

RNA ENCODING PEPTIDOGLYCAN HYDROLASE AND USE THEREOF FOR TREATING BACTERIAL INFECTION

Technical Field

The invention provides agents and methods for treating bacterial infections using RNA. The RNA encoding peptidoglycan hydrolases, e.g., endolysins, is formulated and administered in a way that peptidoglycan hydrolase proteins, e.g., endolysin proteins, can be produced and secreted by cells of a subject to combat bacterial infections.

Background

Despite the widespread use of small-molecule antibiotics, humanity still suffers from more than 2 billion bacterial infections every year. While antibiotics are suitable for a range of these infections, there are several settings in which antibiotics are not able to provide adequate cure rates. Resistance rates are rising fast and for antibiotic-resistant pathogens novel therapies are urgently needed. In addition, antibiotics are often ineffective against bacterial biofilms – yet biofilms are the preferred habitation state of many bacterial pathogens. Finally, beneficial bacteria populate our microbiomes and, in many settings antibiotics do more harm than good (e.g. in the gut microbiome where antibiotic therapy can select for *C. difficile*). From a patient perspective, novel therapies are most urgently required when the patient suffers from a chronic infection in which antibiotics may lead to short-term resolution of symptoms but not to a lasting cure.

For all these reasons, novel anti-bacterials are urgently required.

Bacteriophage-encoded peptidoglycan hydrolases—also called endolysins or “enzybiotics”—are a promising alternative to antibiotics (Fischetti, 2018; Schmelcher et al., 2012). Produced towards the end of the lytic cycle in phage-infected bacteria, these enzymes cleave peptidoglycan (PG) in the bacterial cell wall, thus lysing the cells and releasing the progeny phages. They have been initially described for Gram-positive bacteria, which lack the Gram-negative outer membrane that might limit accessibility to the cell wall peptidoglycan. Endolysins have several advantages over antibiotics; not least, their very narrow host spectrum, which is usually limited to a single genus or even a single species (Fischetti, 2010),

and their low propensity to generate resistance in their hosts (Schuch et al., 2014). Bacteriophages that invade Gram-positive bacteria encode a variety of highly diverse endolysins. In general, they have a modular structure consisting of one or more enzymatically active domains (EADs) connected by a flexible interdomain linker to at least one cell wall-binding domain (CBD), typically located at the C-terminus of the protein. Both domains can contribute to the specificity for a given genus or species of bacteria (Oliveira et al., 2013).

In addition to phage-endolysins, also other bactericidal proteins are known, including autolysins and bacteriocins, whose activities can be combined with those of endolysins (Linden and Nelson, 2018). Autolysins are structurally very similar to endolysins and are used by bacteria to remodel their cell wall by partially lysing it, for growing in size or during cell division. Bacteriocins are typically produced by bacteria to kill other bacteria. They can have architectures homologous to endolysins, such as the bacteriocin lysostaphin, a metallopeptidase produced by *Staphylococcus simulans* to specifically kill its relative *Staphylococcus aureus*. Or they can be antimicrobial peptides not related to endolysins, which for example perforate the inner membrane of Gram-negative bacteria, such as nisin.

Thus far resistance to endolysins has not been reported in bacteria and these enzymes demonstrated activity against bacteria in suspension as well as in biofilms - bacterial communities notorious for conferring antibiotic resistance and causing a range of hospital-associated and recurring infections. Additionally, engineered lysin has been shown to have 10-fold higher bactericidal activity as compared to natural lysins (Landlinger C. et al. (2021) Pathogens 10(1):54).

There is a need of providing agents and treatments for treating bacterial infections.

Summary

Here we describe an approach for treating bacterial infections employing RNA, e.g., nucleoside modified RNA (modRNA) or self-amplifying RNA (saRNA), to express endolysins in cells of a patient. The RNA encoding endolysins is formulated and administered in a way that endolysin proteins can be produced and secreted by patient cells to combat bacterial infections.

In one aspect, the invention provides RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase.

In some embodiments, the peptidoglycan hydrolase breaks down peptidoglycan in bacterial cell wall.

In some embodiments, the peptidoglycan hydrolase is derived from a bacterium.

In some embodiments, the peptidoglycan hydrolase is derived from a bacteriophage.

In some embodiments, the peptidoglycan hydrolase is or is derived from an endolysin.

In some embodiments, the endolysin comprises one or more enzymatically active domains (EADs) connected by a flexible interdomain linker to one or more cell wall-binding domains (CBD).

In some embodiments, the peptidoglycan hydrolase has a modified glycosylation pattern to improve pharmacokinetics in blood.

In some embodiments, the peptidoglycan hydrolase is modified so as to reduce glycosylation.

In some embodiments, the peptidoglycan hydrolase is modified by removing glycosylation sites.

In some embodiments, the peptidoglycan hydrolase is modified so as to reduce immunogenicity.

In some embodiments, the peptidoglycan hydrolase is modified by removing T cell epitopes.

In some embodiments, the peptidoglycan hydrolase is or is derived from endolysin Lys26A (from *P. aeruginosa* phage JD010).

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 2, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of

SEQ ID NO: 2, or a fragment of the nucleotide sequence of SEQ ID NO: 2 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 3, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, or a fragment of the nucleotide sequence of SEQ ID NO: 3 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3.

In some embodiments, the peptidoglycan hydrolase is secreted peptidoglycan hydrolase.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises a signal peptide.

In some embodiments, the peptidoglycan hydrolase is fused to a signal peptide.

In some embodiments, the signal peptide is cleaved off during secretion or export.

In some embodiments, the signal peptide comprises the amino acid sequence of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10, or a fragment of the amino acid sequence of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 5, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, or a fragment of the nucleotide sequence of SEQ ID NO: 5 or the nucleotide

sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises an extended pharmacokinetic (PK) polypeptide.

In some embodiments, the peptidoglycan hydrolase is fused to a pharmacokinetic modifying group.

In some embodiments, the pharmacokinetic modifying group comprises a moiety which is heterologous to the peptidoglycan hydrolase.

In some embodiments, the pharmacokinetic modifying group comprises a moiety selected from the group consisting of albumin, an immunoglobulin fragment, transferrin, Fn3, a functional variant thereof, or a functional fragment of the albumin, immunoglobulin fragment, transferrin, Fn3 or the functional variant thereof.

In some embodiments, the pharmacokinetic modifying group comprises human albumin, a functional variant thereof, or a functional fragment of the human albumin or the functional variant thereof.

In some embodiments, the pharmacokinetic modifying group is fused to the N-terminus of the peptidoglycan hydrolase.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises from N-terminus to C-terminus: N-pharmacokinetic modifying group-GS-linker-peptidoglycan hydrolase-C.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence, wherein the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

In some embodiments, the RNA comprises a modified nucleoside in place of uridine.

In some embodiments, the RNA comprises a modified nucleoside in place of each uridine.

In some embodiments, the modified nucleoside is selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1 ψ), and 5-methyl-uridine (m5U).

In some embodiments, the modified nucleoside is N1-methyl-pseudouridine (m1 ψ).

In some embodiments, the RNA is suitable for expression of the peptidoglycan hydrolase in a eukaryotic cell, e.g., a mammalian cell, such as a human cell. In some embodiments, the RNA comprises sequence elements allowing expression of the peptidoglycan hydrolase in a eukaryotic cell, e.g., a mammalian cell, such as a human cell. In some embodiments, such sequence elements comprise cap, 5'UTR, 3'UTR, and polyA.

In some embodiments, the RNA comprises a cap0 5' cap.

In some embodiments, the RNA comprises the 5' cap $m_2^{7,2'-O}G(5')ppSp(5')G$ (in particular its D1 diastereomer).

In some embodiments, the RNA comprises a cap1 5' cap.

In some embodiments, the RNA comprises the 5' cap $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$.

In some embodiments, the RNA comprises a 5' UTR.

In some embodiments, the 5' UTR comprises the nucleotide sequence of SEQ ID NO: 11, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11.

In some embodiments, the RNA comprises a 3' UTR.

In some embodiments, the 3' UTR comprises the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12.

In some embodiments, the RNA comprises a poly-A sequence.

In some embodiments, the poly-A sequence comprises at least 100 nucleotides.

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

In some embodiments, the RNA is mRNA.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 6, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence of SEQ ID NO: 6 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 7, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7, or a fragment of the nucleotide sequence of SEQ ID NO: 7 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7.

In some embodiments, the RNA is self-amplifying RNA.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 8, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8, or a fragment of the nucleotide sequence of SEQ ID NO: 8 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9, or a fragment of the nucleotide sequence of SEQ ID NO: 9 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9.

In some embodiments, the RNA is formulated as a liquid, formulated as a solid, or a combination thereof.

In some embodiments, the RNA is formulated for injection.

In some embodiments, the RNA is formulated for intravenous administration.

In some embodiments, the RNA is formulated as lipid particles.

In some embodiments, the RNA lipid particles are lipid nanoparticles (LNP).

In some embodiments, the RNA is present in a composition or medical preparation.

In some embodiments, the composition is a pharmaceutical composition.

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

In some embodiments, the medical preparation is a kit.

In some embodiments, the medical preparation further comprises instructions for use of the RNA for treating or preventing bacterial infection.

In some embodiments, the RNA is for administration to a human.

In some embodiments, the RNA is for pharmaceutical use.

In some embodiments, the pharmaceutical use comprises a therapeutic or prophylactic treatment of a disease or disorder.

In some embodiments, the therapeutic or prophylactic treatment of a disease or disorder comprises treating or preventing bacterial infection.

In one aspect, the invention provides a composition or medical preparation comprising RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase.

In some embodiments, the peptidoglycan hydrolase breaks down peptidoglycan in bacterial cell wall.

In some embodiments, the peptidoglycan hydrolase is derived from a bacterium.

In some embodiments, the peptidoglycan hydrolase is derived from a bacteriophage.

In some embodiments, the peptidoglycan hydrolase is or is derived from an endolysin.

In some embodiments, the endolysin comprises one or more enzymatically active domains (EADs) connected by a flexible interdomain linker to one or more cell wall-binding domains (CBD).

In some embodiments, the peptidoglycan hydrolase has a modified glycosylation pattern to improve pharmacokinetics in blood.

In some embodiments, the peptidoglycan hydrolase is modified so as to reduce glycosylation.

In some embodiments, the peptidoglycan hydrolase is modified by removing glycosylation sites.

In some embodiments, the peptidoglycan hydrolase is modified so as to reduce immunogenicity.

In some embodiments, the peptidoglycan hydrolase is modified by removing T cell epitopes.

In some embodiments, the peptidoglycan hydrolase is or is derived from endolysin Lys26A (from *P. aeruginosa* phage JD010).

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 2, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2, or a fragment of the nucleotide sequence of SEQ ID NO: 2 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 3, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, or a fragment of the nucleotide sequence of SEQ ID NO: 3 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3.

In some embodiments, the peptidoglycan hydrolase is secreted peptidoglycan hydrolase.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises a signal peptide.

In some embodiments, the peptidoglycan hydrolase is fused to a signal peptide.

In some embodiments, the signal peptide is cleaved off during secretion or export.

In some embodiments, the signal peptide comprises the amino acid sequence of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10, or a fragment of the amino acid sequence of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least

99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 5, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, or a fragment of the nucleotide sequence of SEQ ID NO: 5 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises an extended pharmacokinetic (PK) polypeptide.

In some embodiments, the peptidoglycan hydrolase is fused to a pharmacokinetic modifying group.

In some embodiments, the pharmacokinetic modifying group comprises a moiety which is heterologous to the peptidoglycan hydrolase.

In some embodiments, the pharmacokinetic modifying group comprises a moiety selected from the group consisting of albumin, an immunoglobulin fragment, transferrin, Fn3, a functional variant thereof, or a functional fragment of the albumin, immunoglobulin fragment, transferrin, Fn3 or the functional variant thereof.

In some embodiments, the pharmacokinetic modifying group comprises human albumin, a functional variant thereof, or a functional fragment of the human albumin or the functional variant thereof.

In some embodiments, the pharmacokinetic modifying group is fused to the N-terminus of the peptidoglycan hydrolase.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises from N-terminus to C-terminus: N-pharmacokinetic modifying group-GS-linker-peptidoglycan hydrolase-C.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is

increased compared to wild type coding sequence, wherein the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

In some embodiments, the RNA comprises a modified nucleoside in place of uridine.

In some embodiments, the RNA comprises a modified nucleoside in place of each uridine.

In some embodiments, the modified nucleoside is selected from pseudouridine (ψ), N1-methyl-pseudouridine ($m1\psi$), and 5-methyl-uridine ($m5U$).

In some embodiments, the modified nucleoside is N1-methyl-pseudouridine ($m1\psi$).

