocarb sodium, penicillamine, pentetate calcium, a sodium salt of pentetic acid, succimer, trientine, nitrilotriacetic acid, trans-diaminocyclohexanetetraacetic acid (DCTA), diethylenetriaminepentaacetic acid (DTPA), and bis(aminoethyl)glycolether-N,N,N′,N′-tetraacetic acid. In some embodiments, the chelating agent is EDTA or a salt of EDTA. In an exemplary embodiment, the chelating agent is EDTA disodium dihydrate. In some embodiments, the EDTA is at a concentration from about 0.05 mM to about 5 mM, from about 0.1 mM to about 2.5 mM or from about 0.25 mM to about 1 mM.
[0612] In an alternative embodiment, the nucleic acid (such as RNA and/or DNA) particle compositions described herein do not comprise a chelating agent.
[0613] Compositions comprising nucleic acids described herein, optionally formulated in particles, may be useful as or for preparing pharmaceutical compositions or medicaments for therapeutic or prophylactic treatments.
[0614] The term “pharmaceutical composition” relates to a composition 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 by administration of said pharmaceutical composition to a subject.
[0615] 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 chemokines may be IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, INFa, INF-γ, GM-CSF, LT-a. Further known adjuvants are aluminum hydroxide, Freund's adjuvant or oil such as Montanide® ISA51. Other suitable adjuvants for use in the present disclosure include lipopeptides, such as Pam3Cys, as well as lipophilic components, such as saponins, trehalose-6,6-dibehenate (TDB), monophosphoryl lipid-A (MPL), monomycoloyl glycerol (MMG), or glucopyranosyl lipid adjuvant (GLA).
[0616] The pharmaceutical compositions of the present disclosure may be in a storable form (e.g., in a frozen or lyophilized/freeze-dried form) or in a “ready-to-use form” (i.e., in a form which can be immediately administered to a subject, e.g., without any processing such as diluting). Thus, prior to administration of a storable form of a pharmaceutical composition, this storable form has to be processed or transferred into a ready-to-use or administrable form. E.g., a frozen pharmaceutical composition has to be thawed, or a freeze-dried pharmaceutical composition has to be reconstituted, e.g. by using a suitable solvent (e.g., deionized water, such as water for injection) or liquid (e.g., an aqueous solution).
[0617] The pharmaceutical compositions according to the present disclosure are generally applied in a “pharmaceutically effective amount” and in “a pharmaceutically acceptable preparation”.
[0618] 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.
[0619] The term “pharmaceutically effective amount” refers to the amount which achieves a desired reaction or a desired effect alone or together with further doses. In some embodiments relating to the treatment of a particular disease, the desired reaction may relate to inhibition of the course of the disease. This comprises slowing down the progress of the disease and, in some embodiments, 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 pharmaceutical 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 pharmaceutical 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.

[0620] The pharmaceutical compositions of the present disclosure may contain buffers, preservatives, and optionally other therapeutic agents. In some embodiments, the pharmaceutical compositions of the present disclosure comprise one or more pharmaceutically acceptable carriers, diluents and/or excipients.

[0621] Suitable preservatives for use in the pharmaceutical compositions of the present disclosure include, without limitation, benzalkonium chloride, chlorobutanol, paraben and thimerosal.

[0622] 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, thickeners, surface active agents, preservatives, stabilizers, emulsifiers, buffers, flavoring agents, or colorants

[0623] 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.

[0624] 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 some embodiments, the pharmaceutical composition of the present disclosure includes isotonic saline.

[0625] 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).

[0626] Pharmaceutical carriers, excipients or diluents can be selected with regard to the intended route of administration and standard pharmaceutical practice.

Routes of Administration of Pharmaceutical Compositions

[0627] In some embodiments, the pharmaceutical compositions described herein may be administered intravenously, intraarterially, subcutaneously, intradermally, dermally, intranodally, intramuscularly, intratumorally, or peritumorally. In some 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 some embodiments, the pharmaceutical compositions are formulated for systemic administration. In some embodiments, the systemic administration is by intravenous administration.

Use of Compositions

[0628] Compositions comprising nucleic acids described herein, optionally formulated in particles, may be used in the therapeutic or prophylactic treatment of various diseases, in particular diseases in which provision of a peptide or polypeptide to a subject results in a therapeutic or prophylactic effect. For example, provision of an antigen or epitope which is derived from a virus may be useful in the treatment of a viral disease caused by said virus. Provision of a tumor antigen or epitope may be useful in the treatment of a cancer disease wherein cancer cells express said tumor antigen. Provision of a functional protein or enzyme may be useful in the treatment of genetic disorder characterized by a dysfunctional protein, for example in lysosomal storage diseases (e.g. Mucopolysaccharidoses) or factor deficiencies. Provision of a cytokine or a cytokine-fusion may be useful to modulate tumor microenvironment.

