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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 March 2025 (27.03.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/061863 A1

(51) International Patent Classification:
A61K 31/7088 (2006.01) *C12Q 1/686* (2018.01)
C12N 15/10 (2006.01) *G01N 30/72* (2006.01)

(21) International Application Number:
PCT/EP2024/076303

(22) International Filing Date:
19 September 2024 (19.09.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
23198654.8 20 September 2023 (20.09.2023) EP

(71) Applicant: **BIONTECH SE** [DE/DE]; An der Goldgrube 12, 55131 Mainz (DE).

(72) Inventors: **SCHLERETH, Julia**; c/o BioNTech SE, An der Goldgrube 12, 55131 Mainz (DE). **DEL CAMPO, Cristian**; c/o BioNTech SE, An der Goldgrube 12, 55131 Mainz (DE). **LE GALL, Mikaelle**; c/o BioNTech SE, An der Goldgrube 12, 55131 Mainz (DE). **KUHN, Andreas**; c/o BioNTech SE, An der Goldgrube 12, 55131 Mainz (DE). **CZECH, Andreas**; c/o BioNTech SE, An der Goldgrube 12, 55131 Mainz (DE).

(74) Agent: **D YOUNG & CO LLP**; 3 Noble Street, London EC2V 7BQ (GB).

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

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

Published:
— *with international search report (Art. 21(3))*
— *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

(54) Title: METHOD

(57) Abstract: A method for determining at least one quality parameter of an RNA sample by using reverse transcription and PCR, preferably droplet digital PCR, is provided. The quality parameter may be (a) the quantitative ratio of two or more RNA molecule species in an RNA sample containing RNA molecules of n RNA molecule species, or (b) the identity of the *n* RNA molecule species in the RNA sample, wherein n is an integer of at least 2; or (c) the integrity of an RNA sample containing RNA molecules of *n* RNA molecule species, wherein n is an integer of at least 1, or (d) the potency of a formulated RNA sample comprising RNA molecules of interest.

WO 2025/061863 A1

METHOD

Field of the Invention

The present invention relates to a method for quality control analysis of an RNA sample comprising *n* different RNA molecule species, i.e. for an RNA mixture. The method is suitable for use in quality control of RNA molecule mixtures during or following production. The invention relates to a method for determining at least one quality parameter of an RNA sample by using reverse transcription and PCR, preferably droplet digital PCR. The quality parameter may be the quantitative ratio of two or more RNA molecule species in an RNA sample containing RNA molecules of *n* RNA molecule species, or the identity of the *n* RNA molecule species in the RNA sample, wherein *n* is an integer of at least 2. Alternatively, the quality parameter may be the integrity of an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 1. Alternatively, the quality parameter may be the potency of a formulated RNA sample comprising RNA molecules of interest. The method is particularly suitable for analyzing mixtures of encapsulated and/or complexed RNA molecule mixtures, in particular for liposome, lipid nanoparticle- and lipoplex-formulated RNA molecule mixtures.

Background to the Invention

Ribonucleic acid (RNA) molecules represent an emerging class of drugs. RNA-based therapeutics may be used in immunotherapy, gene therapy and genetic vaccination, belonging to the most promising and quickly developing therapeutic fields of modern medicine. RNA-based therapeutics may provide highly specific and individual treatment options for the therapy of a large variety of diseases.

For certain medical treatments and applications, it is desired to apply a mixture of RNA species. Examples of such RNA mixture based treatments may include the application of polyvalent RNA mixtures that provide protection against several serotypes of a pathogen (e.g., hemagglutinin (HA) from multiple serotypes of Influenza A and B virus); RNA mixtures that provide different antigens from one

pathogen (e.g., different antigens from influenza, such as HA, nucleoprotein (NP), neuraminidase (NA) etc.); RNA mixtures that provide protection against several isoforms or variants of a cancer antigen (e.g., prostate specific antigen (PSA) in the context of prostate carcinoma); RNA mixtures that provide different epitopes of an antigen; RNA mixtures that contain a cancer specific and/or patient specific mixture of cancer antigens (expressed antigens or mutated antigens); RNA mixtures that encode a variety of antibodies (e.g., antibodies that are targeted against different epitopes of one or more proteins), or any other therapeutically active RNA mixture (e.g., encoding different isoforms of an enzyme for molecular therapy, different therapeutic proteins for treatment of an indication wherein several proteins have to be supplemented).

For certain medical indications, the RNA molecules in the RNA (mixture) sample may be present in complexed form, i.e. in the form of at least one RNA-carrier complex. The complexation or encapsulation of RNA in RNA- carrier complexes facilitates successful *in vivo* delivery. Complexing carrier compounds used in the art typically include various types of peptides, polymers, carbohydrates, cholesterol, polyethylene glycol (PEG), lipids, phospholipids, PEGylated lipids, cationic and polycationic compounds, and combinations thereof, as well as other carrier compounds, which may be assembled into RNA-carrier complexes.

For RNA mixture based therapeutics it is required that the *n* different components (*n* different RNA molecule species, complexed or free) of the drug product and drug substance can be characterized, in terms of presence, integrity, ratio and quantity (quality control parameter). Such quality controls may be implemented during or following the RNA sample production, and/or during or following complexation of the RNA sample and/or as a batch release quality control. As RNA mixture-based therapeutics can be composed of multiple RNA species of highly similar size and sequence (e.g., polyvalent vaccines composed of multiple similar antigens) standard methods for quality control to discriminate between RNA species of similar size such as agarose gel electrophoresis or analytic HPLC are not suitable. An ideal method for the quality control of RNA mixtures should be fast, robust, and cost effective allowing for the characterization of any or all of the quality control parameters selected from the group consisting of presence, quantity, and integrity of at least one

RNA molecule species and ratio of at least two RNA molecule species within an RNA sample comprising *n* different RNA molecule species. Hence, there is a need for a method for RNA quality control analysis in particular in terms of cost-efficiency, robustness and the ability of the methods to discriminate highly similar RNA species in RNA samples comprising n different RNA molecule species.

WO2018/211038 describes a method for quality control analysis of an RNA sample comprising *n* different RNA molecule species using reverse transcription and a polymerase chain reaction (PCR) based assay, wherein each of the *n* different RNA molecule species comprises one or more coding RNA molecules of synthetic origin, wherein *n* is an integer of at least 1, in some aspects at least 2, thereby determining at least one quality parameter, and wherein the PCR based assay is digital PCR (dPCR), preferably droplet digital (ddPCR). The assay may be used to determine various quality parameters, including the quantity of the RNA species present, the presence of the one or more coding RNA molecules, the integrity of the RNA molecules, and the quantitative ratio between the *n* RNA species. The drawbacks of the method described in this publication are that it requires the introduction of artificial sequences in the RNA sequence, which are purely for analytical purposes, rather than for the functionality of the RNA. In addition, multiple primer sets are needed for analysis of the *n* RNA species.

C. Du Cheyne et al. *Anal. Biochem.* 626 (2021) 114217 describes a 3':5' digital PCR assay to determine horse RNA integrity. This assay comprises, firstly, reverse transcription of the RNA into DNA, followed by analysis of the DNA using PCR with 3 probes and 3 pairs of primers at the 3' end, 5' end and centre of the DNA sequence. This method also suffers from the drawback that multiple primer sets are needed for analysis of the *n* RNA species.

Summary of the Invention

In a first aspect, the invention provides a method of determining a quality parameter of an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 2, the quality parameter being selected from the group consisting of:

i) quantitative ratio of two or more RNA molecule species of the *n* RNA molecule species; and

ii) identity of the *n* RNA molecule species in the RNA sample;

the method comprising the steps of:

a) reverse transcription of the *n* RNA molecule species in the RNA sample into cDNA molecules of *n* DNA molecule species; and

b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set and a single second primer, wherein

the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region of only one of the *n* DNA molecule species in the sample, and

the single second primer is capable of annealing to a second target region of all of the *n* DNA molecule species in the sample.

The quality parameter may be the quantitative ratio of two or more RNA molecule species of the *n* RNA molecule species. Therefore, in a second aspect, the invention provides a method of determining the quantitative ratio of two or more RNA molecule species in an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 2, the method comprising the steps of:

a) reverse transcription of the *n* RNA molecule species in the RNA sample into cDNA molecules of *n* DNA molecule species; and

b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set and a single second primer, wherein

the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region of only one of the *n* DNA molecule species in the sample, and

the single second primer is capable of annealing to a second target region of all of the *n* DNA molecule species in the sample.

The quality parameter may also be the identity of the *n* RNA molecule species in the RNA sample. Therefore, in a third aspect, the invention provides a method of determining identity of *n* RNA molecule species in an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 2, the

method comprising the steps of:
a) reverse transcription of the *n* RNA molecule species in the RNA sample into cDNA molecules of *n* DNA molecule species; and
b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set and a single second primer, wherein
the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region of only one of the *n* DNA molecule species in the sample, and
the single second primer is capable of annealing to a second target region of all of the *n* DNA molecule species in the sample.

In a fourth aspect, the invention provides a method of determining integrity of an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 1,
the method comprising the steps of:
a) reverse transcription of the *n* RNA molecules in the RNA sample into cDNA molecules of *n* DNA molecule species; and
b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set, a single second primer, a third primer set and a single fourth primer, wherein
the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region at the 3'-end region of a DNA molecule species in the sample,
the single second primer is capable of annealing at the 3'-end region of all of the *n* DNA molecule species in the sample;
the third primer set comprises *n* primer species, wherein each primer species is capable of annealing to a third target region at the 5'-end region of a DNA molecule species in the sample; and
the single fourth primer is capable of annealing to a fourth target region at the 5'-end region of all of the *n* DNA molecule species in the sample.

In a fifth aspect, the invention provides a method of determining the potency of a formulated RNA sample comprising RNA molecules of interest, the method

comprising the steps of:
a) providing an RNA sample which has been isolated from cells transfected with a formulated RNA sample;
b) reverse transcription of the RNA molecules in the RNA sample into cDNA molecules;
c) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer, a second primer, a third primer and a fourth primer, wherein
the first primer and the second primer are capable of annealing to a first target region and a second target region of the cDNA molecules produced from the RNA of interest in the sample, and
the third primer and the fourth primer are capable of annealing to a first target region and a second target region of the cDNA molecules derived from an endogenous RNA in the sample; and
d) comparing the measured amount of the cDNA produced from the RNA molecules of interest with the measured amount of cDNA produced from endogenous RNA.

In the method according to each of the above aspects of the invention, the PCR-based assay may use a detectable label. In any of these methods, the detectable label may be a fluorescent probe.

Brief Description of the Figures

Figure 1 shows the % integrity of each RNA composing the differently degraded RNA mixtures measured with ddPCR.
Figure 2 shows the copy numbers (CN) measured by digital droplet polymerase chain reaction (ddPCR) measuring the potency of a ribonucleic acid (RNA) of interest compared with a housekeeping gene from total RNA isolated from Chinese Hamster Ovary (CHO) cells which were previously transfected with four different amounts of formulated RNA of interest.

Advantages and Surprising Findings

It has been surprisingly found by the present inventors that carrying out the PCR method using the PCR-based assay using a first primer set and a single second primer, wherein the first primer set comprises *n* primer species, each primer species is capable of annealing to a first target region of only one of the *n* DNA molecule species in the sample (i.e. is specific to each DNA molecule species in the sample) , and the single second primer is capable of annealing to a second target region of all of the *n* DNA molecule species in the sample (i.e. is common to all DNA molecule species in the sample) can measure in parallel the quantitative ratio of two or more RNA molecule species in the original RNA sample and the identity of the *n* RNA molecule species in the original RNA sample, as well as the integrity of the *n* RNA molecule species, optionally also in parallel. In both of these aspects, the method is advantageous compared with the methods described in WO2018/211038 in that the experimental system is simplified compared with the by the fact that two components of each oligonucleotide set (i.e., one common primer and one common dual-labelled probe) are identical for all RNAs and only one primer of each set is RNA-specific. This reduces the amount of reagents needed, the complexity of the system and the pipetting time compared to assembling two sets each composed of RNA-specific oligonucleotides.

It has also been surprisingly found by the present inventors that carrying out the PCR method using four primer sets as defined herein is advantageous, compared with the method described in WO2018/211038 and in Du Cheyne et al. when determining the integrity of the *n* RNA molecule species in the original RNA sample. Similar to the identity and ratio assay method referred to above, the method is advantageous compared with the methods described in both documents in that the experimental system is simplified compared with the by the fact that two components of each oligonucleotide set (i.e., one common primer and one common dual-labelled probe) are identical for all RNAs and only one primer of each set is RNA-specific. This reduces the amount of reagents needed, the complexity of the system and the pipetting time compared to assembling two sets each composed of RNA-specific oligonucleotides. The method can also determine identity and quantitative ratio of the RNA species in parallel with determining integrity.

Finally, it has also been surprisingly found by the present inventors that carrying out the PCR method using four primer sets as defined herein that the method also allows determining the potency of a formulated RNA sample in a simplified manner compared with the methods described in WO2018/211038.

Definitions

Definitions of General Terms

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.

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".

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.

All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by the context.

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.

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.

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.

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.

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.

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.

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.

Nucleic acids

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.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.

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.

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 C1-4 alkyl groups, even more preferably one or more methyl groups. Particular examples of modified purine or pyrimidine base moieties include N7-alkyl-guanine, N6-alkyl-adenine, 5-alkyl-cytosine, 5-alkyl-uracil, and N(1)-alkyl-uracil, such as N7-C1-4 alkyl-guanine, N6-C1-4 alkyl-adenine, 5-C1-4 alkyl-cytosine, 5-C1-4 alkyl-uracil, and N(1)-C1-4 alkyl-uracil, preferably N7-methyl-guanine, N6-methyl-adenine, 5-methyl-cytosine, 5-methyl-uracil, and N(1)-methyl-uracil.

In one embodiment, the nucleic acid is DNA. 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).

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.

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).

"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. 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)).

In some embodiments, the nucleic acids of the present invention, such as one, at least two or all of the nucleic acids of the present invention, are RNA.

In some embodiments, the RNA is single stranded RNA.

In some embodiments, the RNA is mRNA.

In some embodiments, the RNA is generated by RNA in vitro transcription.

In some embodiments, the RNA comprises a 5' cap structure.

In some embodiments, the RNA does not comprise modified ribonucleotides.

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.

In some embodiments, the nucleic acids of the present invention, such as one, at least two or all of the nucleic acids of the present invention, are DNA.

In some embodiments, the DNA is present in the form of a vector.

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.

In some embodiments, the vector is a DNA vector.

In some embodiments, the nucleic acids of the present invention, such as one, at least two or all of the nucleic acids of the present invention comprise a mixture of RNA and DNA.

In some embodiments, the RNA in the mixture is single stranded RNA.

In some embodiments, the RNA in the mixture is mRNA.

In some embodiments, the RNA in the mixture is generated by RNA *in vitro* transcription.

In some embodiments, the RNA in the mixture comprises a 5' cap structure.

In some embodiments, the RNA in the mixture does not comprise modified ribonucleotides.

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.

In some embodiments, the DNA in the mixture is present in the form of a vector. 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.

In some embodiments, the vector in the mixture is a DNA vector.

In some embodiments, the nucleic acid (such as RNA and/or DNA) of the present invention, which can comprise one or at least two or more nucleic acid constructs, is formulated with a delivery vehicle.

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).

In some embodiments, the nucleic acid (such as RNA and/or DNA) is formulated as particles.

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.

In some embodiments, the nucleic acid (such as RNA and/or DNA) is formulated as lipid nanoparticles.

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.

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.

In some embodiments, the different amino acid sequences comprise the amino acid sequence of different peptides or polypeptides having biological activity.

In some embodiments, the different peptides or polypeptides having biological activity comprise different antigens.

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.

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) (m27,2'-OGppSpG) or m27,3'-OGppp(m12'-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.

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.

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.

Two nucleic acid sequences are complementary, if the bases forming a first nucleic acid sequence are able to form base pairs with the bases forming a second nucleic acid sequence by hydrogen bonds. Hydrogen bonds and therefore base pairs can be formed between a pyrimidine (i.e. cytosine, uracil, thymine and analogues thereof) and a purine (i.e. adenine, guanine and analogues thereof). In order to form a double strand two sequences do not need to be 100% complementary to each other. However, within the present invention the single-stranded nucleic acid molecules which serve as nucleic acid probes are preferably 100% complementary to the target sequence to enable a high level of quality control of the in vitro transcribed RNA.

mRNA

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.