In some embodiments, the RNA is suitable for expression of the peptidoglycan hydrolase in a eukaryotic cell, e.g., a mammalian cell, such as a human cell. In some embodiments, the RNA comprises sequence elements allowing expression of the peptidoglycan hydrolase in a eukaryotic cell, e.g., a mammalian cell, such as a human cell. In some embodiments, such sequence elements comprise cap, 5'UTR, 3'UTR, and polyA.

In some embodiments, the RNA comprises a cap0 5' cap.

In some embodiments, the RNA comprises the 5' cap $m_2^{7,2'-O}G(5')ppSp(5')G$ (in particular its D1 diastereomer).

In some embodiments, the RNA comprises a cap1 5' cap.

In some embodiments, the RNA comprises the 5' cap $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$.

In some embodiments, the RNA comprises a 5' UTR.

In some embodiments, the 5' UTR comprises the nucleotide sequence of SEQ ID NO: 11, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11.

In some embodiments, the RNA comprises a 3' UTR.

In some embodiments, the 3' UTR comprises the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12.

In some embodiments, the RNA comprises a poly-A sequence.

In some embodiments, the poly-A sequence comprises at least 100 nucleotides.

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

In some embodiments, the RNA is mRNA.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 6, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence of SEQ ID NO: 6 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 7, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7, or a fragment of the nucleotide sequence of SEQ ID NO: 7 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7.

In some embodiments, the RNA is self-amplifying RNA.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 8, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8, or a fragment of the nucleotide sequence of SEQ ID NO: 8 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9, or a fragment of the nucleotide sequence of SEQ ID NO: 9 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9.

In some embodiments, the RNA is formulated as a liquid, formulated as a solid, or a combination thereof.

In some embodiments, the RNA is formulated for injection.

In some embodiments, the RNA is formulated for intravenous administration.

In some embodiments, the RNA is formulated or is to be formulated as lipid particles.

In some embodiments, the RNA lipid particles are lipid nanoparticles (LNP).

In some embodiments, the composition is a pharmaceutical composition.

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

In some embodiments, the medical preparation is a kit.

In some embodiments, the kit further comprises instructions for use of the RNA for treating or preventing bacterial infection.

In some embodiments, the composition or medical preparation is for administration to a human.

In some embodiments, the composition or medical preparation is for pharmaceutical use.

In some embodiments, the pharmaceutical use comprises a therapeutic or prophylactic treatment of a disease or disorder.

In some embodiments, the therapeutic or prophylactic treatment of a disease or disorder comprises treating or preventing bacterial infection.

In one aspect, the invention provides a cell comprising the RNA described herein. In one aspect, the invention provides a cell transfected with the RNA described herein.

In some embodiments, the cell is a human cell

In one aspect, the invention provides a population of cells, wherein at least one cell is a cell described herein. In one aspect, the invention provides a population of cells, wherein the population of cells comprises a plurality of a cell described herein.

In some embodiments, a cell or a population of cells described herein is present in a subject.

In one aspect, the invention provides a method of producing a peptidoglycan hydrolase *in vivo* comprising transfecting a cell or a population of cells with an RNA described herein.

In some embodiments, a cell or a population of cells described herein is present in a subject.

In some embodiments, the subject has been administered the RNA or the composition described herein

In one aspect, the invention provides a method of treating bacterial infection in a subject comprising administering RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase to the subject.

In some embodiments, the peptidoglycan hydrolase breaks down peptidoglycan in bacterial cell wall.

In some embodiments, the peptidoglycan hydrolase is derived from a bacterium.

In some embodiments, the peptidoglycan hydrolase is derived from a bacteriophage.

In some embodiments, the peptidoglycan hydrolase is or is derived from an endolysin.

In some embodiments, the endolysin comprises one or more enzymatically active domains (EADs) connected by a flexible interdomain linker to one or more cell wall-binding domains (CBD).

In some embodiments, the peptidoglycan hydrolase has a modified glycosylation pattern to improve pharmacokinetics in blood.

In some embodiments, the peptidoglycan hydrolase is modified so as to reduce glycosylation.

In some embodiments, the peptidoglycan hydrolase is modified by removing glycosylation sites.

In some embodiments, the peptidoglycan hydrolase is modified so as to reduce immunogenicity.

In some embodiments, the peptidoglycan hydrolase is modified by removing T cell epitopes.

In some embodiments, the peptidoglycan hydrolase is or is derived from endolysin Lys26A (from *P. aeruginosa* phage JD010).

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 2, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2, or a fragment of the nucleotide sequence of SEQ ID NO: 2 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 3, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, or a fragment of the nucleotide sequence of SEQ ID NO: 3 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3.

In some embodiments, the peptidoglycan hydrolase is secreted peptidoglycan hydrolase.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises a signal peptide.

In some embodiments, the peptidoglycan hydrolase is fused to a signal peptide.

In some embodiments, the signal peptide is cleaved off during secretion or export.

In some embodiments, the signal peptide comprises the amino acid sequence of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10, or a fragment of the amino acid sequence of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 5, a nucleotide sequence having

at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, or a fragment of the nucleotide sequence of SEQ ID NO: 5 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises an extended pharmacokinetic (PK) polypeptide.

In some embodiments, the peptidoglycan hydrolase is fused to a pharmacokinetic modifying group.

In some embodiments, the pharmacokinetic modifying group comprises a moiety which is heterologous to the peptidoglycan hydrolase.

In some embodiments, the pharmacokinetic modifying group comprises a moiety selected from the group consisting of albumin, an immunoglobulin fragment, transferrin, Fn3, a functional variant thereof, or a functional fragment of the albumin, immunoglobulin fragment, transferrin, Fn3 or the functional variant thereof.

In some embodiments, the pharmacokinetic modifying group comprises human albumin, a functional variant thereof, or a functional fragment of the human albumin or the functional variant thereof.

In some embodiments, the pharmacokinetic modifying group is fused to the N-terminus of the peptidoglycan hydrolase.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase comprises from N-terminus to C-terminus: N-pharmacokinetic modifying group-GS-linker-peptidoglycan hydrolase-C.

In some embodiments, the amino acid sequence comprising a peptidoglycan hydrolase is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence, wherein the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

In some embodiments, the RNA comprises a modified nucleoside in place of uridine.

In some embodiments, the RNA comprises a modified nucleoside in place of each uridine.

In some embodiments, the modified nucleoside is selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1 ψ), and 5-methyl-uridine (m5U).

In some embodiments, the modified nucleoside is N1-methyl-pseudouridine (m1 ψ).

In some embodiments, the RNA comprises a cap0 5' cap.

In some embodiments, the RNA comprises the 5' cap m₂^{7,2'-O}G(5')ppSp(5')G (in particular its D1 diastereomer).

In some embodiments, the RNA comprises a cap1 5' cap.

In some embodiments, the RNA comprises the 5' cap m₂^{7,3'-O}Gppp(m₁^{2'-O})ApG.

In some embodiments, the RNA comprises a 5' UTR.

In some embodiments, the 5' UTR comprises the nucleotide sequence of SEQ ID NO: 11, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11.

In some embodiments, the RNA comprises a 3' UTR.

In some embodiments, the 3' UTR comprises the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12.

In some embodiments, the RNA comprises a poly-A sequence.

In some embodiments, the poly-A sequence comprises at least 100 nucleotides.

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

In some embodiments, the RNA is mRNA.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 6, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence of SEQ ID NO: 6 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 7, a nucleotide sequence having

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

In some embodiments, the RNA is self-amplifying RNA.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 8, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8, or a fragment of the nucleotide sequence of SEQ ID NO: 8 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8.

In some embodiments, the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9, or a fragment of the nucleotide sequence of SEQ ID NO: 9 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9.

In some embodiments, the RNA is formulated as a liquid, formulated as a solid, or a combination thereof.

In some embodiments, the RNA is administered by injection.

In some embodiments, the RNA is administered by intravenous administration.

In some embodiments, the RNA is formulated as lipid particles.

In some embodiments, the RNA lipid particles are lipid nanoparticles (LNP).

In some embodiments, the RNA is formulated as a pharmaceutical composition.

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

In some embodiments, the subject is a human.

In some embodiments, the infection is a chronic infection.

In some embodiments, the bacteria causing infection are multi-resistant bacteria.

In some embodiments, the bacteria causing infection are Gram-negative bacteria, Gram-positive bacteria, or both.

In some embodiments, the bacteria are selected from *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacteriaceae*, *Porphyromonas gingivalis*, *Helicobacter pylori*, *Chlamydia trachomatis*, *Enterobacteriaceae*, *E. coli*, *Klebsiella pneumoniae*, *Salmonella*, *Shigella*, *Borellia*, *Campylobacter jejuni*, *Neisseria gonorrhoeae*, *Chlamydia trachomatis*, *Vibrio cholerae*, and *Fusobacterium nucleatum*.

In some embodiments, the bacteria are resistant to carbapenem.

In some embodiments, the bacteria are selected from *Staphylococci*, *Streptococci*, *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Clostridioides difficile*, *Cutibacterium acnes*, *Mycobacterium tuberculosis*, *Gardnerella*, *Enterococcus faecalis*, *Enterococcus faecium*, *Lactobacillus iners*, *Bacillus subtilis* and *Bacillus anthracis*.

In some embodiments, the bacteria are resistant to methicillin and vancomycin or non-susceptible to penicillin.

In some embodiments, the bacteria are methicillin-resistant *Staphylococcus aureus* (MRSA).

In some embodiments, the invention provides an agent described herein, e.g., the RNA, composition or medical preparation described herein, for use in a method described herein, e.g., for treating or preventing bacterial infection or a disease associated with or caused by bacterial infection.

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 of a recipient. In addition to wildtype or codon-optimized sequences encoding an amino acid sequence comprising the amino acid sequence of a peptide or polypeptide having peptidoglycan hydrolase activity, e.g., an endolysin, 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 some embodiments, the RNA contains all of these elements. In some embodiments, beta-S-ARCA(D1) ($m_2^{7,2'-O}GppSpG$) or $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ may be utilized as specific capping structure at the 5'-end of the RNA drug substances. As 5'-UTR sequence, the 5'-UTR sequence of the human 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 encoded amino acid sequence may comprise amino acid sequences other than the amino acid sequence of a peptide or polypeptide having peptidoglycan hydrolase activity. Such other amino acid sequences may support the function or activity of the peptide or polypeptide having peptidoglycan hydrolase activity. In some embodiments, such other amino acid sequences comprise an amino acid sequence enhancing protein stability or persistence in the body.

Furthermore, in the RNA, a secretory signal peptide may be fused to the peptidoglycan hydrolase-encoding regions. Sequences coding for short linker peptides predominantly consisting of the amino acids glycine (G) and serine (S), as commonly used for fusion proteins may be used as GS/Linkers.

The RNA may be formulated in lipid particles to generate serum-stable formulations for intravenous (i.v.) administration. The RNA may be present in lipid nanoparticles (LNP). RNA-nanoparticles may target liver which results in an efficient expression of the encoded protein. In one embodiment, the RNA described herein is N1-methylpseudouridine modified, dsRNA-purified RNA which is formulated as lipid nanoparticles for intravenous administration.

Brief description of the Figures

Figure 1. Bactericidal activity of 11 lysins expressed and secreted in *P.pastoris*. Supernatants of *P.pastoris* cultures expressing different lysin constructs were mixed with lysin buffer at the indicated dilutions and incubated with *S.aureus* cells. Bacterial growth was monitored by measuring the culture OD over the course of 24 hours. **A.)** Growth curves of *S.aureus* in the presence of supernatants from different lysin-expressing *P.pastoris* strains. L0502*: the supernatant of L0502 was diluted with a supernatant from wild type *P.pastoris* not expressing recombinant lysins (wt S/N). **B.)** Supernatants of different clones from lysin L0499 were incubated with *S.aureus* at the indicated dilutions and bacterial growth was monitored over time.

Figure 2. Detection of RNA encoded LysPA26 in HEK293T cells by anti-FLAG BV241 conjugate antibody staining and flow cytometry. HEK293T cells analyzed by flow cytometry were incubated with: no RNA (non-transfected control), 9_LysPA26 and 10_LysPA26. Viability indicates cell survival post transfection, transfection rate is indicative of the percentage of cells with protein expression and expression rate is mean fluorescence of the percentage of positive cells Heights of bars indicate the means of technical triplicates.

Figure 3. Western blot of denatured proteins detected in supernatants (culture media) and cell lysates of HEK cells transfected with 9_LysPA26 and 10_LysPA26 at concentration of 350 ng and 1µg. Positive control used was HVV Aptamer-FLAG (not shown due to overexposure). Expected product size ~17 kDa.