[0629] The term “disease” (also referred to as “disorder” herein) 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, 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.

[0630] 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. 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.

[0631] 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.
[0632] 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.
[0633] 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), or any other non-mammal-animal, including birds (chicken), fish or any other animal species that can be afflicted with or is susceptible to a disease (e.g., cancer, infectious diseases) but may or may not have the disease, or may have a need for prophylactic intervention such as vaccination, or may have a need for interventions such as by protein replacement. In many embodiments, the individual is a human being. Unless otherwise stated, the terms “individual” and “subject” do not denote a particular age, and thus encompass adults, elderlies, children, and newborns. In some embodiments of the present disclosure, the “individual” or “subject” is a “patient”.
[0634] The term “patient” means an individual or subject for treatment, in particular a diseased individual or subject.
[0635] Nucleic acid, in particular RNA, having potency according to assays described herein, may be administered to a subject for delivering the nucleic acid to cells of the subject.
[0636] Nucleic acid, in particular RNA, having potency according to assays described herein, may be administered to a subject for delivering a therapeutic or prophylactic peptide or polypeptide (e.g., a pharmaceutically active peptide or polypeptide) to the subject, wherein the nucleic acid encodes a therapeutic or prophylactic peptide or polypeptide.
[0637] Nucleic acid, in particular RNA, having potency according to assays described herein, may be administered to a subject for treating or preventing a disease in a subject, wherein delivering the nucleic acid to cells of the subject is beneficial in treating or preventing the disease.
[0638] Nucleic acid, in particular RNA, having potency according to assays described herein, may be administered to a subject for treating or preventing a disease in a subject, wherein the nucleic acid encodes a therapeutic or prophylactic peptide or polypeptide and wherein delivering the therapeutic or prophylactic peptide or polypeptide to the subject is beneficial in treating or preventing the disease.
[0639] In some embodiments, the nucleic acid is present in a composition as described herein.
[0640] In some embodiments, the nucleic acid is administered in a pharmaceutically effective amount.
[0641] In some embodiments, the subject is a mammal. In some embodiments, the mammal is a human.
[0642] In some embodiments of the disclosure, the aim is to induce an immune response by providing a vaccine.
[0643] 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, nucleic acids described herein are applicable for inducing or enhancing an immune response. Nucleic acids described herein are thus useful in a prophylactic and/or therapeutic treatment of a disease involving an antigen or epitope.
[0644] In some embodiments of the disclosure, the aim is to treat cancer by vaccination.
[0645] In some embodiments of the disclosure, the aim is to provide protection against an infectious disease by vaccination.
[0646] In some embodiments of the disclosure, the aim is to provide secreted therapeutic proteins, such as antibodies, bispecific antibodies, cytokines, cytokine fusion proteins, enzymes, to a subject, in particular a subject in need thereof.
[0647] In some embodiments of the disclosure, the aim is to provide a protein replacement therapy, such as production of erythropoietin, Factor VII, Von Willebrand factor, β-galactosidase, Alpha-N-acetylglucosaminidase, to a subject, in particular a subject in need thereof.
[0648] In some embodiments of the disclosure, the aim is to modulate/reprogram immune cells in the blood.
[0649] In some embodiments of the disclosure, the aim is to provide one or more cytokines or cytokine fusions which modulate tumor microenvironment to a subject, in particular a subject in need thereof.
[0650] In some embodiments of the disclosure, the aim is to provide one or more cytokines or cytokine fusions which have antitumoral activity to a subject, in particular a subject in need thereof.
[0651] 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.
[0652] The description (including the following examples) 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: Justification Cell Line (CHO)—Uptake Via Macropinocytosis

[0653] Dendritic cells (DCs), the target cells which take up RNA-LPX by macropinocytosis, adopt a specialized process that can be inhibited by rottlerin (Kranz et al. 2016). Previous reports demonstrate that CHO cells are also capable of taking up a variety of nanoparticles (including lipid-based nanoparticles) via macropinocytosis (Hufnagle et al. 2009; Cardarelli et al. 2012; Pozzi et al. 2014; Zhang et al. 2011).

[0654] Characterization studies of CHO cells confirmed efficient and dose-dependent uptake of the RNA-LPX (FIG. **1**). This uptake is antagonized by the macropinocytosis inhibitor rottlerin. These data verify that the RNA-LPX uptake mechanism in both DC and CHO cells is mainly driven by macropinocytosis.