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

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

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.

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, 4th 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, MAXIscript®), 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.

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.

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 José et al., *Future Microbiol.*, 2009, *4*, 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.

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 WO2019/175356, 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).

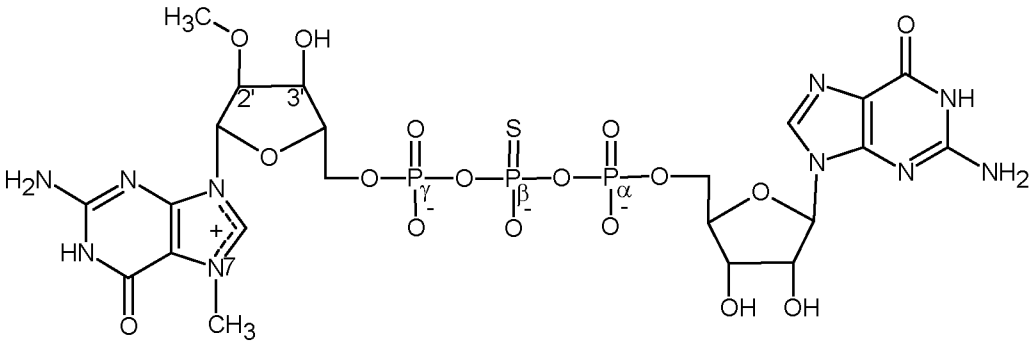
In some embodiments, the mRNA comprises a 5'-cap structure. In some embodiments, the mRNA does not have uncapped 5'-triphosphates. In some embodiments, the mRNA may comprise a conventional 5'-cap and/or a 5'-cap analog. The term "conventional 5'-cap" refers to a cap structure found on the 5'-end of an 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 N7 (resulting in the cap

structure m7Gppp). 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 m7guanosine 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 m27,2'OG(5')ppSp(5')G (referred to as beta-S-ARCA or β-S-ARCA)), as described in WO2019/175356. 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.

In some embodiments, the mRNA comprises a 5'-cap structure selected from the group consisting of m27,2'OG(5')ppSp(5')G (in particular its D1 diastereomer), m27,3'OG(5')ppp(5')G, and m27,3'-OGppp(m12'-O)ApG.

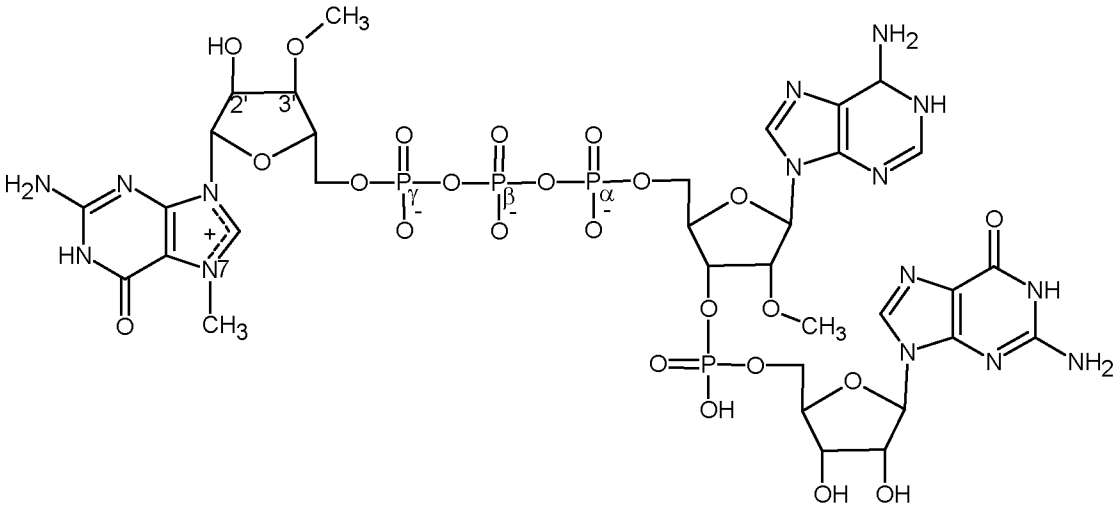
In some embodiments, the mRNA comprises a cap0, cap1, or cap2, preferably cap1 or cap2. According to the present disclosure, the term "cap0" means the structure "m7GpppN", wherein N is any nucleoside bearing an OH moiety at position 2'. According to the present disclosure, the term "cap1" means the structure "m7GpppNm", wherein Nm is any nucleoside bearing an OCH3 moiety at position 2'. According to the present disclosure, the term "cap2" means the structure "m7GpppNmNm", wherein each Nm is independently any nucleoside bearing an OCH3 moiety at position 2'.

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

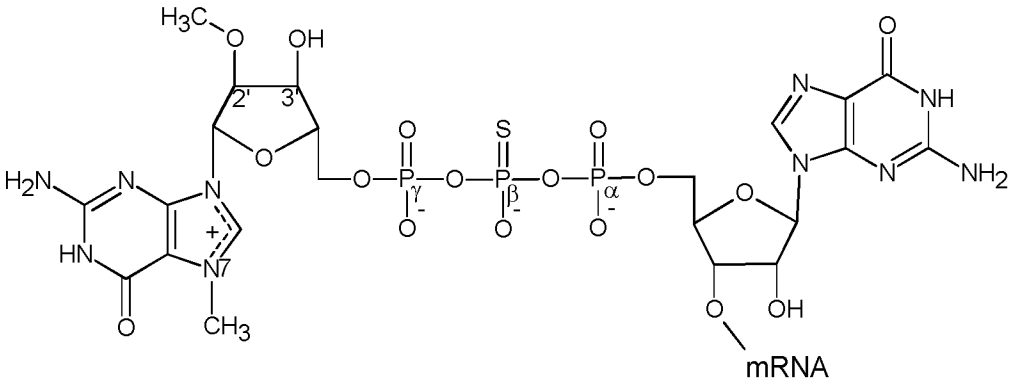


The "D1 diastereomer of beta-S-ARCA" or "beta-S-ARCA(D1)" is the diastereomer of beta-S-ARCA which elutes first on an HPLC column compared to the D2 diastereomer of beta-S-ARCA (beta-S-ARCA(D2)) and thus exhibits a shorter retention time. The HPLC preferably is an analytical HPLC. In some embodiments, a Supelcosil LC-18-T RP column, preferably of the format: 5 µm, 4.6 x 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.

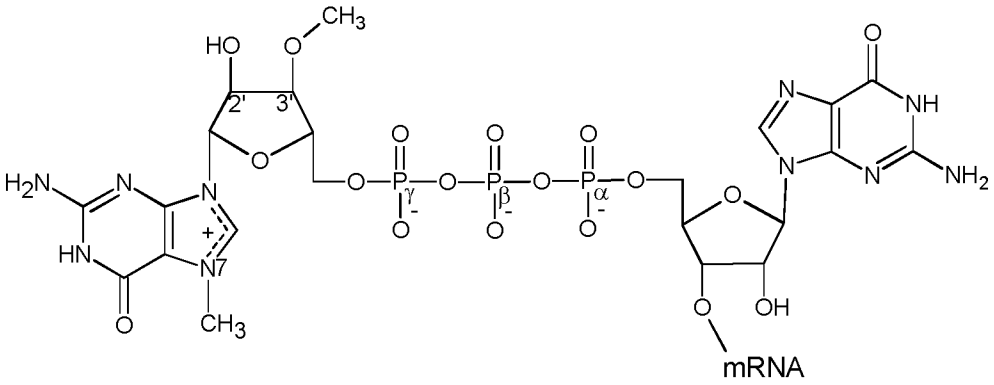
The 5'-cap analog m27,3'-OGppp(m12'-O)ApG (also referred to as m27,3'OG(5')ppp(5')m2'-OApG) which is a building block of a cap1 has the following structure:



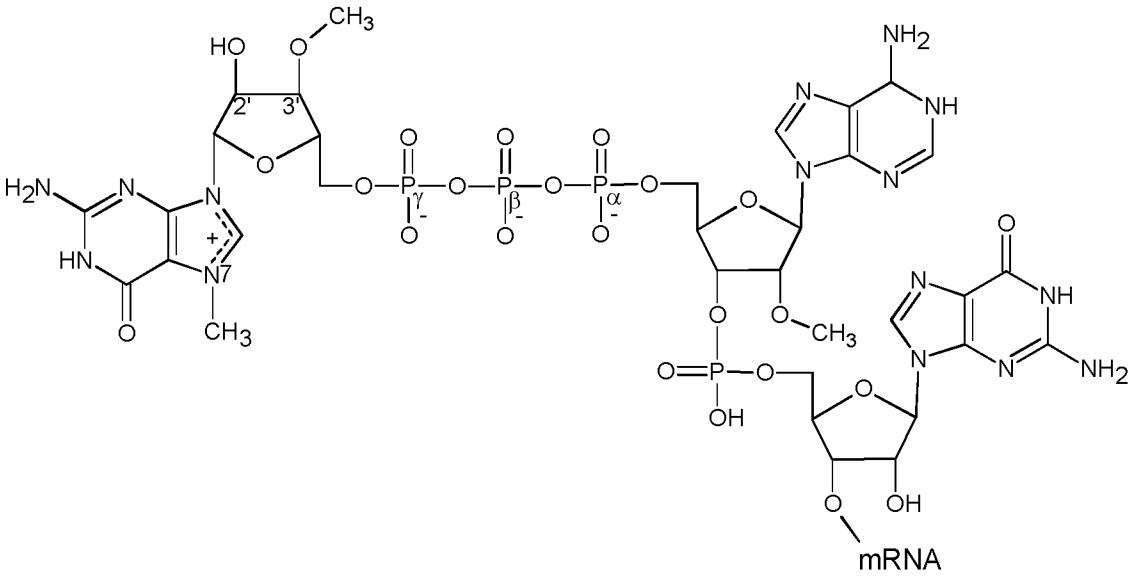
An exemplary cap0 mRNA comprising β-S-ARCA and mRNA has the following structure:



An exemplary cap0 mRNA comprising m27,3'OG(5')ppp(5')G and mRNA has the following structure:



An exemplary cap1 mRNA comprising m27,3'-OGppp(m12'-O)ApG and mRNA has the following structure:



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.

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., *Blood*, **2006**, *108*, 4009-4017).

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.

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.

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.

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.

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 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.

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.

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. 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.

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-uridineor 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-carboxymethyl-aminomethyl-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 (acp3 ψ), 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.

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.

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.
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".

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.
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.

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. 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.

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.

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.

RNA Molecule Species

The term "RNA molecule species" denotes at least one RNA molecule in a population of RNA molecules which do not differ in their RNA sequence and/or their sequence length. Hence, the RNA molecules within one RNA molecule species are encoded by the same template DNA. If the RNA present within the sample is a coding RNA, one RNA species may encode one target peptide or protein or variant thereof.

The term "*n* different RNA molecule species" denotes a group of *n* RNA molecules which may differ with respect to their RNA sequence and/or their sequence length. Hence, if an RNA sample comprises n different RNA molecule species and if *n* is 2, the RNA sample comprises RNA molecules which belong to either of the 2 RNA molecule species, i.e. have the same RNA sequence and/or their sequence length. In a typical RNA sample comprising 2 different RNA molecule species, the one or more RNA molecules of the first RNA molecule species do not differ in their RNA sequence and/or their sequence length among each other but differ from the RNA sequence and/or sequence length of the one or more RNA molecules of the second RNA molecule species. Each RNA molecule species comprises at least one RNA molecule, i.e. each RNA molecule species comprises one or more RNA molecules. Accordingly, if an RNA sample comprises *n* different RNA molecule species, the

RNA sample comprises at least *n* RNA molecules (at least one RNA molecule per RNA molecule species). However, typically an RNA molecule species comprises a higher number of RNA molecules per RNA molecule species in one sample. In the present invention, the one or more RNA molecules of each RNA molecule species are coding RNA species of synthetic origin. If the RNA molecule species are coding RNA molecule species, each of the *n* different RNA molecule species preferably but not necessarily encodes one target peptide/protein or variant thereof. In a sample comprising *n* different RNA species which is analyzed by the method of the present invention, identical, similar or different amounts of each species may be present, preferably the amounts are identical or similar. The number of different RNA molecule species which are present in the RNA sample are reflected by the integer *n* which is at least 1 and thus can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and so forth. Preferably, *n* is an integer of at least 2 or at least 3, or in the range of 1 to 200, or 2 to 200, more preferably of 2 to 150, even more preferably of 2 to 100, and most preferably of 2 to 50.

In the methods of the present invention, the *n* different RNA molecule species may have a similar length, as the detection and discrimination of the *n* different RNA species is not dependent on differences in the length. For example, the length of the n different RNA species within the sample may differ by not more than 10% or 8%, preferably not more than 7% or 6%, more preferably by not more than 5% and most preferably by not more than 4%.

In particular, the present method can be used with RNA molecule species which may have a similar sequences. For example, the sequences of the *n* different RNA molecule species within the sample or a part of these sequences may have a sequence identity of at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% . The sequence identity may be over a region comprising 200 nucleotides, 300 nucleotides, 400 nucleotides, 500 nucleotides or 600 nucleotides or more over the complete sequence.

In the present invention, the RNA sequences of the one or more coding RNA molecules of each of the *n* different RNA molecule species may be at least 80%

identical to each other, which considers the identity of the whole RNA molecule sequence and not only the target sequence.

Integrity of the one or more coding RNA molecules of an or at least two RNA molecule species: The term "integrity" describes whether the complete target RNA sequence is present in the sample of *in vitro* produced RNA. Low integrity could be due to, amongst others, degradation, cleavage, incorrect or incomplete chemical synthesis, incorrect base pairing, integration of modified nucleotides or the modification of already integrated nucleotides, lack of or incomplete capping, lack of or incomplete polyadenylation, or incomplete transcription.

Untranslated Regions

The primers used in the methods of the present invention may target the 3'-untranslated region (3'-UTR) or the 5'-untranslated region (3'-UTR).

Generally, the term "3'-UTR" refers to a part of the RNA molecule, which is located 3' (i.e. "downstream") of a coding sequence and which is not translated into protein. Typically, a 3'-UTR is the part of an mRNA which is located between the protein coding region (coding sequence (CDS)) and the 3' terminus of the mRNA. In the context of the invention, the term 3'-UTR may also comprise elements, which are not encoded in the template, from which an RNA is transcribed, but which are added after transcription during maturation, e.g. a poly(A) sequence (or poly(A) 'tail'). A 3'-UTR of the mRNA is not translated into an amino acid sequence. The 3'-UTR sequence is generally encoded by the DNA template, which is transcribed into the corresponding mRNA during the gene expression process. In the context of the present invention, a 3'-UTR corresponds to the sequence of a mature mRNA, which is located between the stop codon of the protein coding region, preferably immediately 3' to the stop codon of the protein coding region, and the poly(A) sequence of the mRNA.

5'-untranslated region (5'-UTR): A 5'-UTR is typically understood to be a particular section of messenger RNA (mRNA). It is located 5' of the coding sequence of the mRNA. Typically, the 5'-UTR starts with the transcriptional start site and ends one nucleotide before the start codon of the coding sequence. The 5'-UTR may comprise

elements for controlling gene expression, also called regulatory elements. Such regulatory elements may be, for example, ribosomal binding sites. The 5'-UTR may be post-transcriptionally modified, for example by addition of a 5' cap structure. In the context of the present invention, the term "5'-UTR" typically refers to the sequence of an mRNA, which is located between the 5' cap structure and the start codon. Preferably, the 5'-UTR is the sequence, which extends from a nucleotide located 3' to the 5' cap structure, preferably from the nucleotide located immediately 3' to the 5' cap structure, to a nucleotide located 5' to the start codon of the coding sequence, preferably to the nucleotide located immediately 5' to the start codon of the coding sequence.

Preparation of RNA Sample

The methods of the invention may involve the preparation of the sample including RNA. This method may include one or more purification steps.

As used herein, the term "purification" or "purifying" is understood to mean that the desired RNA in a sample is separated and/or isolated from impurities, intermediates, by-products and/or reaction components present therein or that the impurities, intermediates, by-products and/or reaction components are at least depleted from the sample comprising the RNA.