Figure 4. Schematic illustration of the general structure of RNA with 5'-cap, 5'- and 3'-untranslated regions, and poly(A)-tail. (A) Nucleoside-modified mRNA. (B) Self-amplifying mRNA additionally encodes for a viral-replicase and a subgenomic promotor. The individual elements are not drawn exactly true to scale compared to their respective sequence lengths.

mRNA = messenger ribonucleic acid; SGP = subgenomic promotor; sec = secretory signal peptide; UTR = untranslated region; PolyA = poly adenosine.

Figure 5. Expression of endolysins in HEK cells. HEK293T cells were transfected with 0.5 and 1 µg of modRNA and saRNA of construct 9_LysPA26 and construct 10_LysPA26. Extracellular protein or protein in the cell lysate was detected. Left panels show supernatant and cell lysates for modRNA. The constructs with secretion signals (10_LysPA26) are clearly detectable at both concentrations in the supernatant, while the constructs without secretion signals (9_LysPA26) cannot be detected. The same result is seen for saRNA.

Figure 6. Expression of DNA and RNA encoded lysostaphin in cells.

A,B: Concentration and time dependent expression of L0483 in HEK293T/17 cells from mRNA (pST4-T7-AGA-dEarl-hAg-Kozak-SS(IGHV1-2)-His-3C-hL0483-FI-A30LA70) and DNA (pcDNA3.1(+)-SS(IGHV1-2)-His-3C-hL0483). **(A)** 0.9×10^6 HEK293T/17 cells were transfected with 3, 5 or 7.5 µg mRNA. In parallel 2.5 µg of DNA were transfected. Supernatants were collected after 24 (lane 1), 48 (lane 2) or 72 (lane 3) hours of expression. **(B)** Quantification of chemiluminescent signals from panel A using Image Lab 6.

C,D: Expression of L0483 mRNA-LNP in HEK293T/17 and CHO-K1 cells (1.5 µg). Expression was compared to mRNA transfection performed with Ribojuice (construct PL084_pST4-T7-AGA-dEarl-hAg-Kozak-SS(IGkappa)-His-3C-hL0483(PM)-FLAG-FI-A30LA70). **(C)** Expression levels varied substantially between cell lines. Transfection with ribojuice resulted in lower expression. **(D)** Approximately 15 fold higher expression was seen from mRNA-LNPs compared to ribojuice transfection of the same mRNA.

Figure 7. Expression of DNA and RNA encoded lysins in HEK293T/17 cells. Expression of lysin variants from mRNA or DNA transfections. 5 µg of mRNA or 2.5 µg of DNA were transfected into HEK293T/17 cells. Supernatants were analyzed by Western Blot, demonstrating expression from DNA as well as from mRNA.

Figure 8. Bactericidal (A –D) and enzymatic (E) activity of lysins expressed from mRNA in HEK293T/17 cells. (A-D) Growth inhibition assay in duplicate. (A) Concentrated supernatant from HEK293T/17 transfection with 2.5 µg L0483 DNA in duplicate. (B) Growth control in duplicate. (C) Concentrated supernatant from HEK293T/17 transfection with L0483 mRNA-LNPs in duplicate. (D) Concentrated supernatant from HEK293T/17 transfection with Luciferase mRNA-LNPs in duplicate. (E) Enzymatic activity of lysins expressed from mRNA in HEK293T/17 cells.

Figure 9. PK parameters of recombinant and mRNA delivered lysostaphin - Measurement of lysin levels in mouse plasma. (A) 3 mice were treated with 535 µg of recombinant lysostaphin and plasma lysin levels were measured for 24 hours. (B) 3 mice were treated with 107 µg of recombinant lysostaphin and plasma lysin levels were measured for 24 hours. (C) 5 mice were treated with 30 µg of L0483 mRNA-LNP and plasma lysin levels were measured for 72 hours.

Figure 10. In vivo efficacy of lysostaphin in *S. aureus* mouse bacteremia. Shown is the reduction of bacterial burden in different organs after administration of two different doses of recombinant lysin or SOC (vancomycin).

Description of the Sequences

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

TABLE 1: DESCRIPTION OF THE SEQUENCES		
SEQ ID NO:	Description	SEQUENCE
Endolysin LysPA26		
1	LysPA26 (amino acid sequence)	MRTNNIDAIKEHEGLRLVAYLDSVGWVTIGYDGTGPDVWKGLTITKEAEKRLRKRLVEFEGVYNTYVKPLKQHQFDALVSLVYNIGPTNFKTSTLLKLNAGDYGAADQFLVWNKGR VDGKLVV/KGLANRRAKERKQFLGE
2	LysPA26 (nucleotide sequence, wt)	AUGCGAACGAAUAAUUCGACGCCAUUAAAGAGCACGAGGGGCUACGUCUUGUAGCCUACCUAGCUCUGUAGGUGUAUGGACCAUUGGUUAUUGGUGACACUGGCCCUAGUAGUUAAGGGCCUUAACUUAUACCAAGAGAGGCCGAGAAAGCGCCUUAAGGAAGCGUCUGGUGGAGUUUGAAGGUAUGUCAAUACCUACGUCAAGGUUCCGCUCAAACAGCAUCAAUUCGACGUCUGGUGGAGUUUGGUGUACAACAUUGGUCCUACUAAUUAAGACCUUCGACGUUGCUUAAGAAGCUUAAACGUCGGGAUUAUUAUUGGAGCGGCUAGUAGUCCUAGUUGGAAACAAAGGCCGAGUAGAUGGGAAGCUUUGUGAUUAAAGGGUUCUAAUAGGGGGCCAAAGAACGUAACAAUCCUUGGGAGAAUGA
3	LysPA26 (nucleotide sequence, opt10)	AUGAGAACCAACAAUUCGAUGCCAUCAAAGAACAUAGAGGACUGAGACUGGUGGCUUACCUUGGUAUUCUGUGGAGUGUGGACAAUUGGAUUGGACACAGGACCUAGGAGGACUGGUGGAAUUGAAGGAAUUGAAGGAAUUGUAAACACCUAUGUGAAAGUGCCUCUGAAAAGAGACUGAGAAAGAGACUGGUGGAAUUGAAGGAAUUGUAAACACCUAUGUGAAAGUGCCUCUGAAAAGAGACUUGGAGGAGUUUGCUUGGUGGAGCCUGGUGUACAACAUUGGACCCCAACAUUAAGACACACACUGCUGAAGAAGCUGAAUUGCUGGAGAUUACAUUGGAGGCUAGUAGUCCUUGGUGGAAACAAAGGCCGAGUGGAUGGCAACUGGUGGUGAUCAAAGGCCUGGCCAACAGGAGAGCCAAAGAAAGAAACAGUUCUUGGGAGAAUGA
Secreted Endolysin LysPA26		
4	Secreted LysPA26 (amino acid sequence)	MIRVMAPRTLILLSGALALTETWAGSRTNNIDAIKEHEGLRLVAYLDSVGWVTIGYDGTGPDVWKGLTITKEAEKRLRKRLVEFEGYVNTYVKPLKQHQFDALVSLVYNIGPTNFKTSTLLKLNAGDYGAADQFLVWNKGRVDGKLVV/KGLANRRAKERKQFLGE

26

7	nucleotide sequence	AGAAUAAACUAGUAUUUCUUGGUCCCCCAGACUCAGAGAGAAACCCGCCCAUGAGAGUGGCCCCCAGAA CCCUGAUCCUGCUGUCUGUGGCGCCCUUGGCCUCGACAGAGACAUUGGCGCGGAAGCAGAACCAACAUCGAUG CCAUCAAGAAACAUGAAGGACUGAGACUGGGGCUUACCUGGAUUCUGUGGGAGUGUGGACAAUUGGAUAUGG AGACACAGGACCUGACGUGGUGAAGGACUGACAUCACAAAGGAAGCCGAAAAGAGACUGAGAAAAGAGAC UGGUGGAUUUUGAAGGAUAUGUGAACACCUAUGUGAAAGUGCCUCUGAAGAACAGCACAGUUUGAUGCUCUGGU GAGCCUGGUGUACAACAUAUUGGACCCACAACUUAAGACAAGCACACUGCUGAAGAAAGCUGAAUGCUGGAGAUU ACAUUGGAGCUGCUGAUCAGUUCUGGUGUGGAAACAAAGGCAGAGUGGAUUGGCAACUCUGGUGGUAUCAAAGG CCUGGCCAACAGGAGAGCCAAAGAAAGAAACAGUUCUUGGGAGAAUGACUCGAGCUGGUAUCUGCAUGCAC GCAUUGCUAGCUGCCCCUUCUCCGUCUUGGGUACCCCGAGUCCUCCCGACCUCCGGUCCCAGGUUUGCUCCCAC CUCCACCUGCCCCACUCACCAUCUCUGCUAGUUCAGACACUCCCAAGCACGCGAGCAUUGCAGCUCAAAACGCUU AGCCUAGCCACACCCCCACCGGAAACAGCAGUGAUUAACCUUUAAGCAUAAACGAAAGUUUAACUAAGCUAUACU AACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACCCUGGAGCUAGCAAAAAAAAAAAAAAAAAAAAAAAAAA AAGCAUAUGACUAAA AAAAAA
pST1-OPAL-LysPA26-A30LA70		

8	nucleotide sequence	AUGGGGAUGUGACAUACACGACGCCAAAGAUUUUGUUCAGCUCCUGCCACCUCGCUACGCGAGAGAUU AACCAACCACGAUGGCCGCCAAAGUGCAUUGAUUAGGUGCAGAGCCCAUUAUCAAGUCUUUGCAGAAG GCAUUUCCGUCGUCGAGGUGGAGUCAUUGCAGGUCACACCAAAUGACCAUUGCAAUUGCCAGAGCAUUUUCGCA CCUGGCUACCAAAUUGAUCGAGCAGGAGACUGACAAGACACACUCAUUGGAUUAUCGGCAGUGCGCCUUCGAG GAGAUGAUUCUACGCACAAUACCAUCUGUAUGCCUUAUGCGCAGCGCAGAAGACCCCGAAAGGCUUCGUUUG CUACGCAAAGAAACUGGCAGCGCCUCCGGGAAGGUGUGGAUAGAGAGAUUCGCGAGGAAAUACCCGACCUGCA GACCGUCAUGGCUACGCCAGACGUGAAUCUCCUACCUUUUGCCUGCAUACAGAGCUACGUGUCGUAACGGCAGC CGAAGUGGCCGUAUACCAAGGACGUGUAUGCUGUAUGCAUGCAACAACUGCGUACCAUCAGGCGAUGAAAGGUG UCAGAACGGCGUAUUGGAUUGGGUUGACACCCCGUUUAUGUUUAGCGCGUAGCGCGUAUCCCAACC UACGCCACAACUGGGCCGACGAGCAGGUGUACAGGCCAGGAACAAGGACUGUGUGCAGCAUCCUUGACUGA GGGAAGACUGGCAACUGUCCAUUCUCCGAAGAAGCAUUGAAACCUUGCGACACAGUCAUGUUCUGGUAAG GAUCUACAUUGUACACUGAGAGCAGAAAGCUACUGAGGAGCGGCGACUACCCUCCGUUUCCACCUGAAAGGU AAACAUCUUUACCUUGAGGUGCGAUACCAUCGUAUCAUGUAGGAGGUAUGUAUAGAAAUACACUAUGUG CCCCGGCCUGUACGGUAAAAACGGUAGGGUACGCCGUGACGUAUCACGCGGAGGGAUUCUAGUGCAAGACCA CAGACACUGUCAAAGGAGAAAGAGUCUCAUUCUUGUAUGCAUACGUCCCUCAACCAUCUGUGAUCAAUUGA CUGGCAUACUAGCGACCGACGUCACACCGGAGGACGCACAGAAGUUGUAGUGGGAUUGAAUACAGAGGAUAGUU GUGAACGGAAGAACACAGCGAAACACUAACACGAUGAAGAAACCUUGGUGUCCGAGAGAGGUCACUUACU CAAGUGGCGAGGGAUACAAGGCAGACCUUGAUGAUGAAGAAACCUUGGUGUCCGAGAGAGGUCACUUACU UGCUGCUGUUGUGGGCAUUUAAAACGAGGAAGAUACACCAUGUACAAGAAACCAAGAGCCUCCGAAUCCAGUCAGAU GAAGGUGCCUUCAGAGUUUAAUCUGUUCGUAUCCCGAGCCUUAUGGUCUACAGGCCUUCGACGCGUGUCAGCCAGG CAGCAUUAAGAUUGCUUUUGCCAAGAAGACCAAGCGAGAGUUAUUAUACCUUUCGACGCGUGUCGACGCCCAU GAUGCUGAACAAGAGGAGAGGAGGUGUUGGAGGCCGAGCUGACUAGAGAAGCCUUAACCCUCCGUCGCCCAU CGCGCGCGGAGACGGGAGUCGUCGACGUCGACGUUGAAGAACUAGAGUAUCACGCAAGGUGCAGGGGUGGUGG AAACACCUAGCAGCGGUUGAAAGUACCGCACAGCCGACGAGCUACUACUAGGAAAUUACGUUUCUGUCCCC CGCAGACCGUGUCUAAAGUCUCCAAAGUUGGCCCGGACCCUAGCAGAGCAGGUGAAAUAAUAAACACAU ACGGGAGGGCCGCGUUAACAGGUUCGACGGAUUAUGACGGCAGGGUCCUACUACCAUGGAGUCCGUAACAGAAACU GUCCUGAGUUCAGGCUUUGAGCGAGAGCGCCACUUGGUGUACAACGAAAGGAGUUCGUAACAGGAAACU AUACCAUUAUUGCCGUUCACGGACCCUCCUGAACACCGACGAGGAGAAACUACGAGAAAGUCAGAGCUGAAAGAAC UAGCGCGAGUACGUGUUCGACGUAGAUAAAAAUUGCUGCGUCAAGAGAGAGGAAGCGUGCGGUGUUGGUGUUG
---	--------------------------------	---