Example 2: Justification Cell Line (CHO)—Intracellular Localization of RNA-LPX-Encoded Antigen

[0655] Previous reports have shown that the RNA encoded protein-constructs were processed and directed to the plasma membrane from antigen-presenting cells (Kreiter et al. 2008). DCs and CHO cells were co-incubated with an RNA-LPX encoding a tumor antigen (MAGE A3) and localization of the translated MAGE A3 protein was assessed. In both cell types, the MAGE A3 protein that was translated from the RNA was detectable and was localized in the same cell compartment at the plasma cell membrane (FIG. **2**). The data indicate that DC and CHO cells translate and process the RNA-encoded protein in a similar manner

Example 3: Justification Cell Line (CHO)—RNA-LPX Dose-Dependent Detection of RNA-Encoded, Functional Protein

[0656] A typical characteristic of a quantitative potency assay is a sigmoidal dose-response-relation with an initial low response to the drug, a linear portion and a plateau indicating a saturation from the drug. We analyzed the RNA-LPX dose-response-relation in DC and CHO cells with the expression of an RNA-LPX encoding enhanced green fluorescent protein (eGFP) as a reporter protein. Both cell types show a dose-dependency with an initial phase and a linear phase and both are sensitive to a high RNA-LPX dosage (plateau phase) (FIG. **3**, FIG. **4**; for CHO also in FIG. **5**, and LC-MS read-out also in FIG. **6** Panels C). This was also demonstrated for tumor antigens and alternative read-outs (FIG. **8** and FIG. **11**). In contrast, using the same RNAs in a cell-free based assay (reticulocyte lysate) a dose-dependent signal increase for both capped and uncapped RNA was observed (FIG. **5** Panels C). This shows that the potency assay disclosed herein is capable of distinguishing between uncapped and capped RNAs, whereas cell-free assays may not be able to distinguish between uncapped and capped RNAs.

Example 4: Justification Cell Line (CHO)—Analysis with Stressed Samples

[0657] To evaluate sensitivity of the CHO cells in comparison to DC towards reduced product quality of the RNA-LPX, the following experiments with thermally stressed eGFP RNA-LPX samples were performed. First, eGFP RNA-LPX samples were stored for three days at 40° C., a stress condition which is known to reduce the product quality (e.g. reduction of the RNA integrity). The stressed samples, as well as non-stressed control eGFP RNA-LPX reference samples, were lipofected into DC and CHO cells. The impact of RNA-LPX product quality on reporter protein signal (i.e. assessing translation/function of the protein) is directly determined. The assay was found to have sensitivity towards drug product variations in both cell types (FIG. **4**).

[0658] To further analyze the impact of RNA quality on the potency assay readout, we tested a set of eGFP-RNA-LPX drug product batches manufactured with RNA with levels of RNA integrity from 40-95% and with an RNA manufactured without an 5′-cap. We measured the eGFP fluorescence readout for the analysis of the drug product batches since eGFP fluorescence signal is directly proportional to the amount of MS-based, relatively quantified peptides as shown in FIG. **6**. All of the RNA-LPX were tested in a dose dependent manner within 24 hours of incubation time the fluorescence signals were monitored (FIG. **5**).

[0659] We found a direct correlation of the active amount of eGFP protein and the degree of RNA integrity (FIG. **5**A). Furthermore, no signal was detected in the un-capped RNA batch indicating loss of functionality (FIG. **5**B). This was expected since the translation efficiency and stability of eukaryotic RNA strongly depends on cis-acting elements such as 5′-cap (Kuhn et al., 2011).

Example 5: Proof-of-Concept—LC-MS Detection of Encoded Peptides in Cell Lysates

[0660] The proof-of-concept for the MS-based potency assay was shown with the reporter gene (eGFP) as a surrogate marker. Use of eGFP allowed for direct, quantitative comparison of the readouts from quantification of the fluorescence and the MS signal in parallel. The cells were incubated with different doses of eGFP-RNA-lipoplexes and monitored for the expression (FIG. **6**, Panel A). After overnight incubation the cells were washed, harvested and lysed. The fluorescence intensity of the lysates was analysed prior the digestion (FIG. **6**, Panel C blue curve). Afterwards the digested lysates were separated with a nanoUPLC system and the peptides were analysed by MS/MS. We identified 16 eGFP-specific peptides (FIG. **6**, Panel B) in the treated lysates (and none in the mock transfected cell lysate). The most suitable three peptides (regarding e.g. charge, length/mass) were selected for a relative, label free quantification and were directly compared to the eGFP fluorescence intensity signal of the lysates (FIG. **6**, Panel C). Both signals (MS vs fluorescence) were directly correlated in the analysed dose range and show again the expected sigmoidal curve characteristic.