Non-limiting examples of undesired constituents of RNA-containing samples which therefore need to be depleted may comprise degraded fragments or fragments which have arisen as a result of premature termination of transcription, or also excessively long transcripts if plasmids are not completely linearized. Furthermore, intermediates may be depleted from the sample such as e.g. template DNA (for RNA IVT). Additionally, reaction components such as enzymes, proteins, bacterial DNA and RNA, small molecules such as spermidine, buffer components etc. may have to be depleted from the RNA sample. In addition, impurities such as organic solvents, nucleotides, nucleosides or other small molecules may be separated.

Reverse Transcription

Each of the methods of the invention comprises the step of reverse transcription of the *n* RNA molecule species in the RNA sample into cDNA molecules of *n* DNA molecule species.

In this sense, the term "reverse transcription" takes its normal meaning in the art as the process of generating complementary DNA (cDNA) from RNA. In this process an RNA, i.e. one or more coding RNA molecules of at least two RNA molecule species of the *n* different RNA molecule species, is typically incubated with the enzyme reverse transcriptase, deoxynucleotides (dNTPs), and at least one suitable primer for a time and under conditions sufficient for cDNA synthesis to occur, e.g. incubation for thirty minutes to one hour at a temperature of about 37 °C to 42 °C.

The primer(s) used for reverse transcription may be random so that any RNA molecule, e.g. one or more coding RNA molecules of all *n* different RNA molecule species, present in a sample may be reverse transcribed into cDNA or may be target-specific so that only the target RNA, e.g. one or more coding RNA molecules of at least two RNA molecule species of the *n* different RNA molecule species, are reverse transcribed into the corresponding cDNA.

PCR Assay

The methods of the invention further comprise carrying out a polymerase chain reaction (PCR)-based assay on the cDNA molecules resulting from reverse transcription of the RNA.

In its broadest sense, the term "PCR based assay" encompasses any assay that employs a PCR reaction. In one embodiment, the PCR assay is quantitative Polymerase chain reaction (qPCR). In one embodiment, the PCR assay is digital PCR (dPCR). In one embodiment, the PCR assay is droplet digital PCR (ddPCR).

As is known to the person skilled in the art, the polymerase chain reaction (PCR) is a technology in molecular biology used to amplify a piece of DNA across several

orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. The method relies on thermal cycling, consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA. Primers (short DNA fragments) containing sequences complementary to the target sequence along with a heat-stable DNA polymerase, such as Taq polymerase, enable selective and repeated amplification. As PCR progresses, the DNA generated is itself used as a template for replication, setting in motion a chain reaction in which the DNA template is exponentially amplified. The DNA polymerase enzymatically assembles a new DNA strand from DNA building-blocks, the nucleotides, by using single-stranded DNA as a PCR template and DNA oligonucleotides (also called DNA primers), which are required for initiation of DNA synthesis. The vast majority of PCR methods use thermal cycling, i.e., alternately heating and cooling the PCR sample through a defined series of temperature steps. In the first step, the two strands of the DNA double helix are physically separated at a high temperature in a process called DNA melting. In the second step, the temperature is lowered and the two DNA strands become templates for DNA polymerase to selectively amplify the target DNA. The selectivity of PCR results from the use of primers that are complementary to the DNA region targeted for amplification under specific thermal cycling conditions. RT-qPCR: An RT-qPCR assay involves a first step of reverse transcription and a second step of quantitative PCR as described above. The reverse transcription reaction and the quantitative PCR reaction may be performed separately so that in a first reaction the RNA is reverse transcribed into cDNA and in a second reaction the cDNA is transferred into a new reaction mixture for the quantitative PCR. Alternatively, the reverse transcription reaction and the quantitative PCR reaction may be performed in one step so that the reaction mixture comprises both the components of the reverse transcription reaction and the components of the quantitative PCR.

The PCR process generally consists of a series of temperature changes, known as thermal cycles, with each cycle commonly consisting of at least two, optionally three or four, discrete temperature steps. The cycling is often preceded by a single temperature step at a very high temperature (>90°C), and followed by one hold at the end for final product extension or brief storage. The temperatures used and the length of time they are applied in each cycle depend on a variety of parameters, including the

enzyme used for DNA synthesis, the concentration of bivalent ions and dNTPs in the reaction, and the melting point of the primers. The individual steps common to most PCR methods are as follows.

The first step, known as the initialization step, or activation step, is the heating of the reaction chamber in order to activate the DNA polymerase. In one embodiment, the initialization step takes place at a temperature between 90-100 °C. In one embodiment, the initialization step takes place at a temperature between 94-96 °C. In one embodiment, the initialization step takes place at a temperature of 95 °C. In one embodiment, the initialization step lasts 1 to 20 minutes. In one embodiment, the initialization step lasts 5 to 15 minutes. In one embodiment, the initialization step lasts 8 to 12 minutes. In one embodiment, the initialization step lasts 10 minutes. In one embodiment, the initialization step is omitted.

The second step, known as the denaturation step, allows the separation of the nucleic acid's double chain yielding two single-stranded DNA molecules.

In one embodiment, the denaturation step takes place at a temperature between 90-100 °C. In one embodiment, the denaturation step takes place at a temperature between 92-97 °C. In one embodiment, the denaturation step takes place at a temperature of 94 °C. In one embodiment, the denaturation step lasts 1 second to 1 minute. In one embodiment, the denaturation step lasts 20 to 40 seconds. In one embodiment, the denaturation step lasts 30 seconds.

The third step, known as the annealing and extension step, allows the binding of the primers with each of the single-stranded DNA templates, and the subsequent polymerisation carried out by the DNA polymerase. The temperature of the annealing and extension step must be low enough to allow for hybridization of the primer to the strand, but high enough for the hybridization to be specific, i.e., the primer should bind only to a perfectly complementary part of the strand, and nowhere else. If the temperature is too low, the primer may bind imperfectly. If it is too high, the primer may not bind at all. A typical annealing temperature is about 3–5°C below the melting point of the primers used. Stable hydrogen bonds between complementary bases are formed only when the primer sequence very closely matches the template sequence.

During this step, the polymerase binds to the primer-template hybrid and begins DNA formation.

The temperature at which the annealing and extension step is carried out depends on the RNA under investigation and the primers used. In one embodiment, the annealing and extension step takes place at a temperature between 55-65 °C. In one embodiment, the annealing and extension step takes place at a temperature between 60-65 °C. In one embodiment, the annealing and extension step takes place at a temperature of 63 °C. In one embodiment, the annealing and extension step lasts 10 seconds to 5 minutes. In one embodiment, the annealing and extension step lasts 20 seconds to 2 minutes. In one embodiment, the annealing and extension step lasts 40 to 80 seconds. In one embodiment, the annealing and extension step lasts 60 seconds.

The denaturation and annealing / extension steps may be repeated to the number of times necessary to allow the DNA to be sufficiently amplified. In one embodiment, the denaturation annealing and polymerisation steps are repeated 10 to 100 times. In one embodiment, the denaturation annealing and polymerisation steps are repeated 20 to 60 times. In one embodiment, the denaturation annealing and polymerisation steps are repeated 30 to 50 times. In one embodiment, the denaturation annealing and polymerisation steps are repeated 35 to 45 times. In one embodiment, the denaturation annealing and polymerisation steps are repeated 40 times.

In one embodiment, the PCR procedure includes an enzyme inactivation step subsequent to completion of the cycles of denaturation and annealing / extension steps. In one embodiment, the enzyme inactivation step is omitted.

When it forms part of the method of the present invention, in one embodiment, the enzyme inactivation step takes place at a temperature between 97-99 °C. In one embodiment, the enzyme inactivation step takes place at a temperature of 98 °C.

In the methods of the present invention to determine the identity, integrity and quantitative ratio of the RNA, in one embodiment, the enzyme inactivation step lasts 1 to 20 minutes. In one embodiment, the enzyme inactivation step lasts 5 to 15

minutes. In one embodiment, the enzyme inactivation step lasts 8 to 12 minutes. In one embodiment, the enzyme inactivation step lasts 10 minutes.

In the methods of the present invention to determine the potency of the RNA, in one embodiment, the enzyme inactivation step lasts 10 seconds to 5 minutes. In one embodiment, the enzyme inactivation step lasts 20 seconds to 2 minutes. In one embodiment, the enzyme inactivation step lasts 40 to 80 seconds. In one embodiment, the enzyme inactivation step lasts 60 seconds.

The temperatures and the timings used for each cycle depend on a wide variety of parameters, such as: the enzyme used to synthesize the DNA, the concentration of divalent ions and deoxyribonucleotides (dNTPs) in the reaction and the bonding temperature of the primers. The type of quantitative PCR technique used depends on the DNA sequence in the samples, the technique can either use non-specific fluorochromes or hybridization probes.

In one embodiment, the PCR methods used in the present invention employ a detectable label. A detectable label is a detectable compound which is attached directly or indirectly to another molecule. Within the method of the present invention the detectable label is attached to a single-stranded nucleic acid molecule. The skilled person knows methods for attaching labels to nucleic acid molecules. Specific, non-limiting examples of labels include fluorescent probes and fluorogenic moieties, chromogenic moieties, haptens, affinity tags and radioactive isotopes. The label can be detectable directly (e.g. optically) or indirectly (e.g. by interaction with one or more molecules which are in turn detectable).

Within the method of the present invention preferably fluorescent probes are used. As is known to the person skilled in the art, a fluorescent probe (also known as a fluorescent label or a fluorescent tag) is a molecule that is attached chemically to aid in the detection of a biomolecule such as a protein, antibody, or amino acid.

In one embodiment, the fluorescent probe carries one label. In one embodiment, the fluorescent probe carries multiple labels (such as two, three, or four labels).

The fluorescent probe is labelled with a fluorophore. As is known to the person skilled in the art, a fluorophore is a fluorescent substance which can re-emit light upon light excitation. The fluorophore selectively binds to a specific region or functional group on the target molecule and can be attached chemically or biologically. Examples of suitable fluorophores are known to those skilled in the art, and include [6-amino-9-[2-carboxy-4-[5-(2,5-dioxopyrrol-1-yl)pentyl-carbamoyl]phenyl]-4,5-disulfoxanthen-3-ylidene]azanium (Alexa 488), 6-carboxyfluorescein (FAM), 4-(2,7-difluoro-3-hydroxy-6-oxoxanthen-9-yl)benzene-1,3-dicarboxylic acid (Oregon Green), [6-amino-9-(2,5-dicarboxyphenyl)xanthen-3-ylidene]azanium (Rhodamine Green), 6-[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-Hexanoic acid (NBD-X), tetrachlorofluorescein (TET), [9-[6-(2,5-dioxopyrrolidin-1-yl)oxy-6-oxohexyl]-8,8-dimethyl-2-oxo-4-(trifluoromethyl)pyrano[3,2-g]quinolin-6-yl]methanesulfonate (Alexa 430), 3-(4,4-Difluoro-5-phenyl-3a,4a-diaza-4-bora-s-indacen-3-yl)propionic acid (BODIPY R6G-X), (6-carboxy-4',5'-dichloro-2',7'-dimethoxyfluorescein) (JOE), [2',5,5',6-tetrachloro-6'-(2,2-dimethylpropanoyloxy)-7'-[3-[6-[[di(propan-2-yl)amino]-(2-hydroxyethoxy)phosphanyl]oxyhexylamino]-3-oxopropyl]-4'-methyl-3-oxospiro[2-benzofuran-1,9'-xanthene]-3'-yl] 2,2-dimethylpropanoate (Yakima Yellow), 12-[4-(2,5-dioxopyrrolidin-1-yl)oxycarbonylphenyl]-7,8,8,16,16,17-hexamethyl-2-oxa-6,18-diazapenta-cyclo[11.7.0.0^{3,11}.0^{5,9}.0^{15,19}]icosa-1(20),3,5,9,11,13,15(19)-heptaene-4,20-disulfonic acid (Alexa 532), 2'-chloro-7'-phenyl-1,4-dichloro-6-carboxyfluorescein (VIC), and heaxchlorofluorescein (HEX). Of these, 5'-hexachlorofluorescein and 6-carboxyfluorescein are preferred.

The fluorescent probe is also labelled with a quencher. As is known to the person skilled in the art, a quencher is a substance which decreases the fluorescent intensity of a given substance. Examples of quenchers are known in the art, and include 2-[N-(2-hydroxyethyl)-4-[[2-methoxy-5-methyl-4-[(4-methyl-2-nitrophenyl)diazenyl]-phenyl]diazenyl]anilino]ethanol (Black Hole Quencher 1, BHQ1), ZEN quencher (available from IDT), Iowa black fluorescein quencher (IBFQ) (available from IDT) 4-N-methyl-N-(4'-nitro-2'-chloroazobenzen-4-yl)-aminobutanamido-1- (2-O-dimethoxytrityloxymethyl)-pyrrolidin-4-yl-succinoyl long chain alkylamino-CPG (ECLIPSE quencher), and 5-carboxy-N,N,N',N'-tetramethylrhodamine (TAMRA),

and mixtures of any thereof. Of these, BHQ1 and a combination of ZEN and IBFQ are preferred.

In one embodiment, the fluorophore is 5'-hexachlorofluorescein and the quencher is Black Hole Quencher 1. In one embodiment, the fluorophore is 6-carboxyfluorescein and the quencher is Black Hole Quencher 1. Both the fluorophore and the quencher are commercially available, for example from Eurofins Genomics.

The method of detecting the double-stranded nucleic acid molecules produced in the PCR based assay of the invention depends on the detectable label attached to the nucleic acid probe. For example, if the nucleic acid probe is labelled with a radioactive isotope, the double-stranded nucleic acid molecules are detected by autoradiography. If the nucleic acid probe is labelled with a fluorescent probe, the double-stranded nucleic acid molecules are detected by fluorescence spectroscopy.

Digital PCR (dPCR)

In one embodiment, the PCR is digital PCR (dPCR). In one embodiment, the dPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

In digital PCR (dPCR), respective PCR reactions are partitioned into multiple smaller reactions so that individual nucleic acid molecules within the sample are localized and amplified in many separate regions. Micro well plates, capillaries, oil emulsion (droplets), and arrays of miniaturized chambers with nucleic acid binding surfaces can be used to separate the RNA sample in multiple small reactions per reaction vessel. After multiple PCR amplification cycles, typically performed under saturating PCR conditions, the samples are analyzed for fluorescence signals with a binary readout of "0" (no signal) or "1" (signal). Using Poisson's law of small numbers, the distribution of target molecule within the sample can be accurately approximated allowing for a quantification of the target strand in the PCR product. Therefore, dPCR is not dependent on the number of amplification cycles to determine the initial sample amount, and thus eliminates the reliance on uncertain exponential data to quantify target nucleic acids. Therefore, dPCR allows absolute quantification of nucleic acids.

Suitable commercial systems that may be used to perform dPCR comprise chip-based QuantStudio™ 3D digital PCR System (Thermo Fisher, Waltham, MA, USA), Rain Drop Plus™ system (RainDance Technologies, Lexington, MA, USA), or QX200™, AutoDG™, and Droplet Digital™ PCR System (BioRad Laboratories, Hercules, CA, USA).

Chip-based digital PCR measures absolute quantities by counting nucleic acid molecules partitioned in independent reaction wells. Commercially available systems for chip-based dPCR comprise QuantStudio™ 3D digital PCR System (Thermo Fisher, Waltham, MA, USA). A PCR reaction is divided into around 20,000 independent reaction wells on a chip that either contain or not contain template. The sealed chip is then subjected to PCR amplification. Each well containing template is leading to PCR positive signals (positive well) and each well that is not containing template is leading to negative PCR signals (negative wells). Positive and negative wells are counted allowing quantitation of template concentration using Poisson distribution algorithm.

Droplet Digital PCR (ddPCR)

In one embodiment, the PCR is droplet digital PCR (ddPCR). In one embodiment, the ddPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

Droplet digital PCR (ddPCR) measures absolute quantities by counting nucleic acid molecules encapsulated in discrete, volumetrically defined water-in-oil droplet partitions. Commercially available systems for ddPCR comprise Droplet Digital™ PCR System (BioRad Laboratories, Hercules, CA, USA) or Rain Drop Plus™ system (RainDance Technologies, Lexington, MA, USA). A PCR reaction is divided into around 20,000 droplets that either contain or not contain template leading to PCR positive and negative droplets that are counted allowing quantitation of template concentration using Poisson distribution algorithm. The droplets are generated using a droplet generator.

ddPCR enables more precise absolute quantification as compared to classic relative quantitation which is limited due to the doubling during each cycle. ddPCR has several advantages over conventional methods (e.g. qPCR) because no standard curve is needed for quantification, it is a more robust method (even sub-optimal primer pairs that lead to false positive or false negative signals will be eventually give a correct concentration due to the Poisson distribution algorithm), and enables precise (diagnostic) quantification (Resolution of standard qPCR: 0.5 cycles (+/-50%) vs. 10% for ddPCR (using acoustic pipetting 1.5% accuracy possible).