UAGCGCCUAAUUUCUCUGCCACGACUGAAGCGGAAGGGACCGCGAAUUGGCCGCGUCUACCGGGCAGUGGCC GCCGAAGUAAACAGACUGUCACUGAGCAGCGUAGCCAUCCCGCUGCUGUCCACAGGAGUGUUCAGCGCGGUAG AGAUAGGCGUGCAGCAAUCCCUCAACCAUUAUUCACAGCAUUGGACGCCACGGACGCGUGACCAUCUACUG CAGAGACAAAGUUUGGAGAAGAAAUCCAGGAAGCCAUAGACAUAGGACGCGUGUGGAGUUGCUCUAAUGAU GACGUGGAGCUGACCACAGACUUGGUGAGAGUGCACCCGGACAGCAGCCUGGUGGUGCGUAAGGGCUACAGUAC CACUGACGGGUCGCGUACUACUUGAAGGUACGAAUUAACACGCGUGCUAUUAUGGCGAGAGAUAC UGACGUUUGGGCCAGCUGCAAGAGGCAACGACAGAUUGCCUUAJACGCGCUGGGCGAAACAUAUGGACAACA UCAGAUCCAAUUGUCGGUGAACGAUUCGGAUUAUCAACACUCCAGGACAGUCCCGUGGCGCUACUUCUUC CAAUGACAGCAGAACGGAUUGCCGCGCUUAGGUCACACCAAGUUAAGCAUGGUGUUGCUCAUCUUCUCC CUCCGAAUACCAUGAUGGGGUGCAGAAGGUAAGUGCGAGAAGGUUCUCCUGUUCGACCCGACGGUACC UUUCAGUGGUUAGUCCGCGAAGUAUGCCGCAUUCAGACGGACCAUCAGAUCCGUGCUUACGAGGUGUUGACU UGGACUGGACCACCGACUCUUCACUGCCAGCGAUACCAUGUCUACCCAGUUUGCAGUUCGUGGACAUUG ACUCGAUCUACGAGCCAAUGGCUCCCAUAGUAGCGGCGUGACGUAACCCUUAACCCGAGGCAUCGCGGACC UGGCGGCAGAUUGCAUCCUGAACCCGACACCAUGUGGACCGGAGAACCCGAGAAAGCGCAGCGCCCAAGGA GAGCUGCAUACCUUGCCUCCCGCGGGGAGCGACCGGUGCGCGGACUUGACGAGCAGGUCGAUGCGUUGGCCUC CUGCGUUUAGGAACAAGCUGCCUUUGACGUUCGGCGACUUGAGCGCGCGGGUGCAUUAUUUCUCCGACACUG GAUUACUUUCGGAGACUUCGACGACGUCCUGCGAUAGGCGCGGGUGCAUUAUUUCUCCGACACUG GCAGCGGACAUUJACAACAAAUCGUAAGGCGACAAUUCUCCAGUGCGCACAAUCUGGUGCGGUGCGAGGAG GAGAAAUGUACCCGCCAAAUUGGAUACUGAGAGGGAGAGCUGUUGCUGUAAAAUGCAGAUCCACCAUC GGAGGCUAAUAGAGUCGAUACCAUCUCGCAAGUGGAGAACAUAAAGCCAGGUGGAGAGGCUACAU CGGGGGCCAGAUUGUACACGGGAGCGGACGUAGGCCGCAUACCAACAUACGGGUUCGGUACCCCGCCGUGU ACUCCCUACCGUGAUCGAAAGAUUCUACGCCCGAUGUAGCAUUCGAGCGUGCAACAAUACCUAUCCAGAA AUUACCCACACAGUGGCGUCGUACCAAGAUAAACAGAUAAUACGACGCAUACUUGGACAUUGGUGCGGUGCAU AGUUGCUUGGACAGAGCAGAUUCUGCCCGCGAAGCUCGGUGCUACCCGAAACAUCAUGCGUACCCAGCCG ACUGUACGAGUGCCGUCACCCUUCAGAACACACUACAGAACGUGUAGCGGCGUCCACCAAGAGAAAC UGCAACGUCACGCAAAUGCGAGAACUACCCACCAUGGACUGGCGAGUGUUAACGUGGAGUGCUUAAGCGCUA UGCCUGCUCGGGAGAUUUGGGAAGAAUUGCUAAACCAACCUUCCGGAUAAACCAUGAGAAUACUACUACCUA UGUGACCAAAUUGAAAGGCCGAAAGCUGCGCUUGUUCGCUAAGACCCCAACUUGUUCGCGUGCAGGAGG UUCCCAUGGACAGAUUCACGGUCGACAUAGAAACGAGAUUGUCAAAGUCACUCCAGGGACGAAACACACAGAGGAAA	
---	--

31

9	nucleotide sequence	AUGGCGGAUGUGACAUACACGACGCCAAAGAUUUUUGUCCAGCUCGCCACCUCGCUACGCGAGAGAUU AACCAACCACGAUGGCCGCAAGUGCAUUAUUGAGGCGUGACAGCCCAUUAUCAAGUUCUUUGCAGAAG GCAUUCGCGUUCGAGUGGAGUCAUUGCAGGUACACCAAAUGACCAUUGCAAUUGCCAGAGCAUUAUUCGCA CCUGGCUACCAAAUUGAUUCGAGCAGGAGACUGACAAAGACACACUUAUUGGAUUAUCGGCAGUGCGCCUUCAG GAGAAUGAUGUCACGCACAAUACCAUCUGCUAUGCCCUAUGCGCAGCGCAGAAAGACCCGAAAGGCUUGAUUG CUACGCAAAGAAACUGGACGGCCUCCGGGAAGGUGCUGGAUAGAGAGAUUCGAGGAAAUACACCGACCUGCA GACCGUCAUGGCUACGCCAGCGCUAUCUCCUACCUUUUGCCUGCAUACAGACGUCACGUGUCGACGGCAGC CGAAGUGGCCGUAUACCGAGCGUGUAUGCUGUACAUGCAACCAACUACGCUUUAUGUACGCGCUAGCGCGCUAUCCAACC UCAGAACGGCGUAUUGGAUUGGUUUGACACCCCGUUUAUGUACGCGCUAGCGCGCUAUCGCGCUUAGACUGA UACGCCCAAAACUGGCCGACGAGCGGUGUUAACAGGCCAGGAACAUAAGGACUGUGGAGCAUCCUUGACUGA GGGAAGACUCGGCAACUUGUCCAUUCUCCGAAGAACAUAUGAAACCUUGCGACACAGUCAUGUUCGCUAG GAUCUACAUUGUACACUGAGAGCAGAAAGCUACUGAGGAGCUGGCACUUAACCCUCCGUUUCCACCUGAAAGGU AAACAAUCCUUUACCUUGAGGUGCGAUACCAUCGUAUAGUGAAGGUGACGUAUACACGCGGAGGGAUUCUAGUGUCAAGACCA CCCCGGCCUGUACGGUAAACGGUAGGGUACGCCGUGAGCUAUGACCUACGUCUCCCAUUGUGGCGGCUUUAAG CAGACACUGUCAAGGAGAAAGAGUCUUAUUCUUGAUGACCUACGUCUCCCAUUCUACCAUUCUGUAUCAAUUGA CUGGCAUACUAGCGACCGACGUCACACCGGAGGACGCACAGAAGUUUUAUGUGGGAUUAAGUACAGAGGAUAGUU GUGAACGGAAGAACACAGCGAAACACUACACGAUGAAGAAACCUUGGUGUCCGAGAGAGGUCACUUAUAG CAAGUGGCGAGGGAUAUACAAGGCAGACCUUGAUGAUGAAGAAACCUUGGUGUCCGAGAGAGGUCACUUAUACU UGCUGCUGCUUGUGGGCAUUAACGAGGAGAUUCCGAGCCUUAUGGUCUACAGGCCUCCGCAUCCAGUCAGAU GAAGGUGCCUUCAGAGUUUAACUCGUUCGUUACCAAGCAGAGAGUUAUUAUACCUUUCGACGCGUCGUCAGCCAGG CACGCAUUAAGAUUGCUUUUGGCAAGAAAGACCAAGCAGAGAGUUAUUAUACCUUUCGACGCGUCGUCAGCCCAU GAUGCUGAACAAAGAGGAGAGAGGUGGAGGCGCGAGCUGACUAGAGAAGCCUUAACCCCUUCGUCUCCCAU CGCGCCGCGGAGACGGGAGUCGUCGACGUCGACGUUGAAGAACUAGAGUAUACGAGGAGGAGGUGGUGG AAACACCUCGCAGCGGUUGAAAGUCACCGCACGCGGACGACUACUAGGAAAUUACGUAGUUCUGUCCCC CGCAGACCGUUCUAGAGCUCAAGUUGGCCCCGUGACCCUUCUAGCAGAGCAGGUGAAAUUAACACAU ACGGGAGGGCCGCGUUAACCGGUUCGACGGAUUAUGACGGCAGGCGCACUAGGUGUACAACGAAAGGAGUUCGUAACAGGAAACU GUCCCCUAGUUUCAGGCUUUGAGCGAGAGCGCACUAGGUGUACAACGAAAGGAGUUCGUAACAGGAAACU AUACCAUUAUUGCCGUUCACGGACCCUUGACUACCGACGAGGAGAAUACGAGAAAGUCAGAGCUGAAAGAAC UGACGCCGAGUACGUGUUCGACGUAGAUAAAAUUGCUGCGUCAAGAGAGAGGAAAGCGGUGUUGGUGUUG
---	------------------------	---