[0661] The data clearly demonstrate, that the signals from the active, encoded protein, namely the fluorescence and the MS peptide quantification directly correlate, with a highly similar dose-response curve. This demonstrated that specific peptides from the encoded and biological active eGFP protein can be detected and identified in cell lysates after lipofection via the label-free-MS approach. In addition, the specific peptides can be relatively quantified in a dose-dependent manner, and they directly reflect the amount of functional protein expressed by the cells.

[0662] After the proof-of-concept with the eGFP reporter protein, we analyzed if we can in addition detect and quantify iNeST specific peptides encoded by the six representative iNeST constructs (see following Table). Therefore, we transfected the cells separately with the six different RNA-lipoplex representative constructs and analyzed the relative

amount of the C-terminal, constant MITD peptide (see FIG. **13**).
TABLE-US-00004 Tested representative iNeST RNA sequences DNA Sequence Length Template DNA Template GC-Sequence ID Description (bp) Content (%) iNeST 1 Average 1947 56 iNeST 2 Average 1976 54 iNeST 4 Short 1385 54 iNeST 3 Long 2149 54 iNeST 6 Low GC 2065 50 iNeST 5 High GC 2045 58
[0663] For all six iNeST representative constructs we found the encoded MITD-peptide specifically only in the lipofected cells (and again none in the mock or non-transfected cell lysate). This demonstrated that the specific peptides from the expressed neoantigen-construct can be also detected and identified in cell lysates after lipofection via the proposed label-free-MS approach. Interestingly the translation level, namely the amount of the MITD peptide vary for the individualized neoantigen constructs by a factor of .sup.˜20fold (FIG. **7** and FIG. **9**).
[0664] To analyze if also the RNA-LPX of the iNeST-specific constructs show a sigmoidal dose-response relation (see section Justification cell line (CHO)—RNA-LPX Dose-Dependent Detection of RNA-encoded, functional protein, FIG. **3**), the experiments were repeated at different concentrations with the two iNeST constructs with the lowest MITD translation levels (iNeST5 and 6). Analogous to the mentioned data (FIG. **3** and FIG. **6**), again we found a direct, sigmoidal dose-response relation for the iNeST-constructs (FIG. **8**). Comparable to the eGFP data, we were able to demonstrate that also the activity (e.g. surface localization and immunological read out) of encoded tumor antigens follow a dose-response relationship (see, FIG. **11**). In summary our cell model system in combination with our LC-MS approach show a sigmoidal dose-response-relation, a typical characteristic of a quantitative potency assay.
Example 6: Proof-of-Concept—Analysis with Stressed Samples
[0665] One necessity of a potency assay is not only use for release testing but also stability testing. To verify that the demonstrated cellular sensitivity to stressed samples can also be detected using the iNeST LPX samples and the LC-MS drug product potency assay, a series of different experiments were performed. To evaluate product stability, we measured all six representative iNeST design space RNA-LPX from stability studies subjected to accelerate stress from heat. All samples including the respective non-stressed controls were applied at three different dose levels to cells and analysed by the LC-MS potency assay (FIG. **9**). For all stressed iNeST design space RNA-LPX samples, a clear reduction in the quantity of MITD peptide can be seen, moreover, this is visible for all applied RNA-LPX dosages.
[0666] In addition, the effect of two accelerated stress time points (2 and 10 days) on LC-MS potency measurement was analyzed. The rationale for the two time points was to obtain samples with different levels of RNA integrity above and below the RNA integrity specification limit of 80% (FIG. **10**). These samples were analogous added to cells with the untreated control samples and analysed by LC-MS measurement. Again for all stressed iNeST RNA-LPX a clear reduction in potency can be seen which is also stress time dependent. Interestingly, the potency readout is highly sensitive towards decrease in RNA integrity. A loss of 10 to 15% RNA integrity resulted in a decrease of 25% to 75% potency activity, furthermore, a loss below 80% RNA integrity specification resulted in an even more major decrease of potency activity.
[0667] In summary, the data demonstrate, that the LC-MS potency assay can sensitively detect losses in the drug product quality, which can affect the translated amount of protein after administration.
[0668] These changes comprise loss of RNA integrity, variation of capping efficacy, variation of other aspects of RNA quality, as well as variations of complexation state which affect translation. The LC-MS assay is highly stability indicating (which may be considered as a prerequisite for a potency assay).
Example 7: Proof-of-Concept—RNA-LPX Biological Effect Correlates Directly with Biological Function in a Dose-Dependent Manner
[0669] The ultimate correlation should be between (i) antigen-specific peptide amount quantified via the LC-MS approach, (ii) the processing of the antigen-construct measured via immunofluorescence microscopy of natively stained cells and (iii) the functional readout detected via a Jurkat NFAT assay. For two different tumor antigens (human leukocyte antigen [HLA] class I and II restricted) we detected a direct dose-dependent relation between the antigen-specific peptide quantified by the LC-MS approach and the surface localization of the antigen construct (processing). Furthermore, it was demonstrated that tumor antigen MAGE-A3 RNA-LPX transfected CHO cells expressing HLA-A*0101 were able to stimulate MAGE-A3-TCR expressing Jurkat T cells in a RNA-LPX dose-dependent manner (FIG. **11**, A). In order to evaluate if CHO cells are able to process and present HLA class II restricted epitopes to induce antigen-specific T cell stimulation, the same experiment was conducted using an HLA-DQBA1*0102/DQB1*0501-restricted HPV-E7-specific TCR (FIG. **11**, B). HPV-E7 RNA-LPX transfected CHO cells expressing HLA-DQBA1*0102/DQB1*0501 were able to stimulate HPV-E7-TCR expressing Jurkat T cells in a RNA-LPX dose-dependent manner.
[0670] For all three read-outs (LC-MS; AB and Jurkat) again a sigmoidal dose-response-curve which is comparable to relevant measurements of DC (see FIGS. **3** and **4**) can be found. Which demonstrate that the amount of detected peptide (LC-MS) is again directly linked to the amount of active protein (Jurkat) (analog to eGFP see FIG. **6**C). In summary, these data indicate that CHO cells are a suitable model for DCs with regard to the mechanisms involved in how they engulf RNA-LPX, intracellularly process antigen-encoding RNA and (if artificially equipped with an HLA molecule) present it either on HLA class I or class II molecules for cognate activation of T cells. These data also demonstrate that the amount of transfected RNA and of translated protein obtained following RNA-LPX administration to CHO cells are directly linked and robust indicators for the level of activation of a T cell that recognizes the respective antigen in an HLA-restricted manner. The LC-MS potency readout is fully predictive for the biological activity of the RNA-LPX.
Example 8: Testing of Tumor Antigen Constructs with the LC-MS Assay
[0671] After the verification with the reporter eGFP gene, three tumor antigen constructs (FIG. **14**) were tested with the LC-MS assay, to show that the applied lysis protocol is suitable for dissolving the transmembrane (TRM) protein and the