Quality parameters

The method according to the present invention is suitable for determining at least one quality parameter. The term "quality parameter" comprises any parameter of the RNA sample which is related to a property of the RNA sample and typically obtained for quality control during or following production. Examples of quality parameters are quantity of the one or more coding RNA molecules of at least two RNA molecule species, integrity of the one or more coding RNA molecules of at least two RNA molecule species and quantitative ratio between the one or more coding RNA molecules of at least two RNA molecule species.

The quality parameter is useful to determine, e.g., whether an RNA sample comprises all required *n* different RNA molecule species, the quantitative ratio between at least two of the *n* different RNA molecule species, whether the RNA molecules of the *n* different RNA molecule species are present in intact form (integrity) and to determine the amount of each RNA molecule species in the RNA sample. By determining at least one quality control parameter, an RNA sample can be analyzed and rated according to regulatory requirements as necessary for marketing approval of a medicinal product.

Method of determining Quantitative Ratio of RNA Molecules

In one aspect the quality parameter measured is the quantitative ratio of two or more RNA molecule species in an RNA sample. In this sense the quantitative ratio may be between the one or more coding RNA molecules of at least two RNA molecule

species: The quantitative ratio of the one or more coding RNA molecules of at least two RNA molecule species may be determined by determining the quantity of the one or more coding RNA molecules of at least two RNA molecule species and evaluating the quantitative ratio of both or more quantities. Since the method allows determining the quantity of the one or more coding RNA molecules of all RNA molecule species, it is possible to determine the quantitative ratio between all RNA molecule species.

Therefore, in one aspect, the invention provides a method of determining the quantitative ratio of two or more RNA molecule species in an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 2, the method comprising the steps of:
a) reverse transcription of the *n* RNA molecule species in the RNA sample into cDNA molecules of *n* DNA molecule species; and
b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set and a single second primer, wherein
the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region of only one of the *n* DNA molecule species in the sample, and
the single second primer is capable of annealing to a second target region of all of the *n* DNA molecule species in the sample.

In one embodiment, the first target region is within the coding sequence of the DNA molecule. In one embodiment, the first target region spans the coding sequence and the untranslated region of the DNA molecule.

In one embodiment, the first target region is at the 3'-end region of the DNA molecule. In one embodiment, the first target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the first target region spans the coding sequence and the 3'-untranslated region (as defined herein) of the DNA molecule. In one embodiment, the first target region is at the 5'-end region of the DNA molecule. In one embodiment, the first target region is within the coding sequence at the 5'-end region of the DNA molecule. In one embodiment, the first

target region spans the coding sequence and the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the second target region is within the coding sequence of the DNA molecule. In one embodiment, the second target region spans the coding sequence and the untranslated region of the DNA molecule. In one embodiment, the second target region is within the untranslated region of the DNA molecule.

In one embodiment, the second target region is at the 3'-end region of the DNA molecule. In one embodiment, the second target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the second target region spans the coding sequence and the 3'-untranslated region (as defined herein) of the DNA molecule. In one embodiment, the second target region is at the 5'-end region of the DNA molecule. In one embodiment, the second target region is within the coding sequence at the 5'-end region of the DNA molecule. In one embodiment, the second target region spans the coding sequence and the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the PCR is digital PCR (dPCR). In one embodiment, the dPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

In one embodiment, the PCR is droplet digital PCR (ddPCR). In one embodiment, the ddPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

Identity of RNA Molecules

In one aspect the quality parameter measured is the identity of *n* RNA molecule species in an RNA sample. In this aspect, the invention provides a method of determining identity of *n* RNA molecule species in an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 2, the method comprising the steps of:

a) reverse transcription of the *n* RNA molecule species in the RNA sample into cDNA

molecules of *n* DNA molecule species; and
b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set and a single second primer, wherein
the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region of only one of the *n* DNA molecule species in the sample, and
the single second primer is capable of annealing to a second target region of all of the *n* DNA molecule species in the sample.

In one embodiment, the first target region is within the coding sequence of the DNA molecule. In one embodiment, the first target region spans the coding sequence and the untranslated region of the DNA molecule.

In one embodiment, the first target region is at the 3'-end region of the DNA molecule. In one embodiment, the first target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the 3'-untranslated region (as defined herein) of the DNA molecule. In one embodiment, the first target region is at the 5'-end region of the DNA molecule. In one embodiment, the first target region is within the coding sequence at the 5'-end region of the DNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the second target region is within the coding sequence of the DNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the untranslated region of the DNA molecule. In one embodiment, the second target region is within the untranslated region of the DNA molecule

In one embodiment, the second target region is at the 3'-end region of the DNA molecule. In one embodiment, the second target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the 3'-untranslated region

(as defined herein) of the DNA molecule. In one embodiment, the second target region is at the 5'-end region of the DNA molecule. In one embodiment, the second target region is within the coding sequence at the 5'-end region of the DNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the PCR is digital PCR (dPCR). In one embodiment, the dPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

In one embodiment, the PCR is droplet digital PCR (ddPCR). In one embodiment, the ddPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

Integrity of RNA Molecules

In one aspect the quality parameter measured is the integrity (as defined herein) of *n* RNA molecule species in an RNA sample. Therefore, in one aspect, the invention provides a method of determining integrity of an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 1,
the method comprising the steps of:
a) reverse transcription of the *n* RNA molecules in the RNA sample into cDNA molecules of *n* DNA molecule species; and
b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set, a single second primer, a third primer set and a single fourth primer, wherein
the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region at the 3'-end region of a DNA molecule species in the sample,
the single second primer is capable of annealing to a second target region at the 3'-end region of all of the *n* DNA molecule species in the sample;
the third primer set comprises *n* primer species, wherein each primer species is capable of annealing to a third target region at the 5'-end region of a DNA molecule species in the sample; and

the single fourth primer set is capable of annealing to a fourth target region at the 5'-end region of all of the *n* DNA molecule species in the sample.

In one embodiment, the first target region is within the coding sequence of the DNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the untranslated region of the DNA molecule.

In one embodiment, the first target region is at the 3'-end region of the DNA molecule. In one embodiment, the first target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the 3'-untranslated region (as defined herein) of the DNA molecule. In one embodiment, the first target region is at the 5'-end region of the DNA molecule. In one embodiment, the first target region is within the coding sequence at the 5'-end region of the DNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the second target region is within the coding sequence of the DNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the untranslated region of the DNA molecule. In one embodiment, the second target region is within the untranslated region of the DNA molecule.

In one embodiment, the second target region is at the 3'-end region of the DNA molecule. In one embodiment, the second target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the 3'-untranslated region (as defined herein) of the DNA molecule. In one embodiment, the second target region is at the 5'-end region of the DNA molecule. In one embodiment, the second target region is within the coding sequence at the 5'-end region of the DNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the third target region is within the coding sequence of the DNA molecule. In one embodiment, the third target region includes a part of the coding sequence and a part of the untranslated region of the DNA molecule.

In one embodiment, the third target region is at the 3'-end region of the DNA molecule. In one embodiment, the third target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the third target region includes a part of the coding sequence and a part of the 3'-untranslated region (as defined herein) of the DNA molecule. In one embodiment, the third target region is at the 5'-end region of the DNA molecule. In one embodiment, the third target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the third target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the fourth target region is within the coding sequence of the DNA molecule. In one embodiment, the fourth target region includes a part of the coding sequence and a part of the untranslated region of the DNA molecule. In one embodiment, the fourth target region is within the untranslated region of the DNA molecule.

In one embodiment, the fourth target region is at the 3'-end region of the DNA molecule. In one embodiment, the fourth target region is within the coding sequence at the 3'-end region of the DNA molecule. In one embodiment, the fourth target region includes a part of the coding sequence and a part of the 3'-untranslated region (as defined herein) of the DNA molecule. In one embodiment, the fourth target region is at the 5'-end region of the DNA molecule. In one embodiment, the fourth target region is within the coding sequence at the 5'-end region of the DNA molecule. In one embodiment, the fourth target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the DNA molecule.

In one embodiment, the PCR is digital PCR (dPCR). In one embodiment, the dPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

In one embodiment, the PCR is droplet digital PCR (ddPCR). In one embodiment, the ddPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

Potency of Formulated RNA Sample

In one aspect the quality parameter measured is the potency of a formulated RNA sample comprising RNA molecules of interest. In this aspect, the invention provides a method of determining the potency of a formulated RNA sample comprising RNA molecules of interest, the method comprising the steps of:
a) providing an RNA sample which has been isolated from cells transfected with a formulated RNA sample;
b) reverse transcription of the RNA molecules in the RNA sample into cDNA molecules;
c) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer, a second primer, a third primer and a fourth primer, wherein
the first primer and the second primer are capable of annealing to a first target region and a second target region of the cDNA molecules produced from the RNA of interest in the sample, and
the third primer and the fourth primer are capable of annealing to a first target region and a second target region of the cDNA molecules derived from an endogenous RNA in the sample; and
d) comparing the measured amount of the cDNA produced from the RNA molecules of interest with the measured amount of cDNA produced from endogenous RNA.

In one embodiment, the first target region is within the coding sequence of the cDNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the untranslated region of the cDNA molecule.

In one embodiment, the first target region is at the 3'-end region of the cDNA molecule. In one embodiment, the first target region is within the coding sequence at the 3'-end region of the cDNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the 3'-untranslated region (as

defined herein) of the cDNA molecule. In one embodiment, the first target region is at the 5'-end region of the cDNA molecule. In one embodiment, the first target region is within the coding sequence at the 3'-end region of the cDNA molecule. In one embodiment, the first target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the cDNA molecule.

In one embodiment, the second target region is within the coding sequence of the cDNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the untranslated region of the cDNA molecule.

In one embodiment, the second target region is at the 3'-end region of the cDNA molecule. In one embodiment, the second target region is within the coding sequence at the 3'-end region of the cDNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the 3'-untranslated region (as defined herein) of the cDNA molecule. In one embodiment, the second target region is at the 5'-end region of the cDNA molecule. In one embodiment, the second target region is within the coding sequence at the 5'-end region of the cDNA molecule. In one embodiment, the second target region includes a part of the coding sequence and a part of the 5'-untranslated region (as defined herein) of the cDNA molecule.

In one embodiment, the PCR is digital PCR (dPCR). In one embodiment, the dPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

In one embodiment, the PCR is droplet digital PCR (ddPCR). In one embodiment, the ddPCR-based assay uses a detectable label. In one embodiment, the detectable label is a fluorescent probe.

In one embodiment, the endogenous RNA is RNA expressed from a housekeeping gene in the sample.

In one embodiment, the endogenous RNA is RNA expressed from a GAPDH gene in the sample.

In one embodiment, the formulated RNA composition comprises two or more different RNA molecule species of interest and the PCR-based assay uses two or more different primer pairs which are each specific for one of the RNA molecule species of interest.

In one embodiment, steps a) to d) are repeated for further RNA samples which have been isolated from cells transfected with formulated RNA compositions comprising the RNA molecule(s) of interest to determine a potency level for each formulated RNA composition.

In one embodiment, carrying out the method establishes an expected potency level for a particular formulated RNA composition comprising a RNA of interest, the expected potency level being defined as a reference potency level.

In one embodiment, the method is carried out on a further formulated RNA composition comprising a RNA of interest and comparing it with the reference potency level for the formulated RNA composition.

In one embodiment, the method further comprises the following step z) prior to step a):
z) isolating/purifying RNA from cells which have been transfected with the formulated RNA sample.

In one embodiment, the method further comprises the following step y) prior to step z):
y) transfecting the cell with the formulated RNA composition.

In one embodiments the cells are transfected using RNA which is complexed with at least one carrier compound, thereby forming at least one RNA-carrier complex.

RNA Carriers

In one embodiment, at least one RNA of the RNA molecules present in the sample is complexed with at least one carrier compound, thereby forming at least one RNA-carrier complex.

In one embodiment, the at least one carrier compound is a member selected from the group consisting of lipids, phospholipids, PEGylated lipids, cationic and polycationic compounds, and combinations thereof.

In one embodiment, at least one RNA-carrier complex is selected from the group consisting of a liposome, a lipid nanoparticle, a lipoplex, and a mixture thereof.

Particles

RNA of the invention may be present in particles comprising (i) the RNA, and (ii) at least one cationic or cationically ionizable compound such as a polymer or lipid complexing the RNA. 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.

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.
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.

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

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.
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.

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.

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).

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.

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.

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.

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.

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.

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.

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.

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.

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.

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

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

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

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

under reduced pressure yields a milky gel that turns subsequently into a liposomal suspension.

The term "ethanol injection technique" refers to a process, in which an ethanol solution comprising lipids is rapidly injected into an aqueous solution through a needle. This action disperses the lipids throughout the solution and promotes lipid structure formation, for example lipid vesicle formation such as liposome formation. Generally, the 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.

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

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

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.

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.

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.

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.

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.

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

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

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

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.

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

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

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

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

Cationic polymers (including polycationic polymers) contemplated for use herein include any cationic polymers which are able to electrostatically bind nucleic acid. In one embodiment, cationic polymers contemplated for use herein include any cationic polymers with which nucleic acid can be associated, e.g. by forming complexes with the nucleic acid or forming vesicles in which the nucleic acid is enclosed or encapsulated.

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

Lipids

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.

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.

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.

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.

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.

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.

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

The term "lipid-like material", "lipid-like compound" or "lipid-like molecule" relates to substances, 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.

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.

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

Glycerolipids are composed of mono-, di-, and tri-substituted glycerols, the best-known being the fatty acid triesters of glycerol, called triglycerides. The word "triacylglycerol" is sometimes used synonymously with "triglyceride". In these compounds, the three hydroxyl groups of glycerol are each esterified, typically by different fatty acids. Additional subclasses of glycerolipids are represented by glycosylglycerols, which are characterized by the presence of one or more sugar residues attached to glycerol via a glycosidic linkage.

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

Sphingolipids are a complex family of compounds that share a common structural feature, a sphingoid base backbone. The major sphingoid base in mammals is commonly referred to as sphingosine. Ceramides (N-acyl-sphingoid bases) are a major subclass of sphingoid base derivatives with an amide-linked fatty acid. The fatty acids are typically saturated or mono-unsaturated with chain lengths from 16 to

26 carbon atoms. The major phosphosphingolipids of mammals are sphingomyelins (ceramide phosphocholines), whereas insects contain mainly ceramide phosphoethanolamines and fungi have phytoceramide phosphoinositols and mannose-containing headgroups. The glycosphingolipids are a diverse family of molecules composed of one or more sugar residues linked via a glycosidic bond to the sphingoid base. Examples of these are the simple and complex glycosphingolipids such as cerebrosides and gangliosides.
Sterol lipids, such as cholesterol and its derivatives, or tocopherol and its derivatives, are an important component of membrane lipids, along with the glycerophospholipids and sphingomyelins.

Saccharolipids describe compounds in which fatty acids are linked directly to a sugar backbone, forming structures that are compatible with membrane bilayers. In the saccharolipids, a monosaccharide substitutes for the glycerol backbone present in glycerolipids and glycerophospholipids. The most familiar saccharolipids are the acylated glucosamine precursors of the Lipid A component of the lipopolysaccharides in Gram-negative bacteria. Typical lipid A molecules are disaccharides of glucosamine, which are derivatized with as many as seven fatty-acyl chains. The minimal lipopolysaccharide required for growth in E. coli is Kdo2-Lipid A, a hexa-acylated disaccharide of glucosamine that is glycosylated with two 3-deoxy-D-manno-octulosonic acid (Kdo) residues.

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

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

Cationic/cationically ionizable lipids

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.

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.

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.

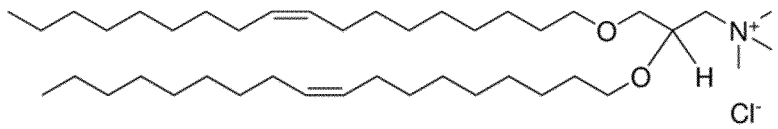
As used herein, a "cationically ionizable lipid" refers to a lipid or lipid-like material which has a net positive charge or is neutral, *i.e.*, which is not permanently cationic. Thus, depending on the pH of the composition in which the cationically ionizable lipid is solved, the cationically ionizable 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.