	UAGCGCCUAAUUUCUCUGCCACGACUGAAGCGGAAGGGACCGCGAAUUGGCCGCGUCUACCGGGCAGUGGCC GCCGAAGUAAACAGACUGUCACUGAGCAGCGUAGCCAUCCCGCGUCGUCCACAGGAGUGUUCAGCGCGGUAG AGUAGGCGUGCAGCAAUCCCUCAACCAUUAUUCACAGCAUUGACGCCACGGACGCGUGACGACCAUCUACUG CAGAGACAAAAGUUUGGAGAAGAAAUCCAGGAAGCCAUAGACAUAGGACGCGCUGUGGAGUUGCUCUAAUGAU GACGUGGAGCUGACCACAGACUUGGUGAGAGUGCACCCGACAGCAGCCUGGUGGGUCGUAAGGGCUACAGUAC CACUGACGGGUCGCGUACUUCUUGAAGGUACGAAAUCAACAGGCGUCUUAUUGAUUGGCGAGAGUAC UGACGUUUGUGGCCAGACUGCAAGAGGCAACGAGAUUUGCCUUAJACGCGCUGGGCGAAACAAGGACAACA UCAGAUCCAAUUGUCCGGUGAACGAUUCCGAUUCAACACCUCGCCAGGACAGUGCCCGUGCCGCUACG CAUAGACAGCAGAACGGAUCCGCCGCUUAGGUCACACCAAGUUAAGCAUGGUGUUGCUCAUCUUCUCC CUCCGAAUACCAUGUAGAUUGGGUGCAGAGGUAAGUGCGAGAGGUUCUCCUGUUCGACCCGACGGUACC UUCAGUGGUUAGUCCGGAAGUAUGCCGCAUCUACGACGGACCACUCAGAUCCGUCGUUACGAGGUGUUGACU UGGACUGGACCAACGACUGUCUUCACUGCCAGCAUACCAUGUCGUAACCCAGUUGGAGUCGUGACAUUG ACUCGAUCUACGAGCAAUUGGCUCCAUAGUAGUGACGGCUGACGUAACCCUGAACCCGAGGCAUCGCGGACC UGGCGGCGAGUUGCAUCCUGAACCCGACACCAUGUGGACCGGAGAACCCGAUUCUCCACCGGCCCAAAGA GAGCUGCAUACCUUGCCUCCCGCGGGGAGCGACCGGUGCGCGCGCGAGAAAGCGACGCCUCCCAAGGA CUGCGUUJAGGAACAAGCUGCCUUGACGUUCGGCGACUUGACGAGCAGGAGGUCGAUGCGUUGGCCUCCGG GAUUAUUUCGGGACUUCGACGACGUCCUGCGACUJAGGCCGCGGGUGCAUUAUUUCUCCUGGACACUG GCAGCGGACAUUACAACAAAUCCGUUAGGCGACACAUCUCCAGUGCGCACAAUCUGGAGCGGUGGAG GAGAAAUGUACCCGCCAAAUUGGAUACUGAGAGGGAGAGCUGUUGCUGUGAAAUGCAGAUCCACCAUC GGAGGCUAAUAGAGUCGUAACAGUCUCGCAAGUGGAGAACAUAAAGCCGUAACGCGUUCGGUACCCCGCCGUGU CGGGGGCCAGAUUGUACACGGGAGCGGACGUAAGCCGCAUACCAACAUAACGCGUUCGCAACGAUACCUAUCAGAA ACUCCCCUACCGUGAUCGAAAGAUUCUACAGCCCGAUUGAGCAUUCGAGCGUGCAACGAUACCUAUCAGAA AUUACCCAAACAGUGGCGUCCAGAUACAGAUACAGAUAAUACGACGCAUACUUGGACAUGGUUACGCGGUGGAU AGUUGCUUGGACAGAGCAGAUUCUGCCCGGAGCUCGCGGUGUACCCGAAACAUCUAGCCACCAAGCCG ACUGUACGAGUGCGUCCGUCACCCUUCAGAACACACUACAGAACGUGCUAGCGGCGUCCCAAGAGAAAC UGCAACGUCACGCAAUUGCGAGAACAACCCACCAUGGACUGGAGUGUUAACGUGGAGUGCUUAAGCGCUA UGCCUGCUCCGGAGAAUUGGGAGAAUUGCUAAACAACCUAUCGGAUAACCAUCUGAGAAACAUCUACUACCUA UGUGACCAAAUUGAAAGGCCGAAAGCUGCGCUUGUUGCUAAGACCCACAACUUGGUUCCGUGCAGGAGG UUCCCAUGGACAGAUUCACGGUCGACAUAGAAACGAGUUGUCAAAGUCACUCCAGGGACGAAACACACAGAGGAAA

		GACCCAAAGUCCAGGUAUUCAAGCAGCGGAGCCAUUUGGCGACCGCUUACCUGUGCGGCAUCCACAGGGAUUUAG UAAGGAGACUAAAUUCUGUGUUAACGCCCUAACGUGCACAUAUUGUUAUUGUUAUUGCGGCCGAAGACUUAUGACGCG AUCAUCGCCUCACUCCACCAGGAGACCCGGUUCUAGAGACGGACAUUGCAUUCGACAAAAGCCAGGAC GACUCCUUGGCUCUACAGGUUUAUGAUCCUGAAGAUUAGGGUGGUAUAGUACCUUGGACUUGAUUG AGGCAGCCUUUGGGAAUAUCCAGCUGUACCUACCAACUGGCGCGCUUAAAGUUCGGAGCUAUGAUGAAA UCGGGCAUGUUUCUGACUUUGUUUAUAACACUGUUUUAACAUACCAUAGCAAGCAGGGUACUGGAGCAGA GACUCACUGACUCCGCCUGUGCGGCCUUAUCGGCGACGACAACAUUGUACGGAGUGAUUCUCCGACAAGCUGA UGGCGGAGAGUGCGCGUGUGGGUCAACAUUGGAGGUGAAGAUCAUUGACGCUUGCCGUGUUCAGACCCACUUAAGC AUUUUUUGUGGGGAUUAUAGUUUUGACAGCGUCACACAGCCCGUGCGUGUUCAGACCCACUUAAGC GCCUGUCAAAGUUGGUAAGCCGUAACAGCUGAAGACAAGCAGGACGAGACGAGCGACGAGCAGUGAGC GAGGUUAGCAAGUGUUCGGACAGGCUUGGGGCCGAAUCUGGAGGUGGCAUAACAUAGGUAGGUAUGAGGUAG AGGGCUGCAAAAGUAUCCUAUAGCCAUUGGCCACCUUGGCGAGGACAUUAAGGCGUUUAAGAAUUUGAGAGGA CCUGUUUAACACCUACCGCGGUCCUAAGAUUGGUGGUUAUAACACAGAAUUCUGAUUAUAGCGCACUAUUAU AGCACCAUAGU AUGAGAGUAGUGGCCCCAGAACCCUGAUCCUGCUGUCUGGGCCCCUGGCCUUGACAGAGACAUGGGCCGG AAGCAGAACCAACAUCGAUGCCAUCAAAGACAUGAAGGACUGAGACUGGUGGCCUUAACUGGAUUCUGUGG GAGUGGGACAAUUGGAUUGGAGACACAGGACCUAGCUGUGAAGGGACUGACAAUCACAAAGGAAGGCC GAAAGAGACUGAGAAAGAGACUGGUGGAUUUUGAAGGAUUGUAAACACCUAUGUGAAAGUGCCUCUGAAAC AGCACCAUUGAUUGCUGGUGAGCCUGGUUACAACAUUGGACCCACAACUUAAGACAAGCACACUGCUG AAGAAGCUGAUGCUGGAGAUUAUUGGAGCUGCUGAUCAGUUCUGGUGUGGAACAAAGGCAGAGUGGAUG GCAACUGGUGUGAUCAAAGGCCUGGCCAACAGGAGAGCCAAAGAAAGAAACAGUUCUCCUGGGGAGAAUGAUGA CUCGAGGGCGCGCAGGAGCUAAUUCGACGAUAUUGGAUUUUUAUUUUGCAUUGGUUUUUUAU AUUUCCAAAAA AA AA
Secretion signal		
10	Sec (amino acid)	MRVMAPRTLIIISGALALTETWAGS
5'-UTR (hAg-Kozak)		
11	5'-UTR	AUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACC

3'-UTR (Fl element)	
12	CUUGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUCCCCGACCUCCGGGUC CCAGGU AUGCUCCCAUCCACCUCCACCUCCACCUCCACCUCCAGUUCAGACACCUCCCAAGCAGCAGCAAU GCAGCUCAAAACGCUUAGCCUAGCCACACCCACCGGAAACAGCAGUUAUAAACCUUUAAGCAAUAAACGAAAGU UUAACUAAGCUAUACUAAACCCAGGGUUGGUAUUCGUGCCAGCCACACC
A30L70	
13	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Lysin 10473	
14	MAKTKKQIEQRLQAYAKGTVDSPYRITKPTSDPSFGVMEVG AIDADRYHAQCQDLITDYVLWLDNKNRVPWGNAKDQI NQSYGTGFKIHKNLPSVAKTGWIAVFTSGTYSQYGHIGIVSPGDTNKFHIEQNWNWGWANKKPLTRWDNYYGLTHFIEV PYAKETITTKTAPKAATSKPAKKISVSHNNINLQSNFAKRGYNPKGVVHNDAGRSTAAQYENSLKGANQNRLNNGIAHAY ADSSHIWYAIIEPERIAVHTGDGVKKGTGNVDYFGIEVCQSMSASDKEFMANEQVWLQEVARLLKWKLPANRNTVRLHLEF SATSCPHRSLTMRTGNPVTQGRPSDAIKKLKDIWIKIRAYMEGKTPSIKTVANKPGSASTPATKNSDGWKVNAYNTYYK TERASFTVGSTRIAVRSTGPFITTCPTGWLNPGVITITYDEVVMVQNGFVWLGYDSYNGRRYVPIRTAQGTPGSSNVVLGPLW GTIS

15	nucleotide sequence AUGGCUAAGACUAAAGAAACAAUUCGAGCAAAAGAUUGCAAGCUUAUGCUAAAGGUACUGUUGAUUCUCCUUACA GAUUUACUAAAGCCAAACUUCUUAUGAUCCUUCUUUCGGUGUUAUGGAAGUUGGUGCUUAUUGAUGCUGAUAGAU ACUAUCAUGCUCUAAUGUCAAGAUUUGACUGAUUACGUUUUUGUGGUUGACUGAUAAACAAAGUUAGACCAUG GGGUAAACGCUAAGGAUCAAAUUAUCAAUCUACGGUACUGGUUUUAAAGAUCCACAAGAAUUUGCCUUCUACU GUUGCUAACACUGGUUGGAUUGCUGUUUUCACUUCUGGUACUUCUCUCAAUUGGUCAUUAUUGGUUAUUGUU UACUCUCCAGGAGAUACUAAACAAGUUCACAUCAUCGAACAAACUGGAACGGUUGGGCUAAACAAGAAACCAUU GACUAGAUGGGAUAAACUACUACGGUUUGACUUAUUCAGGAUUCCUUAACGCUAAGGAGACUACACUACU AAGAAACUGCUCUCCAAAGGCGUACUUCUAAACCUGCUAAGAAAUUUCUGUUUCUCAUAAACAACAUUAACUU GCAUCUAAUUUUGCUAAAAGAGGUUACAACCCAAAGGGUGUUGUUUAUUCACAAUUGAUGCUGGUAGAUUCUACU GCUCAACAUAACGAAAUUUCUUUGAAGGGUGCUAACCCAAAUAAGAUUGAACAAUUGGUUAUUGCUCAUGCUUACG CUGAUUUCUUCACAUUUGGUUUGCUAUUUGAACCCUGAGAGAAUUGCUUACCAUACUGGAGAUUGGUUUAAGAA AGGUACUGGUAAACUACGAUUACUUCGGUAUCGAGUUUUGUCAAUUAUGUCUGCUUCUGAUAAAGAGUUUAU GGCUAAUGAGCAAGUUGUUUUGCAAGAGGUUGCUAGAUUUGUUGAAGAAUUGGAAGUUGCCAGCUAACAGAAA CACUGUUAAGAUUGCAUUGGAUUUUCUGCUACUUCUUGCCUCACAGAUUUIUGACUUAUGAGAACUGGUUU CAUCCAGUUAUCUCAAGGUAGACCUUCUGAUGCUAUUAUUAAGAAUUGAAGGAUUAUCUGGAUCAAGGAAUUC AGAGCUUACAUGGGAGGUAAAACUCCAUUAUAAACUGUUGCUAAAGCCAGGUUCUGCUUCUACUCCUGC UACUAAAUUCUAAUUCUGAUGGUUGGAAGGUUAACGCUUACAACACUUAACUACAAGACUGAGAGCUUCUUUU ACUGUUGGUUUCUACUAGAAUUGCUGUUAAGAUUCUACUGGUCCAUUCACUACUUGUCCUUUCACUGGUUGGUUG AACCCAGGUUGUUAUCUACUUAACGAUGAAGUUUAUGGUUCAAACCGGUUUUCGUUUGGUUGGGUUAACGAUUCU UAUAAUGGUAGAGAUACGUUCCAAUUAAGAACUGCUCAAGGUACUCCUGGUUCUUAUUAUGUUUUGGGU CCUUUGUGGGGUACUUAUUUCUUA
----	--