feasibility of our MS-approach for our relevant TRM domain peptide (MITD). Based on the theoretical characteristic (e.g. length, charge/mass) of the TRM domain peptide (MITD) it is expected that this peptide performs poorly in the MS-analytic. Therefore, we analyze two alternative peptides (P1 and P2, p2p16 region) in the immunological encoded constructs which should perform well in the MS-based relative quantification (FIG. **15**).

[0672] The experiments were performed in a similar manner as for the eGFP-RNA-lipoplex proof-of-concept experiments. Again, lipofection of cells in a dose depended manner was done. After o/n incubation the cell were lysed, digested and relatively quantified via LC-MS/MS measurements. Beforehand the LC-MS/MS method was adapted for the three newly specific peptides. As expected, the three unique peptides (peptide #1; peptide #2 and MITD) of all three constructs were identified in the lysates of lipofected cells (FIG. **15**). Furthermore, the peptides were able to be relatively quantified and were expressed in an RNA-lipoplex dose dependent manner. Again, for most of the peptides derived from the three constructs the detection of the expected sigmoidal dose-response curve (FIG. **15**) can be obtained.

Example 9: LC/MS Potency Assay: —Bridging Data—

[0673] The following studies were performed to analyze if the LC-MS potency assay readout is correlated to orthogonal analytical in vitro readouts (similar to Example 7, FIG. **11**). The applied DP doses were chosen to be within the linear range of the assays.

[0674] Study design: Two different fixed combinations of mRNA-encoded non-mutated antigens shared within specific cancer types (combination 1; four different antigens/combination 2; two different antigens) were tested. For all six DP three samples (non stressed, and heat stressed for 3 days and 10 days, respectively) were subjected to LC-MS potency; Jurkat NFAT (HepG2 cells, GMP release assay performed under R&D); Jurkat NFAT (CHO cells); native immunofluorescence (IF) staining of transfected cells (Main analytics are LC-MS versus Jurkat (HepG2). Not all assays are available for all constructs (e.g. IF; Jurkat CHO)). All three DP per construct (3 days, 10 days stressed and non-stressed) were analyzed in the linear dose level.