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.

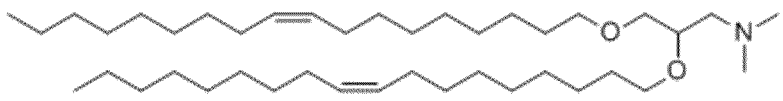
Examples of cationic or cationically ionizable lipids include, but are not limited to N,N-dimethyl-2,3-dioleyloxypropylamine (DODMA), 1,2-dioleoyl-3-trimethylammonium propane (DOTAP); 1,2-di-O-octadecenyl-3-trimethylammonium propane (DOTMA), 3-(N—(N′,N′-dimethylaminoethane)-carbamoyl)cholesterol (DC-Chol), 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), l,2-dimyristoyl-3-trimethylammonium propane (DMTAP), 1,2-dioleyloxypropyl-3-dimethyl-hydroxyethyl ammonium bromide (DORIE), and 2,3-dioleoyloxy- N-[2(spermine carboxamide)ethyl]-N,N-dimethyl-l-propanamium trifluoroacetate (DOSPA), 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLinDMA), 1,2-dilinolenyloxy-N,N-dimethylaminopropane (DLenDMA), dioctadecylamidoglycyl spermine (DOGS), 3-dimethylamino-2-(cholest-5-en-3-beta-oxybutan-4-oxy)-1-(cis,cis-9,12-oc-tadecadienoxy)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_{12}-5), 1-[2-[bis(2-hydroxydodecyl)amino]ethyl-[2-[4-[2-[bis(2 hydroxydodecyl)amino]ethyl]piperazin-1-yl]ethyl]amino]dodecan-2-ol (lipidoid C12-200).

In some embodiments, the cationic or cationically ionizable lipid is DOTMA. In some embodiments, the cationic or cationically ionizable lipid is DODMA.
DOTMA is a cationic lipid with a quarternary amine headgroup. The structure of DOTMA may be represented as follows:



DODMA is an ionizable cationic lipid with a tertiary amine headgroup. The structure of DODMA may be represented as follows:



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

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.

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.

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.

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.

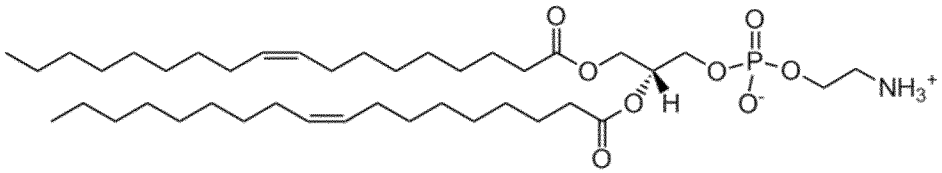
In some embodiments, the one or more additional lipids comprised in the nucleic acid particles (especially in the particles comprising mRNA) described herein comprise one or more of the following: neutral lipids, steroids, and combinations thereof. In some embodiments, the one or more additional lipids 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.

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.

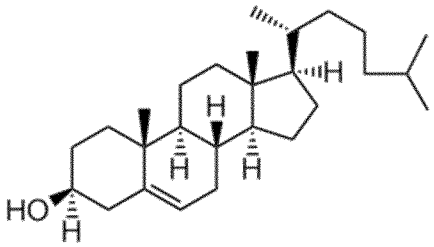
Thus, in some embodiments, the nucleic acid particles (especially the particles comprising mRNA) described herein comprise (1) a cationic or cationically ionizable lipid, and a phospholipid such as DOPE or (2) a cationic or cationically ionizable lipid and a phospholipid such as DOPE and cholesterol.

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.

DOPE is a neutral phospholipid. The structure of DOPE may be represented as follows:



The structure of cholesterol may be represented as follows:



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.

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.

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

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.

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*.

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.

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

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).

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.

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.

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

about 1:1. In an exemplary embodiment, the molar ratio of the at least one cationic lipid to the at least one additional lipid is about 2:1.

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.

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.

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.

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.

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)

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. 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.

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.

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.

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.

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.

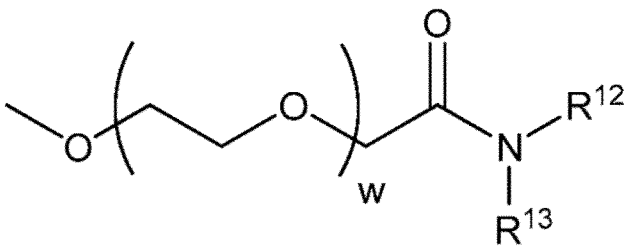
In some embodiments, the LNP comprises from 1 to 10 mol percent, from 1 to 5 mol percent, or from 1 to 2.5 mol percent of the polymer-conjugated lipid.

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. 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.

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.

In some embodiments, the steroid is cholesterol.

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

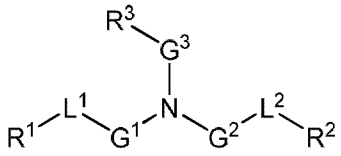


or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof, wherein: R^{12} and R^{13} are each independently a straight or branched, saturated or unsaturated alkyl chain containing from 10 to 30 carbon atoms, wherein the alkyl chain is

optionally interrupted by one or more ester bonds; and w has a mean value ranging from 30 to 60. In some embodiments, R^{12} and R^{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^{12} and R^{13} are each independently a straight, saturated alkyl chain containing about 14 carbon atoms, and w has a mean value of about 45.

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

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



(III)

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

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

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

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

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

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

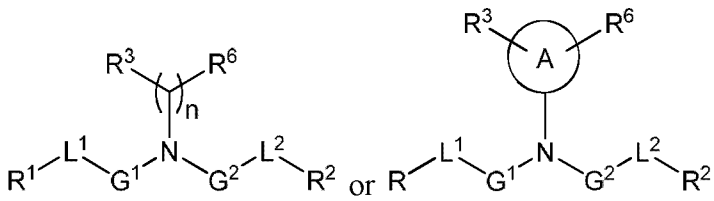
R^3 is H, OR^5, CN, -$C(=O)OR^4$, -$OC(=O)R^4$ or –$NR^5C(=O)R^4$;

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

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

x is 0, 1 or 2.

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



(IIIA) (IIIB)

wherein:

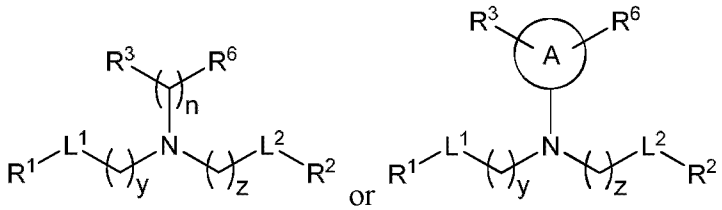
A is a 3 to 8-membered cycloalkyl or cycloalkylene ring;

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

n is an integer ranging from 1 to 15.

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

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

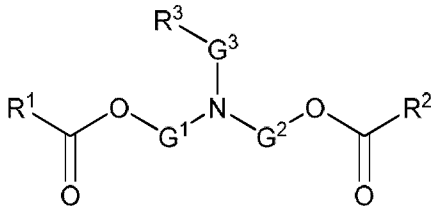
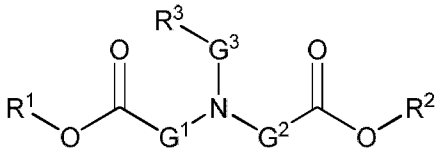


(IIIC) (IIID)

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

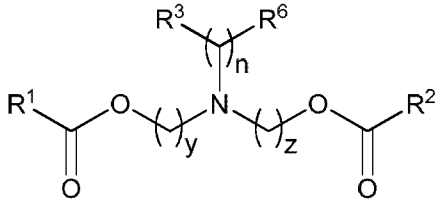
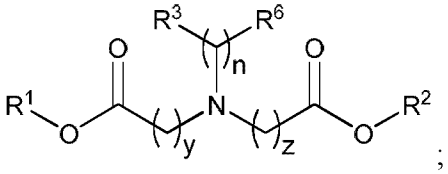
In any of the foregoing embodiments of Formula (III), one of L^1 or L^2 is -O(C=O)-. For example, in some embodiments each of L^1 and L^2 are -O(C=O)-. In some different embodiments of any of the foregoing, L^1 and L^2 are each independently -(C=O)O- or -O(C=O)-. For example, in some embodiments each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

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

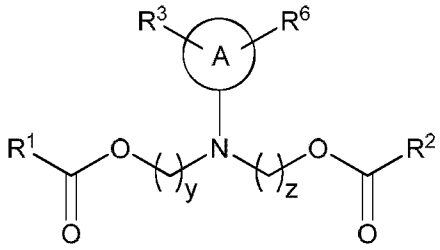
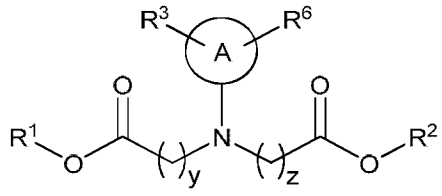



(IIIE) or (IIIF)

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

(IIIG) (IIIH)

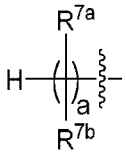
(IIII) or (IIIJ)

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

In some other of the foregoing embodiments of Formula (III), y and z are each independently an integer ranging from 2 to 10. For example, in some embodiments, y and z are each independently an integer ranging from 4 to 9 or from 4 to 6.
In some of the foregoing embodiments of Formula (III), R^6 is H. In other of the foregoing embodiments, R^6 is C_1-C_{24} alkyl. In other embodiments, R^6 is OH.

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

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



,

wherein:

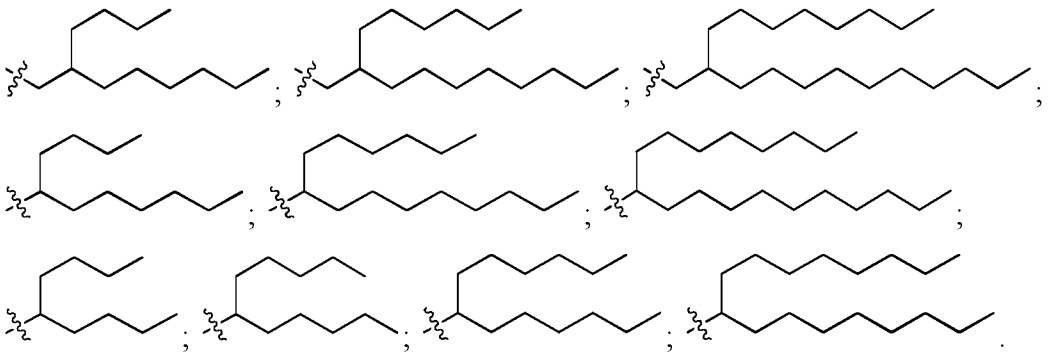
R^{7a} and R^{7b} are, at each occurrence, independently H or C_1-C_{12} alkyl; and

a is an integer from 2 to 12,

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

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

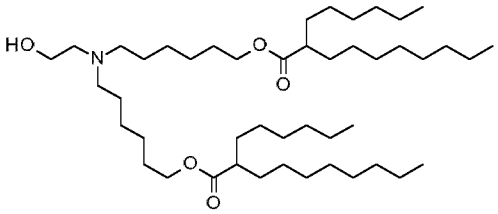
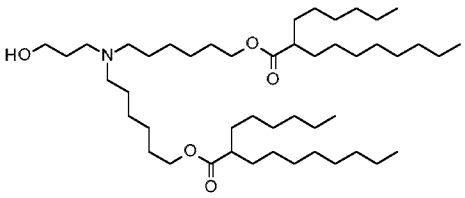
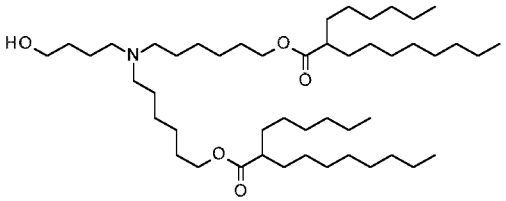
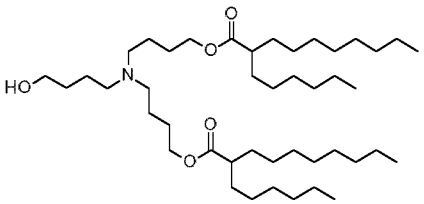
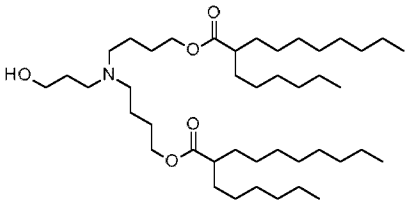
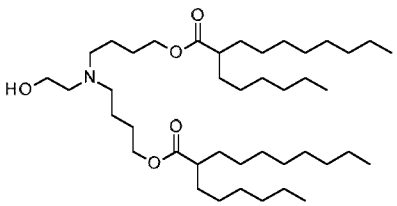
In different embodiments of Formula (III), R^1 or R^2, or both, has one of the following structures:

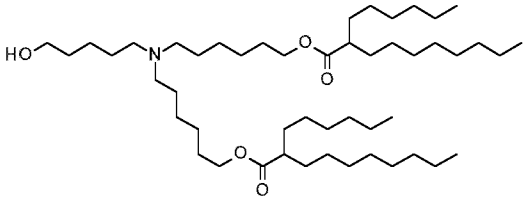
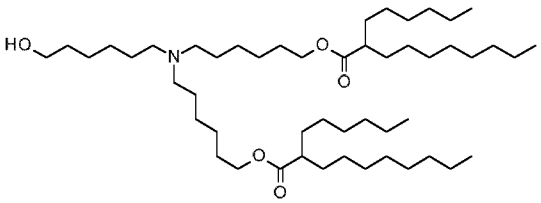
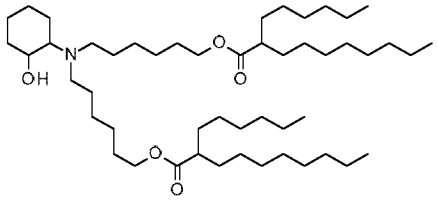
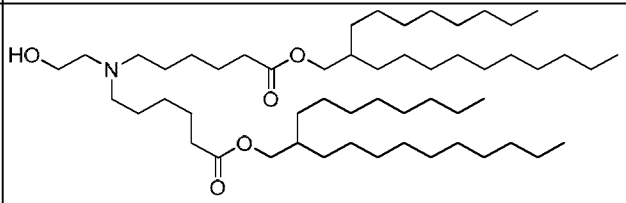
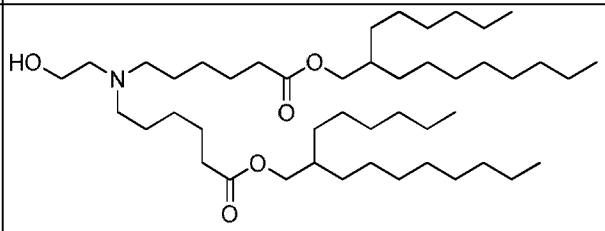
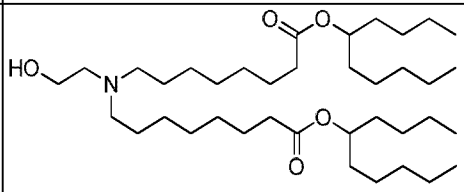
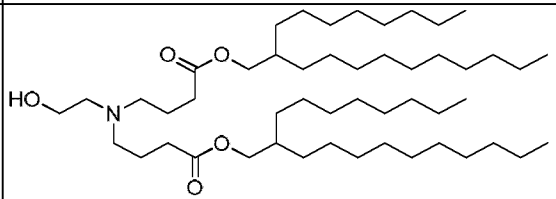
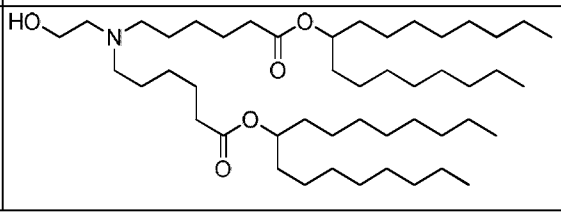