16	nucleotide sequence (opt)	AUGGCCAAACCAAGAAACAGAUAGAACAGAGACUGCAAGCUUACGCAAAAGGGGACCGUGGAUUCCCCUUACCGA AUCACCAAGCCAAACAGCUACGACCCCUCCUUCGGUGUCAUGGAGGUUGGGGCUAUAGAUUGCGGAUCGCUACUA UCAUGCCAGUGUCAAGAUUGAUAAACUGACUAUGUACUUUGGCUACUGACAAACAAAGUAGAACCUUGGGU AAUGCCAAAGGACCAAGAUAAACAGAGCUAUGGACUGGUUUCAAAUCCACAAGAAUCUGCCAGUACCGUGGC UAAAACAGGAUGGAUUGCCGUGUUACAUCGCGGACCUACUCACAGUACGCGCAUCGCAUCGUGUACAGUC CAGGAGAUACCAACAAUUCCAUAUCAUCGAGCAGAAUUGGAUUGGAUUGGGCAACAAAGAAACCUUGACACGU UGGGACAACUACUACGGGUGACCCACUUUAUCGAGGUCCCAUUGCGAAGAGACAUAUCACACGAAGAAAC UGCACCUAAAAGCAGCUACCUCAAAACCGGCUAAGAGAUUCAGUUUCACACAACAACUACCAACUACAGGCAA CUUUGCAAAGAGAGGGUACAACCCAAAGGGAGUGUGAUUACAAUAGACGCGACGCUACUACUGCAUAUGCUGACAGC AUGAAAACAGCCUGAAGGGGCUAAUCAGAACAGGCUGAUAAACGGCAUUGCCCAUGCAUAUGCUGACAGCAGC CACAUUUGGAUUGCUAUCGAACCCGAGAGAAUCGCCUAUCACACAGGCGAUGGCGUUAAGAAGGGAACAGGCAA UUACGACUACUUCGGUAUUGAAGUAUGCCAGAGCAUGAGCGCAAGUGACAAGAGUUCAUUGGCCAACGAACAG GUGGUACUGCAGGAGGUGGCUAGGUUGCUCAAGAAUUGGAAGCUUCCCGCAUUGGGAUACCGUGCGGUCC AUCUGGAUUUCUCGCCACAUUCUUGCCCUAAGGAGCCUACAUAUGCGGACCGGUUCAAUCCGGUGACUCAG GGUAGACCCUCAGACGCCAUCAUCAAGAAGUUAGAGGAUAUCUGGAUAAAAGAGAUUCCGCGCAUAUAUUGGAGG GUAAAACCCCAUCCAUAAGACUGUCGCCAAUAAAGCCCGGAUCAGCCAGUACCCUUGCAACCAAAUCCAAUCCGA UGGUGGAAGGUCAUUGCCUACAUAACCUACAACAAACGGACGGCCUCCUUUACAGUGGGGCUACUCGGA UUGCUGCGUAGUACAGGACCCUUUACAUAUGCCCGUUUACAGGGUGGCUGAACCCCGGAGUCACGAUAACC UAUGACGAGGUUAUGGUGCAGAACGGGUUUGUGUGGUUGGGCUAUGAUAGCUACAUAUGGGAGGAGGUUGU CCCCAUUCGCACUGCCCAAGGAACUCCUGGCAGCUCCAUAUAUGUGCUUGGCCUUCUGUGGGGACGGAUUUCUU GA
Lysin L0475	amino acid sequence	MQAKLTKEFIEWLKTSEKQFNADLWYGFQCFDYANAANKVLFGLLLKGLGAKDIPFANNFDGLATVYQNTPDFLAQPG DMVVFSGNYGAGYGHVAVWIEATLDYIIVYEONWLGSGWTDGIEQPGWGWKEVTRRQHYDFPMWFIRPNFKSEIAPR SVQSPTQAPKKETAKPQKAVELIKIDVVKGYDLPKRGSNPKGIVIHNDAGSKGATAEAYRNLVNAPLSRLEAGIAHSYVS GNTVWQALDESQVWHTANQLGNKYYYGIEVCQSMGADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEFTSTS CPHRSSVLHTGFDPTVTRGLLPEDKRLQLKDYFIKQIRAYMDGKIPVATVSNDSASSNTVKPVASAWKRNKYGTYYMEESAR FTNGNQPIVTRKVGPFYCPVGYQFPQPGGYCDYTEVMLQDGHVWVGYTWEGQRYVLPRTWNGSAPPNQILGLDLWGEIS
17		

39

19	nucleotide sequence (opt)	AUGCAAGCGAAACUGACCAAGAGGAGUUUAUAGAGUGGCGUGAAAAACCCUCUGAGGGUAAACAGUUCAACGCCGA UCUGUGGUAUGGUUUCAGUGCUUUAUUAUACGCAACGCGUCUUGGAAAGUGCUGUUUGGCCUCGCCUGAA GGACUCGGUGCCAAAGACAUCUUUUGCCAAUUAUUGACGGGCGUGCAACCGUACCAAGAAUACUCCUG ACUUUCGCGUCAGCCAGGGGACUAGGUGUUGCGGCUCCAAUUAUUGGGGCGAGGCUAUGGACACGUAAGCAUG GGUAUAGAACCCACUCUGGACUACAUAUCGUACGAGCAAAUUGGCUUGGCCGUGGGUGGACCGGAUGGC AUUGAGCAGCCUGGCGUGGGAGGGAGAGGUUAUCUGUAGACAGCAUGCUUACGACUUAUCCCAUGUGGUUCA UCCGUCCCAACUUAAGUCCGAGAUUGCGCCAGAGCGUCCAGUCUCCAAACCCAGCUCGAAAGGAAACAG CCAAGCCACAGCCUAAGGCCGUUGAACUGAAGAUUAUUAAGGAUGUGGUAAGGGAUAUAGCCUCCCGGAAAGA GGAGCAAUCCCAAGGGAUUGUGAUCCACAACGAUUGCGGGUUAUUAAGGGGCAACAGCAGAAAGCUAUAGGAA CGGUUGGUAAUUGCCUUCAGUCGCUUGGAGGCCGGAUUGCACACUUAUUGUCCGGGAAUACGCUU UGCAAGCUCUCGACGAAAGCCAGUAGGCGGCAUACAGCAACAGCUGGGCAACAAUACUACUAGGCAU CGAGGUGGCGAGUAUGGCGCAGAUUAUGCCACGUUCCUGAAGAACGAGCGCUACCUUUAAGAGUGU GCGCGGUUGCUGAAGAGUGGGGACUGCCAGCUAAUCGCAACACAAUAGGUAGACAGGGGACUGCGCCGAAUA UAGCUGUCCACACAGAAAGUAGCGUGCUGCACACAGGAUUCGAGUCCUGUAGACAGGAAUUAUCCGUCGCUACU AGCGGCGCAGCUGAAAGACUACUUAUCAAACAGAUUCGGGCUUAUAGGACGGAAGAUUCCCGUCGCUACU GUUAGCAUUGACAGCUCUGCCUUAUCAAUACCGUAAACCGGAGCAAGCGCCUGGAAAGGAAACAGUAUGG CACUUAUUAUAGGAGGAUUCGCGGCUUAUUAUACCAAGGUAUACAGCCCAUACAGUCCGCAAGGUCGGGCAU UCCUGUACUGUCCCGUAGGCUACCAUUAUACGCGGUGGUGGUAUCUGGACUUAUACCGAGGUGAUGCUGCAAGA UGGCCAUGUCUGGGUUGGCUUAUACCGGGGAGGCGAGCUCUACUUGCCAAUACGGACUUGGAAUGGUUCU GCCCCACCCAAACAGAUCCUCGGAGAUUCUGGGGCGGAAUUCUCCUGA
Lysin L0479	amino acid sequence	MQAKLTKEFIEWLKTSEGKQFNVDLWYGFQCFDYANAGWKVLFGLLLKGLGAKDIPFANNFDGLATVYQNTPDFLAKPG DMVVFSGSNYGAGYGHVAWVIEATLDYIIVYEQNWLGSGWTDRIEQPGWGWEKVTTRQHALRFPYVVPSPKSETAPRS IQSPTQASKKETAKPQKAVELKIKDVVKGYDLPKRGGNPKGIVIHNDAGSKGATAEAYRNLVNAPLSRLEAGIAHSYVSG NTVWQALDESQVWHTANQLGNKYGYGIEVQCSMGADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEFTSTSC PHRSSVLHTGDFPVTGRLLPEDKRLQLKDYFIKIRAYMDGKIPVATVSNESASSNTVKPVASAWKRNYGTYYMEESARF TNGNQPTVRKVGPFLSCPVGYYQFQPGGYCDYTEVMLQDGHVWVGYTWEGQRYLPIRTWNGSAPPNQILGDLWGEIS

21	nucleotide sequence	AUGCAGGCUAAGUUGACUAAGAAAGAGUUUAUUGAGUGGUUGAAGACUUCUGAAGGUAAACAAUUCACCGUU GAUUUGUGGUACGGUUUUCAAUGUUUCGAUUAUGCUAUAUGCCGUUGGAAAGUUUUGUUUGGUUUUGUUUU GAAGGGUUUUGGGUGCUAAGGAUAUCCCUUUCGCUAACAAUCUUGCUUUGGCUACUGUUUACCAAAACACU CCAGAUUUCUUGGCUAAGCCUGGAGAUUUGGUUUGUUUUGGUUUAUUAACGGUGGUGGUUAUGGUCAUGUU GCUUGGGUUUAUUGAAGCUACUUUGGAUUAUCAUCAUGUUUACGAGCAAAUUGGUUGGGUGGUUGGACU GAUAGAAUUGAACACCAGGUUGGGGUUGGGAGAGGUUACUAGAAGACAACACGCUUUGAGAUUCCCUUAUG UUGUUUACCCAUUUAUCCUAAUUCUGAAACUGCUCCAAGAUUAUUAUCAAUCCUACUCAAGCUUCUAAAGAA GAAACUGCUAAGCCACAACCUAAAGCUGUUGAGUUGAAGAUUAUUAAGGAUGUUUAAGGGUUACGAUUUGC CAAAGAGAGGUGGUAACCCUAAAGGUUUCGUUAUCCAUAAACGAUGCUGGUUUCUAAAGGUGCUACUGCGAAGC UUAACAGAAACGGUUUUGGUUAAUGCUCUCCAUUGUCCAGAUUGGAAGCUGGUUAUUGCUCAUUCUUAUGUUUCUGG UAACACUGUUUGGCAAGCUUUUGGAUUAUCUCAAAGUUUGGUUGGCACACUGCUAACCAUUGGGUAAUAAAGUAC UAUUAACGGUAUUGAGGUUUUGCAAUCUUAUGGGUGGCUGAUAACGCUACUUAUUAAGAAACGAACAAAGCUACU UCCAAGAGUGUGCUAGAUUGUUGAAGAAUUGGGGUUUUGCCAGCUAACAGAAACACUACAGAUUGCACAACGA GUUUAUUAUUAUUGUCCUUAAGAUUCUUGUUGUUUGCACACUGGUUUCGAGUUCAGUUACUAGAGGUUU GUUGCCUGAGGAUAGAGAUUGCAAUUGAAGGAUUAUCUUAAGCAAUUCAGCAAUUCAGAGCUUACAUUGGUA AUCCAGUUGCUACUGUUUUAACGAUUCUUGCUUCUUAUUAUACUUAUUAACCUUGUUGCUUGCUUGGA AGAGAAACAAUACGGUACUUAACAUGGAAGAGUCUGCUAGAUUACUAAACGGUAAUCAAACCAUUAACUGUU AGAAAGGUUGGUCCAUUUUUUGUCUUGUCCUUGGUUUAACCAUUCCAACCUUGGUUACUGUGAUUAUACU GAAGUUAUGUUGCAAGUGGUCAUGUUUGGUUGGUUAUACUUGGGGUGCAAGAUUAUUAUUAUUGCCAAU AGAACUUGGAACGGUUCUGCUCCACCUAUCAAUUUUUGGGAGAUUUGGGGUGAAUUUUCUUA
----	------------------------	--

42

24	nucleotide sequence	AUGGCUAAGACUCAAGCUGAGAUUAUAAAGAGAUUGGAUGCUUACGCUAAGGGUACUGUUGAUUCUCCAUACA GAGUUAAGAAAGCUACUUCUUAUGAUCCUUCUUCGGUGUUUUGGAAGCUGGUGCUAUUUGAUGCUGAUGGUU ACUAUCAUGCUCUAAUGUCAAGAUUUGACUGAUUACGUUUUGUGGUGACUGAUAAACAAGGUUAGAACUU GGGUAAACGCUAAGGAUCAAAUUAAGCAUUCUACGGUACUGGUUUUAAAUCCAUGAAAAACAAGCCAUUCUACU GUUCCUAAAGAAAGGUUGGAUUGCUGUUUACUUCUGGUUUCUACGAGCAUUGGGGUACAUUGGUUUAUUGU UACGAUGGUGGUAAACACUUCUACUUUUAUUAUUGGAACAAACUGGAAUGGUUACGCUAACAAGAAACCAA CUAAGAGAGUUUGAUAAACUACUACGGUUUGACUCACUUCAUUGAAUUCUGUUAAGGCGUGUACUACUGUAA GAAAGAGACUGCUAAGAAUUCUGCUUCUAAACUCCAGCUCUAAAGAAAGGCUACUUGAAAGUUUCUAAAGA ACCAUAUCAACUACACUAUGGAUAAAGAGAGGUAAAAGCCAGAGGUUUGGUUUUUCACACGAGUCUGGUAG AUCUUCUGGUCAACAAUACGAGAACUUCUUGGCUAAUUGCUGGUUACGCUAGAUUAGCUAACGGUUAUUGCUCAU UACUAUGGUUUCUGAAGGUUAUGUUUGGGAGGCUAUUGAUGCUAAAAUCAAUUUGCUUGGCACACUGGGAGAU GGUACUGGUGCUAACUCUGGUAUUUUAGAUUCGUGGUUUAUGAGUUUGCAAUUUGCAUUCUUGCUGCUUCUGAU GCUCAUUUUUGAAGAACGAGCAAGCUGUUUUCCAAUUCACUGCUGAAAAAUCAAAGGAGUGGGGUUUGACUC CUAACAGAAAGACUGUUAGAUUGCAUUGGAUUUGUUCUCCACUGCUUGCCUUAJAGAUCAUUGGUUUUGCA CACUGGUUUCAAUCCAGUUACUCAAGGUAGACCUUCUCAAGCUAUCAUGAACAUAUUGAAGGAUUAUUAUUAUCA AGCAAUUUAAGAAUUAUACUUGGAUAAAGGUACUUCUUCUACUGUUGUUAAAGAUUGGUAAAAACUUCUUCUGC UUCUACUCCAGCUACUAGACCUGUUACUGGUUUCUUGGAAAAAGAACCAUACGGUACUUGGUUAUAAACCCAGAGA AUGCUACUUUUUGUUAACGGUAAUACAACCUAUUUGUUAACUAGAAUUGGUUCUCCAUUCUUGAACGCCUCCUGUUGG UGGUAAUUUGCCAGCUGGUGCUACUACGUUUAACGAUGAAGUUUGUAUCCAAAGCUGGUCAUUAUUGGAUUUGG UUACAACGCUUAUAAACGGUAAUAGAGUUUAUUGUCCAGUUAGAAACUUGUCAAGGUGUUCACCACCUAAUCAAUU CCUGGUGUUGCUUGGGGUGUUUUAAGUAA
----	------------------------	---