[0675] FIG. **16** shows experiments for cells which were lipofected with DP (RNA-LPX; combination 1) encoding for MAGE A3 (FIG. **16**A) or TPTE (FIG. **16**B) tumor antigen in a dose-dependent manner. Afterwards different analytical readouts (LC-MS; Jurkat NFAT (with HepG2 or CHO cells) and quantitative immunofluorescence microscopy (IF)) were done in parallel and the results are linearly fitted and compared. Each DP which was accelerated heat stressed at two conditions (3 days (squares) and 10 days (circles)) and in addition the non-stressed DP (triangles) was applied on the cells in a dose-dependent manner. Error bars represent standard deviation.

[0676] FIG. **17** shows an evaluation and comparison of the different analytical results (LC-MS vs Jurkat and RNA integrity (CE) vs LC-MS) of the non-stressed (right value), 3 days (middle value) and 10 days (left value) stressed DP. The slopes of the linear fits for the different analytical measurements were plotted on a graph and fitted by a linear regression.

[0677] All assays show a direct linear correlation to the LC-MS assay.

[0678] FIG. **18** shows a further evaluation and comparison of the LC-MS (right bars in the diagrams) and Jurkat assay (left bars in the diagrams) results. Therefor, the data points of the dose-response curve were analyzed with the GMP-software PLA. The different stressed samples are compared to the non-stressed samples (set as 100% potency) and the respective potency is calculated using the PLA software. The values of the LC-MS and Jurkat assay were in very good agreement, indicating that the LC-MS is directly indicative for the T cell stimulation readout.

[0679] FIG. **19** shows experiments for cells which were lipofected with DP (RNA-LPX; combination 1) encoding for TYR (FIG. **19**A) or NY-ESO (FIG. **19**B) in a dose-dependent manner. Afterwards different analytical readouts (LC-MS; Jurkat NFAT (with HepG2)) were done in parallel and the results are linearly fitted and compared. Each DP which was accelerated heat stressed at two conditions (3 days (squares) and 10 days (circles)) and in addition the non-stressed DP (triangles) was applied on the cells in a dose-dependent manner. Error bars represent standard deviation.

[0680] FIG. **20** shows an evaluation and comparison of the different analytical results (LC-MS vs Jurkat and RNA integrity (CE) vs LC-MS) of the non-stressed (right value), 3 days (middle value) and 10 days (left value) stressed DP. The slopes of the linear fits for the different analytical measurements were plotted on a graph and fitted by a linear regression.

[0681] All assays show a direct linear correlation.

[0682] FIG. **21** shows a further evaluation and comparison of the LC-MS (right bars in the diagrams) and Jurkat assay (left bars in the diagrams) results. Therefor, the data points of the dose-response curve were analyzed with the GMP-software PLA. The different stressed samples are compared to the non-stressed samples (set as 100% potency) and the respective potency is calculated using the PLA software. The values of the LC-MS and Jurkat assay were in very good agreement, indicating that the LC-MS is directly indicative for the T cell stimulation readout.

[0683] FIG. **22** shows experiments for cells which were lipofected with DP (RNA-LPX; combination 2) encoding for E6 (FIG. **22**A) or E7 (FIG. **22**B) tumor antigen in a dose-dependent manner. Afterwards different analytical readouts (LC-MS; Jurkat NFAT (with HepG2 or CHO) and quantitative immunofluorescence microscopy (IF)) were done in parallel and the results are linearly fitted and compared. Each DP which was accelerated heat stressed at two conditions (3 days (squares) and 10 days (circles)) and in addition the non-stressed DP (triangles) was applied on the cells in a dose-dependent manner. Error bars represent standard deviation.

[0684] FIG. **23** shows an evaluation and comparison of the different analytical results (LC-MS vs Jurkat and RNA integrity (CE) vs LC-MS) of the non-stressed (right value), 3 days (middle value) and 10 days (left value) stressed DP. The slopes of the linear fits for the different analytical measurements were plotted on a graph and fitted by a linear regression.

[0685] All assays show a direct linear correlation.

[0686] FIG. **24** shows a further evaluation and comparison of the LC-MS (right bars in the diagrams) and Jurkat assay (left

bars in the diagrams) results. Therefor, the data points of the dose-response curve were analyzed with the relative potency GMP-software PLA. The different stressed samples are compared to the non-stressed samples (set as 100% potency) and the respective potency is calculated using the PLA software. The values of the LC-MS and Jurkat assay were in good agreement, indicating that the LC-MS is directly indicative for the T cell stimulation readout.