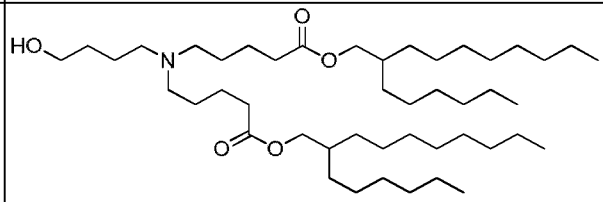
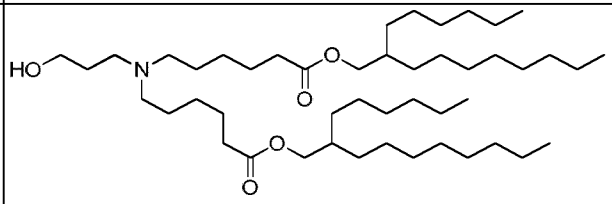
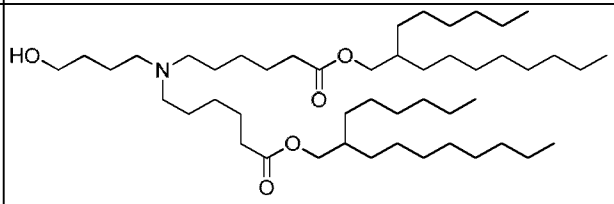
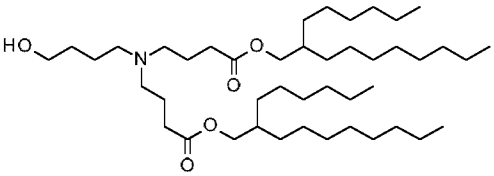
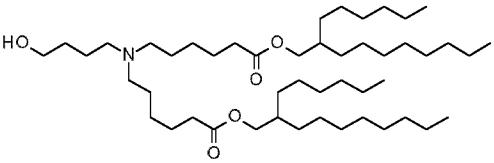
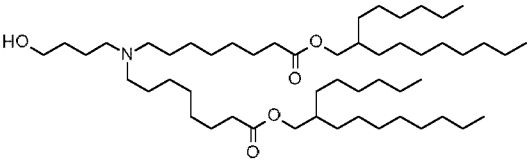
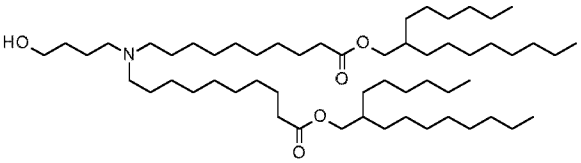
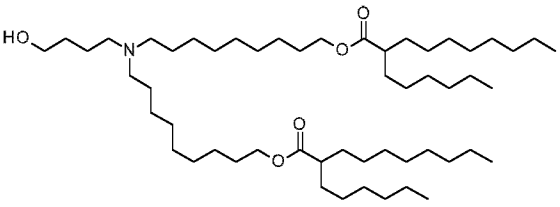
In some of the foregoing embodiments of Formula (III), R^3 is OH, CN, -C(=O)OR^4, -OC(=O)R^4 or –NHC(=O)R^4. In some embodiments, R^4 is methyl or ethyl.

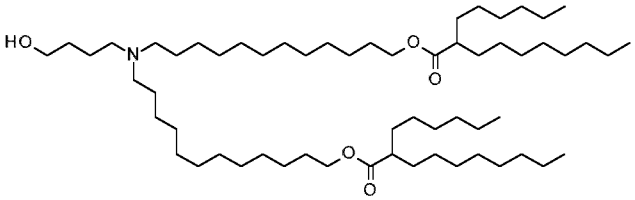
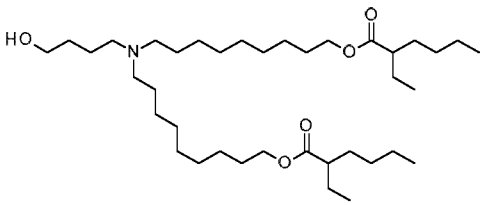
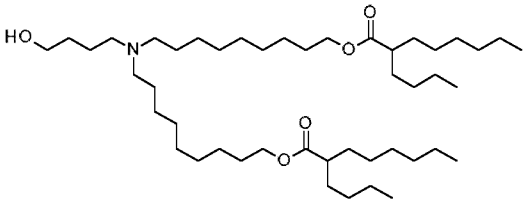
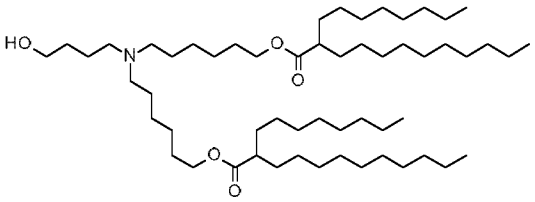
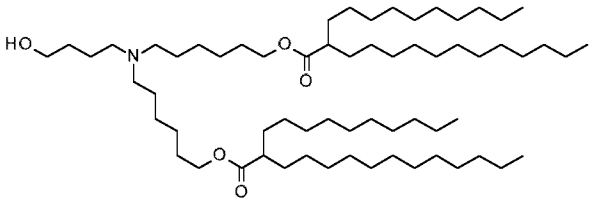
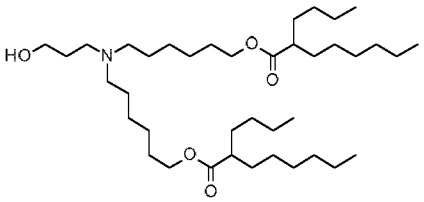
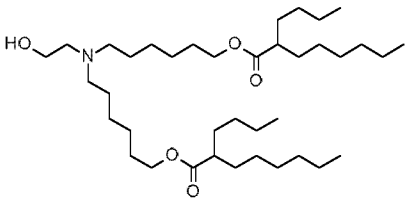
In various different embodiments, the cationic lipid of Formula (III) has one of the structures set forth in the table below.

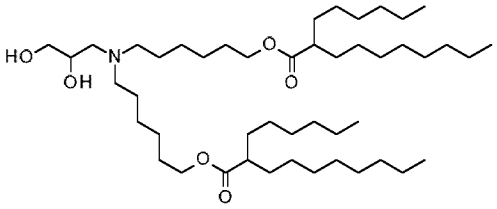
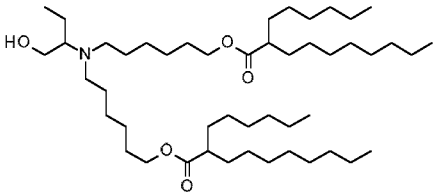
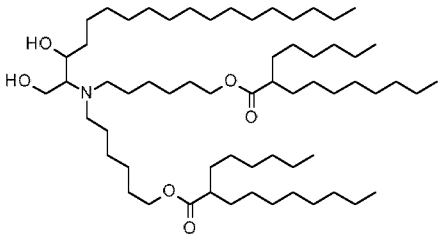
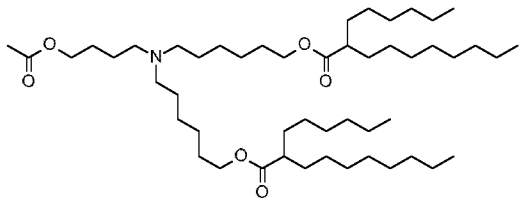
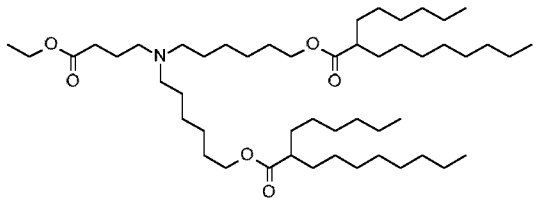
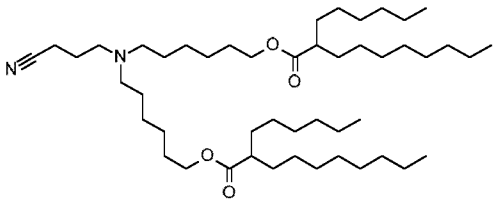
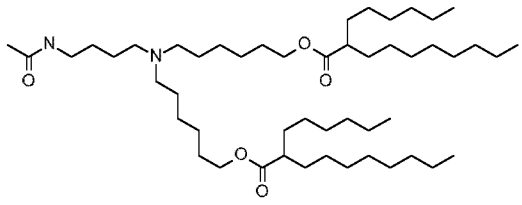
Representative Compounds of Formula (III).

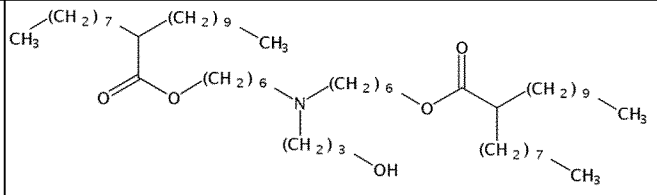
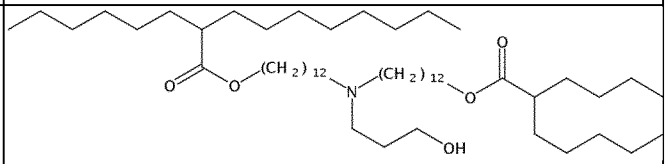
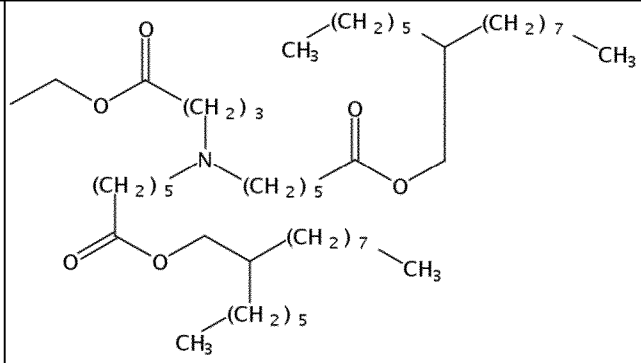
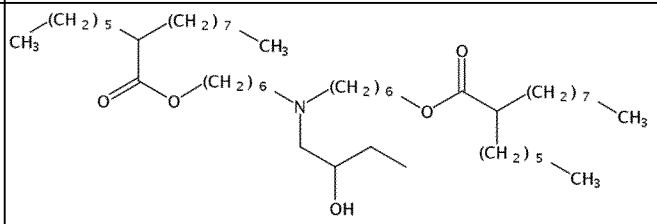
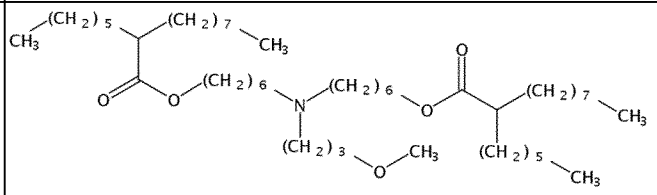
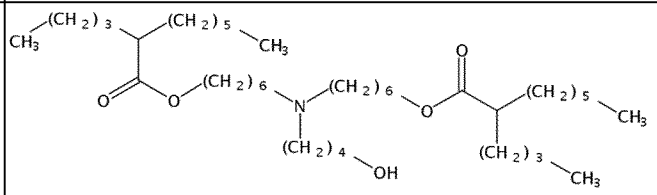
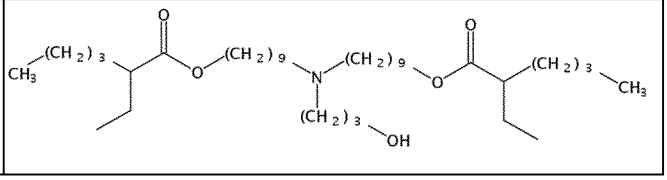
No.	Structure
III-1	
III-2	
III-3	
III-4	
III-5	
III-6	

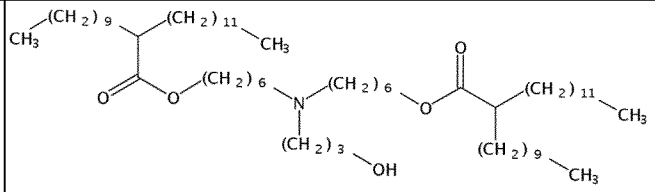
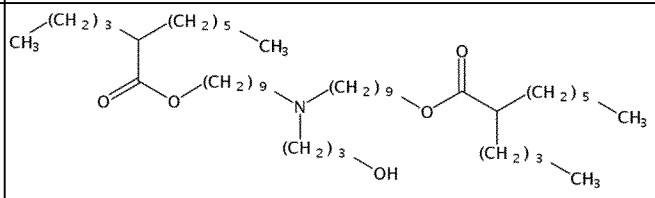
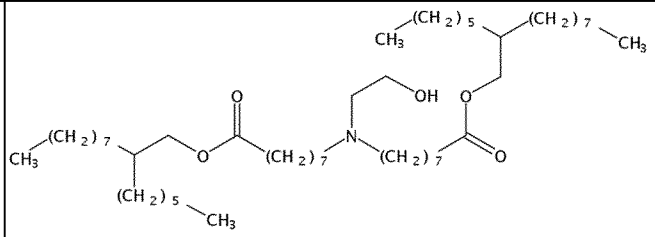
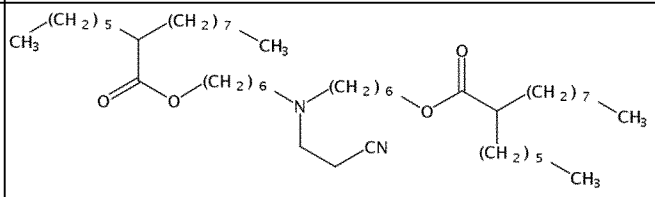
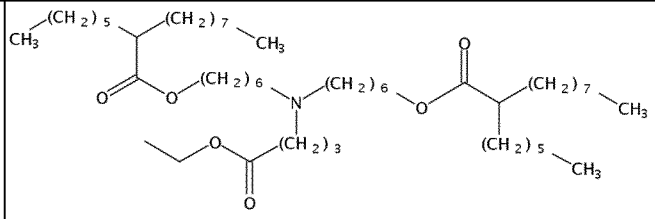
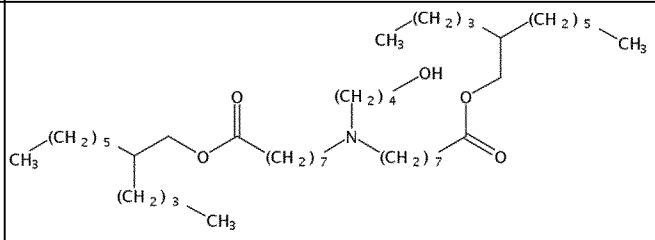
No.	Structure
III-7	
III-8	
III-9	
III-10	
III-11	
III-12	
III-13	
III-14	

No.	Structure
III-15	
III-16	
III-17	
III-18	
III-19	
III-20	
III-21	
III-22	

No.	Structure
III-23	
III-24	
III-25	
III-26	
III-27	
III-28	
III-29	

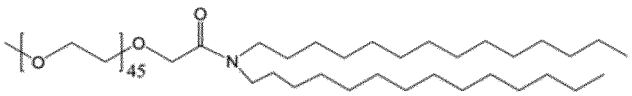
No.	Structure
III-30	
III-31	
III-32	
III-33	
III-34	
III-35	
III-36	

No.	Structure
III-37	
III-38	
III-39	
III-40	
III-41	
III-42	
III-43	

No.	Structure
III-44	
III-45	
III-46	
III-47	
III-48	
III-49	

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). 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, which has the structure shown below.



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.

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.

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.

Pharmaceutical compositions comprising nucleic acid particles

In some embodiments, the composition comprising nucleic acid molecules of the present invention, such as that produced by the methods of the present invention, is a pharmaceutical composition. In an embodiment, this composition may comprise salts, buffers, or other components as further described below. In the most preferred embodiment, the composition is that which comprises two or more different RNA molecules.

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).
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.
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.

Compositions described herein may also comprise a cryoprotectant 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.

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

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.

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.

The term "trisaccharide" means three sugars linked together to form one molecule. Examples of a trisaccharides include raffinose and melezitose.
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.

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 γ cyclodextrin.

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 γ cyclodextrin) 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.

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).

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. Exemplary polysaccharide cryoprotectants include glycogen, amylase, cellulose, dextran, maltodextrin and the like.

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.

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.

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 molar 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.

In an alternative embodiment, the nucleic acid (such as RNA and/or DNA) particle compositions described herein do not comprise a chelating agent.

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.

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.
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).

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).

The pharmaceutical compositions according to the present disclosure are generally applied in a "pharmaceutically effective amount" and in "a pharmaceutically acceptable preparation".

The term "pharmaceutically acceptable" refers to the non-toxicity of a material which does not interact with the action of the active component of the pharmaceutical composition.