25	nucleotide sequence (opt)	AUGGCCAAGACACAGGCCGAGAUCAACAAGCGGCGUGGACGCCUUAUGCCAAAGGGCACAGUGGAUAGCCCCUACAGA GUGAAGAGGCCACACGCUACGACCCCGAGCUUCGGCGUUAUGGAAGCGGCGCUUAUGACGCCGCGCUAUUA UCAUGCCAGUGCCAGGACCUAGUACACCGACUAGUGUGGCGUAGACCGACACAAAGUGCGGACAUUGGGGCA ACGCCAAGGACCAGAUCAAGCAGAGCUACGGCACCGGCUUAAGAUCCACGAGAACAAAGCCUAGCACCGUGCCUA AGAAAGGCGUGGUGUUAUACCAAGCGGCGAGCUAUGAGCAGUGGGGCCACAUCGGCAUCGUGUACGAUGG CGCAAUACACGACCUUACCAUCCUGGAACAGAACUGGAACGCUACGCCAACAAAGAACCCCAAGAGAGU GGACAAUACUACGGCCUGACACUUAUCGAGAUCCCGUGAAGCGCGCACCCGUGAAGAAAGAGACAGC CAAGAAGUCCGCCAGAACCCUGCUCCUAAGAAGAAAGCCACACUGAAGGUGUCCAAAGAACCAUCAACUA CACCAUGGACAAGCGGGGCAAAAGCCGAGGUAUCCAGAUACGCCAAUGGAUUGCCCAUACUAGCGCGCAGCA GUACGAAAACAGCCUGGCCAAUGCCGCGUAGGUAUCCACAACGACGCGGCGAGCGGAAACCGGCCAAUUC GCUAUGUGUGGGAAGCAUCGACGCCAAGUUGGCCAGAGCAUUCUGCCAGCGACGCCAGUCCUGAAGAAUGA GGCAUUCAGAUUCGCCGCAUCGAGAGUUGGCCAGAGCAUUCUGCCAGCGACGCCAGUCCUGAAGAAUGA GCAGGCGUGUUCAGUUAACCGCGAGAGUUAAGAGUGGGGACUGACCCUAAACAGAAAGACCGUGCGGC UGCACAUUGGAUUCGUGCCUACCGCAUGUCCCAACCGCUAUGGUGUCACACCGGCUUAAGCAGAUCAAGGAU AGGGCAGACCUAGCCAGGCCAUCAUGAAACAGCUGAAGGACUACUUAAGCAGAUCAAGAACUACAUUGGAU AAGGGCACCCAGCAGCACCGUGGUCAAGGAUGGCAAGACCAGCGCGCCUACACCCAGCCACAAGACCCUGUG ACAGGCAGCUGGAAGAAACAGUACGGCACCUUGGUACAAGCCGAGAACGCCACCCUUCGUGAACGGCAACCAG CCUUCGUGACCAAGUUCGGCAGCCCCUUCUGAAUGCCCCUGUCGGAGGAAUUCUUCGUGGCGGCCCAUC GUGUAUGACGAAGUCUGUAUUCAGGCGGCCACAUCUGGAUCGGCUACAACGCCUACAACGGCAACAGAGUGUA CUGCCCGUGCGGACUUGUCAGGGCGUUCACCCUAUACAGAUUCCAGGCGUGGCGUGGCGUGUUCAAAUGA
Lysin L0482		
26	amino acid sequence	REAPKTIQVTVKKGDTLSAIAIKYKTTVSNIQNTNNIANPNLIFIGQKLKVPMTPLVEPKPTVSSNNKSNSSSTLNLYLKTLEN RGWDFDGSYGWQCQFDLVNYYWNHLYGHGLKGYGAKDIPYANNFNSEAKIYHNTPTFKAEPGDLVVFSGRFGGYYGHTAI VLNGDYDGKLMKFQSLDQNNWNGGWWRKAEVAHKVHVHNYENDMIFIRPFKKA

27	nucleotide sequence	AGAGAGGCUCCUAAAGACUCAAUACACUGUUUAAAGAAAGGAGAUACUUUUGUCUGCUAUCGCUUUUGAAGUACA AGACUACUGUUUCUAAACAUCCAAAACACUAAACCAUCCGCUAACCCAAUUUUGAUUUUUCUUGGUCAAAAGUUG AAAGUCCAAUGACUCCUUUGGUUGAACCAAGCCUAAAGACUGUUUCUUCUAAACAAUAAAGUCUAAACUCUAAUUC UUCUACUUUGAACUACUUGAAGACUUUGGAGAAUAGAGGUUGGGAUUUUUGAUUGGUUCUUAUGGUUGGCAAU GUUUCGAUUUGGUUAAACGUUUACUGGAAUACUUUGAUUGGUCACGGUUUGAAGGGUUACGGUGCUAAGGGAUA UCCCUAACGCUAAACAUCAACUCUGAAGCUAAGAUACCAUAAUACUCCAAACUUUCAAAGCUGAGCCUGGAG AUUUGGUUUUUUCUCUGGUAGAUUCGGUGGUGGUUACGGUACACACUGCUAUUUGUUUGAACGGAGAUUAC GAUGGUAAAUGAUGAAGUCCAAUCUUUGGAUCAAUUUGGAACAAUUGGUUGGAGAAAGGCGUGAAGUU GCUCAUAAAGGUUGUUCACAAUCAGAGAACGAUAUGAUUUUUCAUUAGACCAUUCAAAGAAAGCUUAA
28	nucleotide sequence (opt)	AGAGAGGCUCCCAAAACACAGAUUUUACACCGUCAAGAAAGGGAGAUACGCUUAGCGCCAUUUGCACUGAAAUACAA GACCACCGUAUCAAAACAUCCAGAACACCAAUAAACAUAGCCAAUCCCAUUCUACUUCUUAUUGGCCAGAACUCUCAA GGUCCCAUGACUCCACUGGUCGAAACCGAAACCUAAGACAGUGUCCUCCAAUAAACAAGAGUAACAGCAACUCCUC UACACUGAAACUUCUCAAGACGUUGGAGAAUCGCGGUUGGGACUUUGACGGCAGUUACGGCUGGCAUUGCUUC GACUUGGUGAAUGUGUACUGGAAUACCUUGAUUGGGCAUGGACUGAAGGCUUUGGGGCCAAAGACAUACCUU AUGCUAAACAAUUUCAACUCUGAGGCCAAGAUUCACCAACACUCCUACCUUCAAAGGCAGAACCCAGGAGUCUGG UUUGUUUAGCGGACGGUUUGGCGGAGGCUAUGGGCAUACUGCCAUUUGCUUAAUUGGGGAUUACGAUGGCA AACUGAUGAAAUUCCAGUCACUGGACGAGAACUUGGAACAUUGGUGGUUGGCCGAAAAGCGGAAGUAGCCCAAG GUCGUCCACAAUACGAGAACGACAUCAUGAUCCGUCCCUUUAAGAAAGCGUGA
Lysin L0483		
29	amino acid sequence	AATHEHSAQWLNNYKKGYGYPPLGINGGMHYGVDFFMNIGTPVKAISSGKIVEAGWSNYGGNQIGLIENDGVHRQ WYMHLSKYNVKVGDYVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPLKSAGYKAGGTVTPTNTGWK TNKYGTLYKSESASFPTNTDIITRTTGPFRSMPQSGVLKAGQTHYDEVMMKQDGHVWVGYTNSGQRILYLPVRTWNKSTNT LGVLWGTIK

46

33	nucleotide sequence	GAUCAAUUCUUGGAUAAAAAGAGGGUAAUACUACUGCUUAUAGAGAUUGGUUCUGGUUUUUGGACUAAU UGUAGAGGUGCUACUUAUGGUUGAUGGUAAACACAGUUAUUCUGGUUAUGAGAUUGUUAAGGAAAAAGUGUGAU CAAGUUAACGCUAUUGAGAGAGAUAGGCUUUGGCUUGGGUUGAAAGAAUUAUAAAGGUUCCAUUGACUGAG CCUCAAAGAGCUGGUUUUGCUUUUUGUCCAUAACAUAUGGUCCUGGUAAAUGUUUCCCAUCUACUUAUCU ACAAGAGAUUGAAUUGCUGGAGAUAGAAAGGGUGCUUGUGAAGCUUAUAGAUUGGUGGAUUAAGGAUGGUGGUA GAGAUUGAGAAUUAAGAUACAACUACUGUUAACGGUAAUUAUAGAAAGAGAUCAAGAAUCUGCUUUUGGCUU GUUGGGUUAUUGAUCAAUA
34	nucleotide sequence (opt)	GAUCAGUUCUCGACGAGAAAGGCAAUACACACCACUGCGUAUAGGGAUGGCUUGGGAUAUGGACCAUCU GUAGAGGGCCACAAUUGGUGAUGGCAACCCGUAUACCCGGAUUGCGUCUCAAGGAGAGAGUGCGAUUCAG GUGAAUGCCAUUCGAACGCGACAAAGCUCUGGCCUGGGUUGAGCGGAACAUAAGGUUCCACUGACGGAACCA GAAAGCCGGAUUGCGUCCUUCUGUCCGUACAACAUAUGGCCUGGAAAGUGCUUUCUCCUACUUAACAAGC GCCUGAAUGCAGGGGACAGAAAAGGAGCCUUGUGAGGCUAUACAGGUGGGAUCAAAGAUAGGAGGAGGACUG UCGAAUCCGGAGCAACAACUGCUAUGGGCAAGUGAUUCGGAGAGAUACAGGAGAGUGCUCUGGCUUGCUGGGGU AUUGACCAGUGA
Lysin L0487		
35	amino acid sequence	MRYIPAAHHSAGSNKPVNRVVIHATCPDVGFPSASRKGRAVSTANYFASPSSGSAHYVCDIGETVQCLSESTIGWHAPPN PHSLGIEICADGGSHASFRVPGHAYTREQWLDPRVWPAVERAAVLCRQLCDKHGVPKRKLSVADLKAGRRGICGHTDVT AWHQSDHDDPGPWFWDKFMVWNGHGGSSSEELTVADVKALHNQIKLSAQLSGSVNKLHHDVGVVQVQNGDLS KRVDALSWVKNPVTGKLWRTKDALWSVWVYVLECRSRIDRLESVAVNGLKK