[0687] In summary, all assays show (i) a direct linear DP dose-relationship for all tested antigens, (ii) a lower activity for the stressed DP samples, this can be seen for all dosages tested, and (iii) a highly similar correlation in activity (non-stressed>3 days stressed>10 days stressed) for all tested DPs and dosages. Furthermore, the LC-MS readout is also correlated to RNA integrity (CQA) and is directly correlated with T cell activation (Jurkat NFAT).

Example 10: LC/MS Potency Assay: —Bridging Data—

[0688] The following studies were performed to (i) analyze if the LC-MS potency assay readout is applicable for other drug product formulations, e.g., LNP-based; enable the LC-MS assay as a generic potency platform assay for RNA-based nanoparticles, (ii) provide PoC that multiple antigens in one DP (one DP contains four different RNAs encoding for four different antigens, “onevial” approach) can be analyzed by the LC-MS potency assay, (iii) demonstrate that the LC-MS assay is applicable to different RNA formats (uRNA and mod RNA), (iv) provide PoC that the LC-MS potency assay is applicable for DNA transfection, and (v) demonstrate that LC-MS potency is applicable based on p2p16 peptide quantification data.

[0689] Study design: Two different LNPs were tested. Both DP were subjected to the LC-MS potency assay (cell culture part is slightly adapted (prolonged and dosing) for LNP and the LC-MS part is adapted for the specific encoded peptides). All two LNP DP were analyzed at different dose levels to generate a dose-response curve. RNA encoding a T-cell-string (RNA encoding a fusion protein of SARS-CoV-2 T-cell epitopes) and RNA encoding four different tuberculosis antigens was used.

[0690] Regarding the RNA encoding a T-cell-string, CHO cells were transfected with ATM or CTM DP (LNP). The DPs were applied in a dose-dependent manner (0.75 to 9 µg, also untranslated cells (UT) were analyzed as control) on the cells. After RNA uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. Two T cell specific peptides were analyzed by LC-MS/MS (PRM) and the normalized expression values were plotted.

[0691] The results in FIG. **25** demonstrate that the LC-MS approach is applicable to the LNP format. T cell string fusion constructs were consistently identified in all samples except of blanks (UT). Furthermore, a dose-dependent increase of the peptide amount was observed. Saturation starts between 3-6 µg of transfected DP LNP and decreasing peptide amounts were observed when transfecting 9 µg. Transfection of ATM results overall in higher translational levels of fusion construct compared to CTM.

[0692] Regarding the RNA encoding tuberculosis antigens, CHO cells were transfected with DP ATM (LNP) with uRNA or mod RNA. Each DP contains 4 RNAs encoding for 4 fusion proteins composed of 8 TB antigens (Ag85A+Hrp1; ESAT6+RpfD; M72+VapB47; RpfA+HbhA). The DP was applied in a dose-dependent manner (150 to 2400 ng) on the cells. After uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. TB-antigen specific peptides were analyzed by PRM. Error bars represent standard deviation.

[0693] The results in FIG. **26** demonstrate that the LC-MS approach is applicable to the LNP format, to other antigens, and to other specific LC-MS-analyzed peptides. It is demonstrated for a DP containing four RNAs (“onevial” approach) that all RNAs are translated. The LC-MS potency is applicable to the onevial approach. All antigens can be detected independent of the RNA format (mod RNA vs uRNA). Blank samples show no translation. Three RNAs were translated in a dose-dependent manner (ESAT6-RpfD is maybe a strongly secreted fusion protein and/or has low translation efficacy).

[0694] In further experiments, CHO cells were transfected with LNP control containing RNA encoding for the reporter luciferase, the luciferase being a secreted version. The DP was applied in one dose in triplicates on the cells. After uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. Luciferase specific peptides were analyzed by PRM. Error bars represent standard deviation.

[0695] The results in FIG. **27** demonstrate that luciferase specific peptides can be detected, even when the protein is secreted. Blank samples show no signal, no translation.

[0696] Further experiments involved a duplicate fluorescence measurement of CHO cells after 24 h of lipofection with DNA-lipofectamine 2000 (Lipofectamine 2000 (Fisher Scientific GmbH, 11668030), LOT: 2418953), the DNA encoding for eGFP. Four different DNA-lipofectamine doses (225 to 1800 ng) were applied onto the cells. After DNA uptake and translation, the cells were harvested, lysed and subjected to MS sample preparation. The eGFP fluorescence signal (fluorescence intensity) of the cells was correlated in comparison to the eGFP MS/MS peptide-signal intensities (MS signal) derived from two different GFP-specific-peptides (Top1 and Top2). The results in FIG. **28** demonstrate that the LC-MS potency assay is also applicable to DNA. Both LC-MS-quantified peptides show a direct linear correlation to the fluorescence signal of the cells.