The term "pharmaceutically effective amount" 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 severity 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 several of such parameters. In the case that a reaction in a patient is insufficient with an initial dose, higher doses (or effectively higher doses achieved by a different, more localized route of administration) may be used.

The pharmaceutical compositions of the present disclosure may contain 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.

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

The term "excipient" as used herein refers to a substance which may be present in a pharmaceutical composition of the present disclosure but is not an active ingredient. Examples of excipients, include without limitation, carriers, binders, diluents,

lubricants, thickeners, surface active agents, preservatives, stabilizers, emulsifiers, buffers, flavoring agents, or colorants.

The term "diluent" relates a diluting and/or thinning agent. Moreover, the term "diluent" includes any one or more of fluid, liquid or solid suspension and/or mixing media. Examples of suitable diluents include ethanol, glycerol and water.

The term "carrier" refers to a component which may be natural, synthetic, organic, inorganic in which the active component is combined in order to facilitate, enhance or enable administration of the pharmaceutical composition. A carrier as used herein may be one or more compatible solid or liquid fillers, diluents or encapsulating substances, which are suitable for administration to subject. Suitable carrier include, without limitation, sterile water, Ringer, Ringer lactate, sterile sodium chloride solution, isotonic saline, polyalkylene glycols, hydrogenated naphthalenes and, in particular, biocompatible lactide polymers, lactide/glycolide copolymers or polyoxyethylene/polyoxy-propylene copolymers. In some embodiments, the pharmaceutical composition of the present disclosure includes isotonic saline. Pharmaceutically acceptable carriers, excipients or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (A. R Gennaro edit. 1985). Pharmaceutical carriers, excipients or diluents can be selected with regard to the intended route of administration and standard pharmaceutical practice.

Routes of administration of pharmaceutical compositions

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.

Uses of pharmaceutical compositions

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.

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.

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.

The term "therapeutic treatment" relates to any treatment which improves the health status and/or prolongs (increases) the lifespan of an individual. Said treatment may eliminate the disease in an individual, arrest or slow the development of a disease in an individual, inhibit or slow the development of a disease in an individual, decrease the frequency or severity of symptoms in an individual, and/or decrease the recurrence in an individual who currently has or who previously has had a disease.

The terms "prophylactic treatment" or "preventive treatment" relate to any treatment that is intended to prevent a disease from occurring in an individual. The terms "prophylactic treatment" or "preventive treatment" are used herein interchangeably.

The terms "individual" and "subject" are used herein interchangeably. They refer to a human or another mammal (*e.g.*, mouse, rat, rabbit, dog, cat, cattle, swine, sheep, horse or primate), 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".

The term "patient" means an individual or subject for treatment, in particular a diseased individual or subject.

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.

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.

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.

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.

In some embodiments, the nucleic acid is present in a composition as described herein.

In some embodiments, the nucleic acid is administered in a pharmaceutically effective amount.

In some embodiments, the subject is a mammal. In some embodiments, the mammal is a human.

In some embodiments of the disclosure, the aim is to induce an immune response by providing a vaccine.

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.

In some embodiments of the disclosure, the aim is to treat cancer by vaccination.
In some embodiments of the disclosure, the aim is to provide protection against an infectious disease by vaccination.

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.

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.

In some embodiments of the disclosure, the aim is to modulate/reprogram immune cells in the blood.

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.

In some embodiments of the disclosure, the aim is to provide one or more cytokines or cytokine fusions which have anti-tumor activity to a subject, in particular a subject in need thereof.

Citation of documents and studies referenced herein is not intended as an admission that any of the foregoing is pertinent prior art. All statements as to the contents of

these documents are based on the information available to the applicants and do not constitute any admission as to the correctness of the contents of these documents. The 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: RNA identity

Materials and Methods:

cDNA synthesis

Complementary DNA synthesis was performed according to the manufacturer's protocol using the SuperScript IV First-Strand synthesis kit (Invitrogen). The RNA mixture samples were diluted to 5 ng/µL. For one reaction, 5 ng of RNA mixture, 1 µL of 10 µM cDNA primer, and 1 µL of 10 mM dNTPs were combined, and the volume was adjusted to 13.5 µL with H_2O. The cDNA primer anneals in the 3'UTR, partially covering the poly(A)-tail. Generally, a master mix containing RNA sample, primers, dNTP, and water was made and divided into four tubes to measure the sample in triplicate and one negative control. The RNA was denatured at 80°C for 5 min, snap-cooled on ice for more than 1 min, and 4 µL of 5× SuperScript IV buffer, 1 µL of 40 U/µL RNase inhibitor, 1 µL of 0.1 M DTT, and 0.5 µL of 200 U/µL SuperScript IV reverse transcriptase were added. For the negative control, 0.5 µL of water were added instead of the reverse transcriptase. Samples were incubated in a PCR cycler with the following program: 55°C for 10 min, 80°C for 10 min, hold at 4°C. Afterwards, 0.5 µL RNase H (2 U/µL) was added and the sample incubated for 20 min at 37°C. The cDNA samples were either stored at -20°C in DNA low binding tubes or directly processed for ddPCR.

Droplet digital PCR

Droplet digital PCR was conducted on a QX200/C1000 system (Bio-Rad) according to the manufacturer's instructions. The cDNA was always freshly diluted to a concentration of ~1000 CN/µL. For one reaction, 5.5 µL of cDNA, 11 µL of 2x ddPCR SuperMix (Bio-Rad), 0.25 µM dual-labelled HEX-BHQ1-probe (Eurofins Genomics), 0.9 µM of RNA-specific forward primer (Eurofins Genomics), and 0.9 µM of common reverse primer were combined in a final volume of 22 µL. Each sample was measured in triplicate, together with one negative control per sample. After oil droplet generation, the samples were incubated in the C1000 thermocycler (Bio-Rad) and the following thermal program was executed:
Activation step 1): 600s at 95°C,
Denaturation step 2): 30s at 94°C,
Annealing and extension step 3) 60s at 63°C,
Enzyme inactivation step 4): 600s at 98°C.
Steps 2 and 3 were repeated 40 times. After PCR completion, the droplet fluorescence was read with the QX200 droplet reader (Bio-Rad).

Data Analysis

Data were analyzed using QuantaSoft version 1.7.4.0917 (Bio-Rad). QuantaSoft is generally able to automatically distinguish and define positive and negative populations. In the rare case that this was not possible, the threshold was set manually. The software automatically calculates the CN/µL. For RNA identity confirmation, the sample must show a minimum of 300 CN/µL and the negative control a maximum of 10 CN/µL.

Results:

To verify that the method correctly assigns the RNA identity, each oligonucleotide set was tested against each single RNA composing the sample RNA mixture. When the 3´-end region was analyzed with the respective oligonucleotide set, the ddPCR generated positive droplets only if the RNA-specific forward primer was combined with the target RNA (e.g., when the RNA A was amplified with the RNA A forward primer). In all other forward primer/RNA combinations, only negative droplets were observed.

	CN/µL			
Primer pair / **RNA**	**A**	**B**	**C**	**D**
A	**7,900**	0	0	0
B	1	**4,220**	0	0
C	0	0	**5,313**	5
D	0	0	0	**3,610**

Table 1

The data in Table 1 report the CN/µL (CN = copy number) measured for each RNA of a four-RNA mixtures tested with forward primer specific for each RNA (i.e., primer pair A is specific for RNA A). Each sample was measured in triplicate. The numbers in **bold type** in the table indicate high CN/µL, those in normal type indicate low CN/µL.

Example 2: RNA ratio

Materials and Methods:

cDNA synthesis

Complementary DNA synthesis was performed according to the manufacturer's protocol using the SuperScript IV First-Strand synthesis kit (Invitrogen). The RNA mixture samples were diluted to 5 ng/µL. For one reaction, 5 ng of RNA mixture, 1 µL of 10 µM cDNA primer, and 1 µL of 10 mM dNTPs were combined, and the volume was adjusted to 13.5 µL with H_2O. The cDNA primer anneals in the 3'UTR, partially covering the poly(A)-tail. Generally, a master mix containing RNA sample, primers, dNTP, and water was made and divided into four tubes to measure the sample in triplicate and one negative control. The RNA was denatured at 80°C for 5 min, snap-cooled on ice for more than 1 min, and 4 µL of 5× SuperScript IV buffer, 1 µL of 40 U/µL RNase inhibitor, 1 µL of 0.1 M DTT, and 0.5 µL of 200 U/µL SuperScript IV reverse transcriptase were added. For the negative control, 0.5 µL of water were added instead of the reverse transcriptase. Samples were incubated in a PCR cycler with the following program: 55°C for 10 min, 80°C for 10 min, hold at

4°C. Afterwards, 0.5 µL RNase H (2 U/µL) was added and the sample incubated for 20 min at 37°C. The cDNA samples were either stored at -20°C in DNA low binding tubes or directly processed for ddPCR.

Droplet digital PCR

Droplet digital PCR was conducted on a QX200/C1000 system (Bio-Rad) according to the manufacturer's instructions. The cDNA was always freshly diluted to a concentration of ~1000 CN/µL. For one reaction, 5.5 µL of cDNA, 11 µL of 2x ddPCR SuperMix (Bio-Rad), 0.25 µM dual-labelled HEX-BHQ1-probe, 0.9 µM of RNA-specific forward primer, and 0.9 µM of common reverse primer were combined in a final volume of 22 µL. Each sample was measure in triplicate, together with one negative control per sample. After oil droplet generation, the samples were incubated in the C1000 thermocycler (Bio-Rad) and the following thermal program was executed:

Activation step: 1) 600s at 95°C,

Denaturation step 2) 30s at 94°C,

Annealing and extension step 3) 60s at 63°C,

Enzyme inactivation step 4) 600s at 98°C.

Step 2 and 3 were repeated 40 times. After PCR completion, the droplet fluorescence was read with the QX200 droplet reader (Bio-Rad).

Data Analysis

Data were analyzed using QuantaSoft version 1.7.4.0917 (Bio-Rad). QuantaSoft is generally able to automatically distinguish and define positive and negative populations. In the rare case that this was not possible, the threshold was set manually. The software automatically calculates the CN/µL. For RNA ratio calculation, the CN/µL were converted into mass/volume concentration (g/µL) according to the following formula:

$$\frac{g}{\mu L} = \left(\frac{CN}{\mu L}\right) * \frac{MW}{6.022 * 10^{23}}$$

where MW is the molecular weight of the RNAs. Afterwards, the ratios (in %) were calculated with the following formulas:

$$\%A = \frac{A\,[\frac{mass}{volume}]}{(A + B + C + D)\,[\frac{mass}{volume}]} * 100$$

To calculate the ratio for RNA B, C or D, the concentration of the respective RNA must be given in the numerator.

Results:

To test the accuracy of the method for RNA ratio measurement, several solutions containing RNAs mixed in different ratios were measured. To evaluate the accuracy of the method, the "%recovery" was calculated, which expresses how close the measured value is to the theoretical value according to the following equation:

$$\%recovery = \frac{mean\ of\ measured\ values}{theoretical\ value} * 100$$

The %recovery spanned between 89.3% and 119.4%. The ratio of the four RNAs is calculated as the amount of one RNA in solution with respect to the other three; that is, if the amount of one RNA is lower, the amount of the other three will increase in a complementary fashion. Thus, the same difference between the measured and the theoretical values results in different %recovery values. For completeness in estimating the accuracy of the method, "delta" was also evaluated, where delta is the mathematical distance of the measurement for both the RNAs in one mixture from the target value. The maximum delta over all the tested RNA mixes is = 3, which further confirms the accuracy of the method.

RNA	Theoretical RNA ratio (%)	Measured RNA ratio (%)					%recovery	Delta
		Replicate 1	Replicate 2	Replicate 3	Mean (%)	SD (%)		
A	26	27	27	27	27	0.3	103.5	1
B	27	27	27	27	27	0.2	100.4	0
C	24	23	23	23	23	0.3	95.8	-1
D	23	23	23	23	23	0.3	99.9	0
A	30	32	34	25	30	4.5	101.4	0
B	32	32	33	35	33	1.6	104.3	1
C	19	17	14	20	17	3.1	89.3	-2
D	18	19	19	20	19	0.4	107.0	1
A	21	22	22	22	22	0.1	106.4	1
B	22	22	22	22	22	0.2	99.6	0

C	29	27	28	28	28	0.2	95.2	-1
D	28	28	28	28	28	0.3	100.4	0
A	23	23	24	24	24	0.2	102.3	1
B	26	26	26	27	26	0.3	101.8	0
C	27	26	26	26	26	0.2	97.2	-1
D	24	24	24	24	24	0.3	99.0	0
A	35	38	38	37	37	0.3	106.7	2
B	27	26	27	27	27	0.2	98.5	0
C	24	22	22	22	22	0.2	92.4	-2
D	14	14	14	14	14	0.4	99.1	0
A	15	18	18	18	18	0.3	119.4	3
B	28	27	27	27	27	0.4	95.9	-1
C	24	24	22	23	23	0.7	96.3	-1
D	32	32	32	33	32	0.4	100.3	0

Table 2

The data in Table 2 report the mean and standard deviation (SD) of the triplicates as well as the %recovery (mean of measured values ÷ theoretical value × 100) and delta (the mathematical distance of the measurement from the theoretical value). Compare the “Theoretical RNA ratio” column with the “Mean” to verify that the measured ratio reflects the theoretical value with minimal deviation.

Example 3: RNA integrity

Materials and Methods:

cDNA synthesis

Complementary DNA synthesis was performed according to the manufacturer’s protocol using the SuperScript IV First-Strand synthesis kit (Invitrogen). The RNA mixture samples were diluted to 5 ng/µL. For one reaction, 5 ng of RNA mixture, 1 µL of 10 µM cDNA primer, and 1 µL of 10 mM dNTPs were combined, and the volume was adjusted to 13.5 µL with H_2O. The cDNA primer anneals in the 3’UTR, partially covering the poly(A)-tail. Generally, a master mix containing RNA sample, primers, dNTP, and water was made and divided into four tubes to measure the sample in triplicate and one negative control. The RNA was denatured at 80°C for 5 min, snap-cooled on ice for more than 1 min, and 4 µL of 5× SuperScript IV buffer, 1 µL of 40 U/µL RNase inhibitor, 1 µL of 0.1 M DTT, and 0.5 µL of 200 U/µL

SuperScript IV reverse transcriptase were added. For the negative control, 0.5 µL of water were added instead of the reverse transcriptase. Samples were incubated in a PCR cycler with the following program: 55°C for 10 min, 80°C for 10 min, hold at 4°C. Afterwards, 0.5 µL RNase H (2 U/µL) was added and the sample incubated for 20 min at 37°C. The cDNA samples were either stored at -20°C in DNA low binding tubes or directly processed for ddPCR.

Droplet digital PCR

Droplet digital PCR was conducted on a QX200/C1000 system (Bio-Rad) according to the manufacturer's instructions. The cDNA was always freshly diluted to a concentration of ~1000 CN/µL. For one reaction, 5.5 µL of cDNA, 11 µL of 2x ddPCR SuperMix (Bio-Rad), 0.25 µM dual-labelled HEX-BHQ1-probes, 0.9 µM of RNA-specific primers, and 0.9 µM of common reverse primers were combined in a final volume of 22 µL. The reaction mix includes two oligonucleotides sets, one targeting the 5'-end and the other the 3'-end of the RNA. The 5'-end oligonucleotide set is composed of one common forward primer, one RNA-specific reverse primer, and one common FAM-BHQ1 dual-labelled probe. The 3'-end oligonucleotide set is composed of one RNA-specific forward primer, one common reverse primer, and one common HEX-BHQ1 dual-labelled probe. Each sample was measure in triplicate, together with one negative control per sample. After oil droplet generation, the samples were incubated in the C1000 thermocycler (Bio-Rad) and the following thermal program was executed:

Activation step 1) 600s at 95°C,

Denaturation step 2) 30s at 94°C,

Annealing and extension step 3) 60s at 63°C,

Enzyme inactivation step 4) 600s at 98°C.

Step 2 and 3 were repeated 40 times. After PCR completion, the droplet fluorescence was read with the QX200 droplet reader (Bio-Rad).