[0697] In further experiments, cells were lipofected with DP (RNA-LPX) encoding for five different tumor antigens in a dose-dependent manner (eleven different dosages per antigen). After 24 h incubation time the cells were harvested and subjected to LC-MS analysis. The Peptide 3 (of the p2p16 domain) is quantified via the LC-MS approach and the normalized expression is plotted against the applied dosages.

[0698] The results in FIG. **29** demonstrate that the LC-MS potency is applicable to the drug products based on p2p16 peptide quantification data.

[0699] In further experiments, cells were lipofected with DP (RNA-LPX) encoding for one tumor antigen in a dose-dependent manner (eleven different dosages per antigen). After 24 h incubation time the cells were harvested and subjected

to LC-MS analysis. The Peptide 3 (of the p2p16 domain) is quantified via the LC-MS approach and the normalized expression is plotted against the applied dosages.
[0700] The results in FIG. **30** demonstrate that the LC-MS potency is applicable to the drug product based on p2p16 peptide quantification data.
[0701] In summary, the data described herein demonstrate that the potency assay described herein is applicable to: [0702] Multiple cell lines (adherent and suspension cells) [0703] Multiple DP formats (LPX, Lipofectamin2000, LNP, also electroporation works) [0704] The detection of multiple proteins/antigens (e.g., multiple RNAs per DP) [0705] Different RNA formats (mod RNA and uRNA) and DNA [0706] Different cell lysis approaches [0707] Different proteins (reporter (GFP and Luc), tumor antigens, neoantigens, virus antigens, secreted proteins).

Claims

1. A method for analyzing different batches of the same RNA encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity comprising the following steps: (i) providing the RNA; (ii) transfecting cells in vitro with the RNA; (iii) allowing expression of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity; (iv) determining the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof, and (v) using the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof as an indication for the potency of the RNA to induce the biological activity in a biological system.
2. The method of claim 1, wherein the cells are cells of an animal cell line, such as Chinese hamster ovary (CHO) cells.
3. The method of claim 1, wherein the fragment of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity is specific for the expressed amino acid sequence.
4. The method of claim 1, wherein the fragment of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity is not comprised by the amino acid sequence of a peptide or polypeptide having biological activity.
5. The method of claim 1, wherein the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is determined using mass spectroscopy.
6. The method of claim 1, wherein the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is determined using one or more amino acid sequences expressed by the cells as reference for quantification, wherein the one or more amino acid sequences expressed by the cells comprise one or more amino acid sequences of housekeeping proteins.
7. The method of claim 1, wherein the potency of the RNA to induce the biological activity in a biological system comprises the therapeutic potency of the RNA.
8. The method of claim 1, wherein the RNA has sufficient potency to induce the biological activity in a biological system such as therapeutic potency if the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is above a pre-determined cut-off or the RNA does not have sufficient potency to induce the biological activity in a biological system such as therapeutic potency if the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is below a pre-determined cut-off.
9. The method of claim 8, wherein the RNA used to determine the pre-determined cut-off and the RNA to be analyzed have the same chemical composition.
10. The method of claim 1, wherein RNA batches having sufficient potency to induce the biological activity in a biological system such as therapeutic potency are used or are to be used for therapy and/or RNA batches not having sufficient potency to induce the biological activity in a biological system such as therapeutic potency are not used or are not to be used for therapy.
11. The method of claim 1, wherein the potency of the RNA to induce the biological activity in a biological system such as therapeutic potency of the RNA reflects the quality such as the therapeutic quality of the RNA.
12. The method of claim 11, wherein the quality of the RNA reflects whether and/or to what extent the RNA was exposed to detrimental conditions.
13. The method of claim 1, wherein the potency of the RNA to induce the biological activity in a biological system such as therapeutic potency of the RNA reflects the integrity of the RNA and/or the capping of the RNA.
14. The method of claim 1, which is for analyzing whether the quality and/or quantity of the RNA is sufficient to induce the biological activity in a biological system.
15. The method of claim 1, wherein the amount of the amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having biological activity or a fragment thereof is indicative for whether the quality and/or quantity of the RNA is sufficient to induce the biological activity in a biological system.
16. The method of claim 1, wherein the biological activity comprises an ability to elicit a specific response in a disease-relevant system.
17. The method of claim 16, wherein the specific response comprises an immune response.
18. The method of claim 1, wherein the RNA is formulated as particles.
19. The method of claim 1, wherein the peptide or polypeptide having biological activity is a vaccine.

20. The method of claim 19, wherein the vaccine is a T cell vaccine.
21-56. (canceled)