Data Analysis

Data were analyzed using QuantaSoft version 1.7.4.0917 (Bio-Rad). QuantaSoft is generally able to automatically distinguish and define positive and negative populations. However, because of the complex mixture of RNA templates and primers/probes, the negative, single-positive, and double-positive populations are not

always clearly separated, and the software fails the automatic assignment. Thus, definition of the populations must be done manually visualizing the data on a 2D amplitude plot. Once the four populations are assigned, the software automatically calculates the HEX-positive and FAM-positive CN/µL for each sample. In addition, the linkage is calculated. Linkage was initially defined for the analysis of proximity of two sequences physically linked on the same piece of DNA (JF Regan et al. *PLOS ONE*, 2015, 10(3), e0118270). Since the 5´ end and 3´-end are physically linked in an intact RNA, this number can be applied to the study of RNA integrity too. The linkage describes the CN/droplet concentration of 5´- and 3´-end double-positive droplets, once normalized for the possibility of random co-partitioning of single- and double-positive templates. To determine the fraction of intact RNA, the linkage automatically calculated by the ddPCR software and converted in CN/µL is divided by the total number of 3´-end (HEX) CN/µL to obtain the %linkage. This number correlates with the RNA integrity.

$$\%linkage = \frac{linkage\ number}{total\ 3'\text{-}end\ CN} * 100$$

Results:

To test the accuracy of the method for RNA integrity measurements, each RNA was individually degraded to different integrity levels by incubation at 90°C for different time spans. The RNAs from the same degradation time point were then mixed and measured with ddPCR. The method successfully detected the decrease in RNA integrity over time for all the RNAs.

Figure 1 shows the integrity of each RNA composing the differently degraded RNA mixtures was measured with ddPCR. The longer the RNA was degraded at 90°C, the lower the integrity measured with ddPCR.

For the calculation of the integrity, the 3'-end CN/µL information is quantified. This is the same information needed for the assessment of RNA ratio and RNA identity. Thus, the method presented here can measure RNA identity, RNA ratio, and RNA integrity in parallel.

In addition, the expermental system is simplified by the fact that 2 components of each oligonucleotide set (i.e., one common primer and one common dual-labelled probe) are identical for all RNAs and only one primer of each set is RNA-specific. This reduces the amount of reagents needed, the complexity of the system and the pipetting time compared to assembling two sets each composed of RNA-specific oligonucleotides.

Example 4: RNA potency

Method description:
Total ribonucleic acid (RNA) was isolated from Chinese Hamster Ovary (CHO) cells which were previously transfected with different amounts of formulated RNA of interest. Total RNA was isolated using a commercial kit (RNeasy Micro kit, Qiagen) and reverse transcribed (SuperScript IV, Invitrogen) into complementary deoxyribonucleic acid (cDNA) with the help of an oligo(dT) primer (Oligo (dT)20, part of the SuperScript IV kit from Invitrogen).

The cDNA was diluted and analyzed by digital droplet polymerase chain reaction (ddPCR) using two sets of specific primers and fluorescently labelled probe against RNA of interest and housekeeping RNA. Probes were labelled with hexachloro-fluorescein (HEX) and 6-carboxyfluorescein (FAM) and carried a Black Hole Quencher 1 (BHQ1). PCR cycling included initial activation for 600 s at 95°C and 40 cycles of denaturation for 30 s at 94°C and annealing / extension for 60 s at 61°C. Subsequently, a final enzyme inactivation for 60 s at 98°C followed.

Figure 2 shows that copy numbers (CN) measured by digital droplet polymerase chain reaction (ddPCR) for a ribonucleic acid (RNA) of interest and a housekeeping gene from total RNA isolated from Chinese Hamster Ovary (CHO) cells which were previously transfected with four different amounts of formulated RNA of interest.

All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the present invention will be apparent to those skilled in the art without departing from the scope and spirit of the present invention. Although the present invention has

been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in chemistry, biochemistry, molecular biology, biotechnology or related fields are intended to be within the scope of the following claims.

CLAIMS

1. A method of determining a quality parameter of an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 2, the quality parameter being selected from the group consisting of:
   i) quantitative ratio of two or more RNA molecule species of the *n* RNA molecule species; and
   ii) identity of the *n* RNA molecule species in the RNA sample;
   the method comprising the steps of:
   a) reverse transcription of the *n* RNA molecule species in the RNA sample into cDNA molecules of *n* DNA molecule species; and
   b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set and a single second primer, wherein
   the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region of only one of the *n* DNA molecule species in the sample, and
   the single second primer is capable of annealing to a second target region of all of the *n* DNA molecule species in the sample.

2. A method according to claim 1, wherein the quality parameter is the quantitative ratio of two or more RNA molecule species of the *n* RNA molecule species.

3. A method according to claim 1, wherein the quality parameter is the identity of the *n* RNA molecule species in the RNA sample.

4. A method of determining integrity of an RNA sample containing RNA molecules of *n* RNA molecule species, wherein *n* is an integer of at least 1,
   the method comprising the steps of:
   a) reverse transcription of the *n* RNA molecules in the RNA sample into cDNA molecules of *n* DNA molecule species; and
   b) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer set, a single second primer, a third primer set and a single fourth primer, wherein

the first primer set comprises *n* primer species, wherein each primer species is capable of annealing to a first target region at the 3'-end region of a DNA molecule species in the sample,
the single second primer is capable of annealing to a second target region at the 3'-end region of all of the *n* DNA molecule species in the sample;
the third primer set comprises *n* primer species, wherein each primer species is capable of annealing to a third target region at the 5'-end region of a DNA molecule species in the sample; and
the single fourth primer is capable of annealing to a fourth target region at the 5'-end region of all of the *n* DNA molecule species in the sample.

5. A method of determining the potency of a formulated RNA sample comprising RNA molecules of interest, the method comprising the steps of:
a) providing an RNA sample which has been isolated from cells transfected with a formulated RNA sample;
b) reverse transcription of the RNA molecules in the RNA sample into cDNA molecules;
c) carrying out a polymerase chain reaction (PCR)-based assay on the resulting cDNA molecules, the PCR-based assay using a first primer, a second primer, a third primer and a fourth primer, wherein
the first primer and the second primer are capable of annealing to a first target region and a second target region of the cDNA molecules produced from the RNA of interest in the sample, and
the third primer and the fourth primer are capable of annealing to a first target region and a second target region of the cDNA molecules derived from an endogenous RNA in the sample; and
d) comparing the measured amount of the cDNA produced from the RNA molecules of interest with the measured amount of cDNA produced from endogenous RNA.

6. A method according to claim 5, wherein the endogenous RNA is RNA expressed from a housekeeping gene in the sample, preferably wherein the endogenous RNA is RNA expressed from a GAPDH gene in the sample.

7. A method according to claim 5 or 6, wherein the formulated RNA composition comprises two or more different RNA molecule species of interest and the PCR-based assay uses two or more different primer pairs which are each specific for one of the RNA molecule species of interest.

8. A method according to any one of claims 5 to 7, wherein steps a) to d) are repeated for further RNA samples which have been isolated from cells transfected with formulated RNA compositions comprising the RNA molecule(s) of interest to determine a potency level for each formulated RNA composition.

9. A method according to any one of claims 5 to 8, wherein carrying out the method establishes an expected potency level for a particular formulated RNA composition comprising a RNA of interest, the expected potency level being defined as a reference potency level.

10. A method according to claim 9, wherein the method is carried out on a further formulated RNA composition comprising a RNA of interest and comparing it with the reference potency level for the formulated RNA composition.

11. A method according to any one of claims 6 to 10, further comprising the following step z) prior to step a):
z) isolating/purifying RNA from cells which have been transfected with the formulated RNA sample, and
optionally further comprising the following step y) prior to step z):
y) transfecting the cell with the formulated RNA composition.

12. A method according to any preceding claim, wherein the PCR-based assay in step b) is digital PCR (dPCR), preferably wherein the PCR-based assay in step b) is droplet digital PCR (ddPCR).

13. A method according to any preceding claim, wherein the PCR-based assay uses a detectable label, preferably wherein the detectable label is a fluorescent probe.

14. A method according to any preceding claim, wherein at least one RNA of the RNA molecules present in the sample is complexed with at least one carrier compound, thereby forming at least one RNA-carrier complex.

15. A method according to claim 14, wherein the cells are transfected using RNA which is complexed with at least one carrier compound, thereby forming at least one RNA-carrier complex.

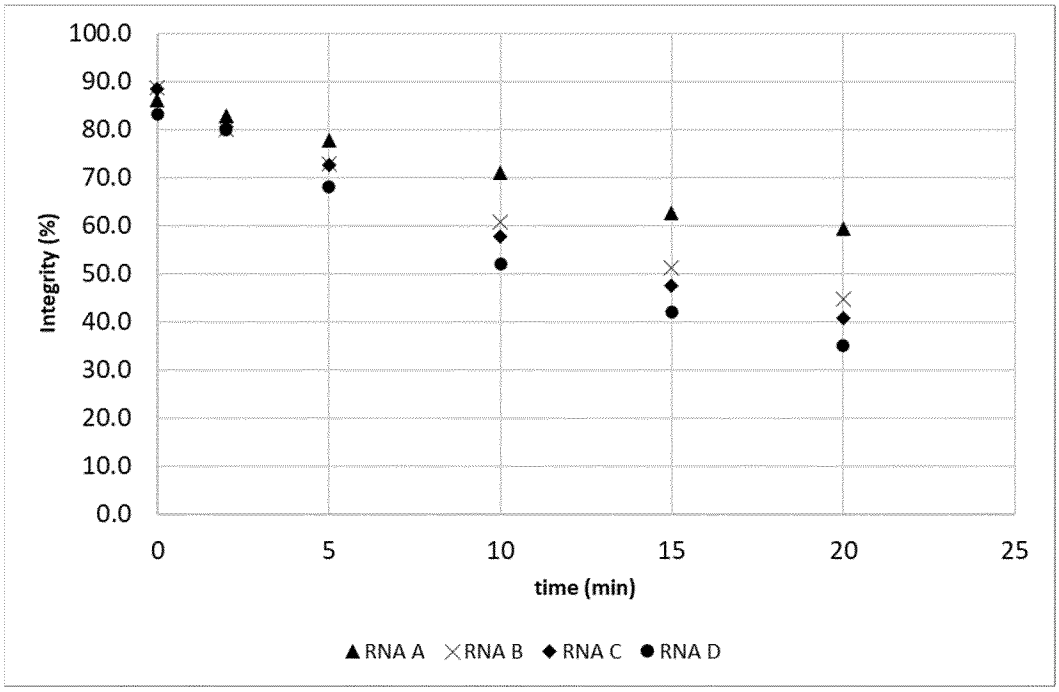


Figure 1

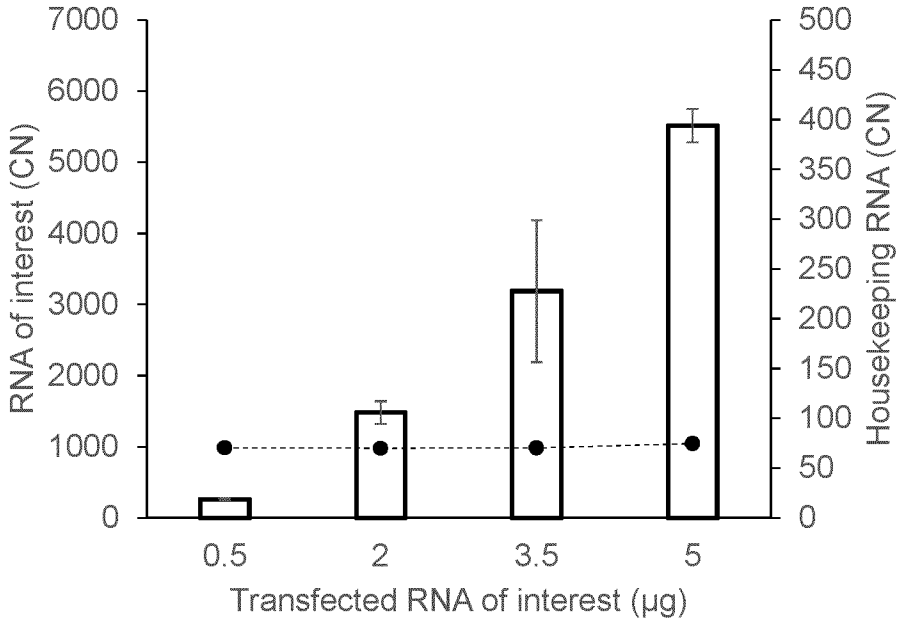


Figure 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/076303

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K31/7088 C12N15/10 C12Q1/686 G01N30/72
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N C40B A61K C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2018/211038 A1 (CUREVAC AG [DE]) 22 November 2018 (2018-11-22) cited in the application claims 1-34 page 2, paragraph 6 - page 4 page 10, paragraph bridging - page 11 page 19, paragraph 2 - page 25, paragraph 1 -----	1-15
A	WO 2006/119439 A2 (ALTHEA TECHNOLOGIES INC [US]; MONFORTE JOSEPH [US] ET AL.) 9 November 2006 (2006-11-09) paragraphs [0018], [0023] - [0028], [0055] - [0056], [0083], [0152] - [0158], [0175] - [0176], [0190] - [0193], [0199] figure 2 examples 1-4 -----	1-15
	-/--	

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 December 2024	19/12/2024

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Bruma, Anja

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2017/137095 A1 (CUREVAC AG [DE]) 17 August 2017 (2017-08-17) claims 29-85	1-15
A	JOSE V DIE ET AL: "Characterization of the 3:5 ratio for reliable determination of RNA quality", ANALYTICAL BIOCHEMISTRY, ACADEMIC PRESS, AMSTERDAM, NL, vol. 419, no. 2, 8 August 2011 (2011-08-08), pages 336-338, XP028313894, ISSN: 0003-2697, DOI: 10.1016/J.AB.2011.08.012 [retrieved on 2011-08-16] the whole document	1-15
A	PADHI BHAJA K. ET AL: "A PCR-based quantitative assay for the evaluation of mRNA integrity in rat samples", BIOMOLECULAR DETECTION AND QUANTIFICATION , vol. 15 1 May 2018 (2018-05-01), pages 18-23, XP093109882, ISSN: 2214-7535, DOI: 10.1016/j.bdq.2018.02.001 Retrieved from the Internet: URL:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6006387/pdf/main.pdf the whole document figure 1	1-15
Y	MAHAJAN ROHIT ET AL: "A TaqManReverse Transcription Polymerase Chain Reaction (RT-PCR) In Vitro Potency Assay for Plasmid-based Vaccine Products", MOLECULAR BIOTECHNOLOGY, SPRINGER US, NEW YORK, vol. 40, no. 1, 26 March 2008 (2008-03-26) , pages 47-57, XP037124922, ISSN: 1073-6085, DOI: 10.1007/S12033-008-9058-3 [retrieved on 2008-03-26] the whole document	5-15

-/--

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CLARNER PETE ET AL: "Development of a one-step RT-ddPCR method to determine the expression and potency of AAV vectors", MOLECULAR THERAPY- METHODS & CLINICAL DEVELOPMENT , vol. 23 1 December 2021 (2021-12-01), pages 68-77, XP093130199, GB ISSN: 2329-0501, DOI: 10.1016/j.omtm.2021.05.003 Retrieved from the Internet: URL:https://www.cell.com/action/showPdf?pii=S2329-0501(21)00087-5 the whole document -----	5-15
Y	WO 2023/030635 A1 (BIONTECH SE [DE]) 9 March 2023 (2023-03-09) claims 1-39 -----	5-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2024/076303

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018211038 A1	22-11-2018	EP 3625363 A1	25-03-2020
		US 2023064106 A1	02-03-2023
		WO 2018211038 A1	22-11-2018
WO 2006119439 A2	09-11-2006	AU 2006243757 A1	09-11-2006
		CA 2607454 A1	09-11-2006
		EP 1883710 A2	06-02-2008
		JP 5280841 B2	04-09-2013
		JP 2008541699 A	27-11-2008
		US 2006281108 A1	14-12-2006
		US 2015275270 A1	01-10-2015
		WO 2006119439 A2	09-11-2006
WO 2017137095 A1	17-08-2017	EP 3414340 A1	19-12-2018
		EP 4289965 A2	13-12-2023
		US 2021180106 A1	17-06-2021
		WO 2017137095 A1	17-08-2017
WO 2023030635 A1	09-03-2023	AU 2022336160 A1	25-05-2023
		AU 2023237054 A1	12-10-2023
		CA 3198815 A1	09-03-2023
		CN 117157400 A	01-12-2023
		EP 4208552 A1	12-07-2023
		JP 7446527 B2	08-03-2024
		JP 2023551735 A	12-12-2023
		WO 2023030635 A1	09-03-2023
		WO 2023031367 A1	09-03-2023