49

50

51

45	nucleotide sequence	AUGAACAGGUUGAUGCUUGGUUGUCUAAAGCAUGAUUUGAAAAUUUGACUACUUGAUCGUUGUUGGUGCU UUUAUUUUUGUUCUGAUCUUGUUGUUUAUGUUGAUGAACCAACUCUGAUCAGGUUUCUACAGAAUUAACUGUU UCUGAAAAUUGCUGAGUUGAGAACUGAACAAUUGCUGCUUACCCAGUUUAUCUACCAAGUUGAAAAAGGAGAU CUUUCACUAGAUUGUCUAAAGUCUGGUAUUGGAUUGAAGUUGAGUCCAGAGUUGGUUCUGAAAAAGGUUGGA UUGCUGGUUGGCAUACUUCUUGAACAUUGAGGCGUUAUCAAAGAAAGUUAACCAUUGAAGAACAAAGU UAUUGUUUGGAUCCUGGUCACGGUGGUAACGAUCAAGGUCUUCUUAUACUUAAGAGUUCUUGGA AAAGAACUACACUUUGAAGACUCGCUAAAGAAUUGCGUAGAGCUUUGGAAAAAGGCGUGCUACUGUUAAGUU ACUAGAUUCUGAUGAUACUUAAGUUUCUUUGGAUUGAAGAAAGUUGAACGGAGUGCUUACUACCAAGAGCUUUGGC ACGAUUCUUGGAUUCUCCUAAACGCUAAUUGGUGCUACUGUUUACUGGUACAAGAACCAAGAGCUUUGGC UGAGACUUUGAACACUACAUCCAAAGAAAGCUAUGUUUGUUAACAGAGGUACUAGACAAGAAACUUCCAAG UUUUGAGACAAACUAAACAGCCAGCUUUUUGUUGGAAUUGGUUAUUAUUAACCCUACUAGAGGUUAU GAUAGAAAUCAAUCCACAGACAAAGUUUGUUGAAGAGGCGUUGUUGAUGGUUUUGAAACAAUUAUUUCGCUUC UUA
46	nucleotide sequence (opt)	AUGAAUAAGGUGGUGCUUGGCUAGCAAAACAGACCUCUUAAGAACCUACUACCUUGAUCGUUGAGGAGCCU UCAUUCUGUUCUGAUAUCUGCUUGUUAUGCUGAUGAAUACUCUCUGAUCAGGGGUCAAACCGCAUAAACUGUUUC CGAGAAUGCCGAACUUCGGACAGAUCCAAACGCCCUUACCGGUCAUCUACAGGUCGAGAAAGGGGACACUU UCACCAGACUGUCUAAAGCGGCAAUUGGAUUGAGGUCGAGAGUAGAGUUGGAGUGAGAAAGGGGUGGUAJAGC AGGAUGGCACACAUCCUCGAAUUGAAAGCCGAUAGCAACAAGAAAGUGAACCCUCUCAAGAACAAAGUAGC UGCUGGAUCCUGGUCAUGGUGGCAACGAUCAGGGCGCUUCAAGCAUACCAAGAAAGUACUGGAGAGAA UUACACGCUGAAACAGCAAGGAACUGCGACGGGCUUUGAGAAAGCAGGAGCUACGGUGAGAAACUGACACGGU CUGACGACACCUAUGUGAGUCUGGACGACAGGAAGCUGAAUUGGCGACGCGUAUUAUCCAUUCACAAUGACAGC CUGGAUUCUCCCAUUGCAAUUGGGGCUACCGUCUACUGGUACAAGAACCAAGAGGGCCUGGCAGAAACCCUC AAUACCAACAUCCAGAGAGGCGAUUGUUGCCAAUAGGGGUACUAGGCAGGAAACUUUCAGGUGCUUAGAC AGACAAACAAACCCGAGUUCUGUUGGAACUCGGCUACUUCUCCAAACCAACAGACGAGGUGAUGAUCGCCAAUC AGAUUACCCGUCAGUAGUGGAAGAGCAGUGGUGGUGGCGGCUAGAAACAGUACUUUGCCAGUUUGA

Detailed Description of the Invention

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

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

The practice of the present disclosure will employ, unless otherwise indicated, conventional
20 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
25 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"
30 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".

- 5 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.

10 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.

15 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.

- 20 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 $\pm 10\%$, $\pm 5\%$, $\pm 4\%$, $\pm 3\%$, $\pm 2\%$, $\pm 1\%$, $\pm 0.9\%$, $\pm 0.8\%$, $\pm 0.7\%$, $\pm 0.6\%$, $\pm 0.5\%$, $\pm 0.4\%$, $\pm 0.3\%$, $\pm 0.2\%$, $\pm 0.1\%$, $\pm 0.05\%$, and for example $\pm 0.01\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 10\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 5\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 4\%$. In some embodiments, "about" indicates

deviation from the indicated numerical value by $\pm 3\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 2\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 1\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.9\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.8\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.7\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.6\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.5\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.4\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.3\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.2\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.1\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.05\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.01\%$. 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.

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

Definitions

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

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

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

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

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

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

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

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

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

5
$$I = \frac{1}{2} \cdot \sum_i z_i^2 \cdot c_i$$

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

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

"Osmolality" refers to the concentration of a particular solute expressed as the number of osmoles of solute per kilogram of solvent.

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

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

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

The term "recombinant" in the context of the present disclosure means "made through genetic engineering". In some embodiments, a "recombinant object" in the context of the present disclosure is not occurring naturally.

The term "naturally occurring" as used herein refers to the fact that an object can be found in nature. For example, a peptide or nucleic acid that is present in an organism (including viruses) and can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally occurring. The term "found in nature" means "present in nature" and includes known objects as well as objects that have not yet been discovered and/or isolated from nature, but that may be discovered and/or isolated in the future from a natural source.

As used herein, the terms "room temperature" and "ambient temperature" are used interchangeably herein and refer to temperatures from at least about 15°C, e.g., from about 15°C to about 35°C, from about 15°C to about 30°C, from about 15°C to about 25°C, or from about 17°C to about 22°C. Such temperatures will include 15°C, 16°C, 17°C, 18°C, 19°C, 20°C, 21°C and 22°C.

The term "EDTA" refers to ethylenediaminetetraacetic acid disodium salt. All concentrations are given with respect to the EDTA disodium salt.

The term "cryoprotectant" relates to a substance that is added to a formulation in order to protect the active ingredients during the freezing stages.

The term "lyoprotectant" relates to a substance that is added to a formulation in order to protect the active ingredients during the drying stages.

According to the present disclosure, the term "peptide" refers to substances which comprise about two or more, about 3 or more, about 4 or more, about 6 or more, about 8 or more, about 10 or more, about 13 or more, about 16 or more, about 20 or more, and up to about 50, about 100 or about 150, consecutive amino acids linked to one another via peptide bonds.

The term "polypeptide" refers to large peptides, in particular peptides having at least about 151 amino acids. "Peptides" and "polypeptides" are both protein molecules, although the terms "protein" and "polypeptide" are used herein usually as synonyms.

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

The terms "part" and "fragment" are used interchangeably herein and refer to a continuous element. For example, a part of a structure such as an amino acid sequence or protein refers to a continuous element of said structure. When used in context of a composition, the term "part" means a portion of the composition. For example, a part of a composition may be any
5 portion from 0.1% to 99.9% (such as 0.1%, 0.5%, 1%, 5%, 10%, 50%, 90%, or 99%) of said composition.

"Fragment", with reference to an amino acid sequence (peptide or polypeptide), relates to a part of an amino acid sequence, *i.e.* a sequence which represents the amino acid sequence shortened at the N-terminus and/or C-terminus. A fragment shortened at the C-terminus (N-terminal fragment) is obtainable, *e.g.*, by translation of a truncated open reading frame that
10 lacks the 3'-end of the open reading frame. A fragment shortened at the N-terminus (C-terminal fragment) is obtainable, *e.g.*, by translation of a truncated open reading frame that lacks the 5'-end of the open reading frame, as long as the truncated open reading frame comprises a start codon that serves to initiate translation. A fragment of an amino acid
15 sequence comprises, *e.g.*, at least 50 %, at least 60 %, at least 70 %, at least 80%, at least 90% of the amino acid residues from an amino acid sequence. A fragment of an amino acid sequence comprises, *e.g.*, at least 6, in particular at least 8, at least 10, at least 12, at least 15, at least 20, at least 30, at least 50, or at least 100 consecutive amino acids from an amino acid
20 sequence. A fragment of an amino acid sequence comprises, *e.g.*, a sequence of up to 8, in particular up to 10, up to 12, up to 15, up to 20, up to 30 or up to 55, consecutive amino acids of the amino acid sequence.

"Variant," as used herein and with reference to an amino acid sequence (peptide or polypeptide), is meant an amino acid sequence that differs from a parent amino acid sequence by virtue of at least one amino acid (*e.g.*, a different amino acid, or a modification of the same
25 amino acid). The parent amino acid sequence may be a naturally occurring or wild type (WT) amino acid sequence, or may be a modified version of a wild type amino acid sequence. In some embodiments, the variant amino acid sequence has at least one amino acid difference as compared to the parent amino acid sequence, *e.g.*, from 1 to about 20 amino acid differences, such as from 1 to about 10 or from 1 to about 5 amino acid differences compared
30 to the parent.

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

- For the purposes of the present disclosure, "variants" of an amino acid sequence (peptide or polypeptide) may comprise amino acid insertion variants, amino acid addition variants, amino acid deletion variants and/or amino acid substitution variants. The term "variant" includes all mutants, splice variants, post-translationally modified variants, conformations, isoforms, allelic variants, species variants, and species homologs, in particular those which are naturally occurring. The term "variant" includes, in particular, fragments of an amino acid sequence.
- 10 Amino acid insertion variants comprise insertions of single or two or more amino acids in a particular amino acid sequence. In the case of amino acid sequence variants having an insertion, one or more amino acid residues are inserted into a particular site in an amino acid sequence, although random insertion with appropriate screening of the resulting product is also possible. Amino acid addition variants comprise amino- and/or carboxy-terminal fusions
- 15 of one or more amino acids, such as 1, 2, 3, 5, 10, 20, 30, 50, or more amino acids. Amino acid deletion variants are characterized by the removal of one or more amino acids from the sequence, such as by removal of 1, 2, 3, 5, 10, 20, 30, 50, or more amino acids. The deletions may be in any position of the protein. Amino acid deletion variants that comprise the deletion at the N-terminal and/or C-terminal end of the protein are also called N-terminal and/or C-
- 20 terminal truncation variants. Amino acid substitution variants are characterized by at least one residue in the sequence being removed and another residue being inserted in its place. Preference is given to the modifications being in positions in the amino acid sequence which are not conserved between homologous peptides or polypeptides and/or to replacing amino acids with other ones having similar properties. In some embodiments, amino acid changes in
- 25 peptide and polypeptide variants are conservative amino acid changes, *i.e.*, substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains. Naturally occurring amino acids are generally divided into four families: acidic (aspartate, glutamate), basic (lysine, arginine, histidine), non-polar (alanine, valine, leucine, isoleucine, proline,
- 30 phenylalanine, methionine, tryptophan), and uncharged polar (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine) amino acids. Phenylalanine, tryptophan, and tyrosine are

sometimes classified jointly as aromatic amino acids. In some embodiments, conservative amino acid substitutions include substitutions within the following groups:

- glycine, alanine;
- valine, isoleucine, leucine;
- 5 - aspartic acid, glutamic acid;
- asparagine, glutamine;
- serine, threonine;
- lysine, arginine; and
- phenylalanine, tyrosine.

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

The terms "% identical" and "% identity" or similar terms are intended to refer, in particular, to the percentage of nucleotides or amino acids which are identical in an optimal alignment between the sequences to be compared. Said percentage is purely statistical, and the differences between the two sequences may be but are not necessarily randomly distributed over the entire length of the sequences to be compared. Comparisons of two sequences are usually carried out by comparing the sequences, after optimal alignment, with respect to a segment or "window of comparison", in order to identify local regions of corresponding sequences. The optimal alignment for a comparison may be carried out manually or with the aid of the local homology algorithm by Smith and Waterman, 1981, *Adv. App. Math.* 2, 482, with the aid of the local homology algorithm by Needleman and Wunsch, 1970, *J. Mol. Biol.* 48, 443, with the aid of the similarity search algorithm by Pearson and Lipman, 1988, *Proc. Natl Acad. Sci. USA* 88, 2444, or with the aid of computer programs using said algorithms (GAP, BESTFIT, FASTA, BLAST P, BLAST N and TFASTA in Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis.). In some embodiments, percent identity of two sequences is determined using the BLASTN or BLASTP algorithm, as available on the United States National Center for Biotechnology Information (NCBI) website (*e.g.*, at blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=align2seq). In some embodiments, the algorithm parameters used for BLASTN algorithm on the NCBI website include: (i) Expect Threshold set to 10; (ii) Word Size set to 28; (iii) Max matches in a query range set to 0; (iv) Match/Mismatch Scores set to 1, -2; (v) Gap Costs set to Linear; and (vi) the filter for low complexity regions being used. In some embodiments, the algorithm parameters used for BLASTP algorithm on the NCBI website include: (i) Expect Threshold set to 10; (ii) Word Size set to 3; (iii) Max matches in a query range set to 0; (iv) Matrix set to BLOSUM62; (v) Gap Costs set to Existence: 11 Extension: 1; and (vi) conditional compositional score matrix adjustment.

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

In some embodiments, the degree of similarity or identity is given for a region which is at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90% or

about 100% of the entire length of the reference sequence. For example, if the reference nucleic acid sequence consists of 200 nucleotides, the degree of identity is given for at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, or about 200 nucleotides, in some embodiments continuous nucleotides. In some embodiments, the degree of similarity or identity is given for the entire length of the reference sequence.

Homologous amino acid sequences exhibit according to the disclosure at least 40%, in particular at least 50%, at least 60%, at least 70%, at least 80%, at least 90% and, e.g., at least 95%, at least 98 or at least 99% identity of the amino acid residues.

The amino acid sequence variants described herein may readily be prepared by the skilled person, for example, by recombinant DNA manipulation. The manipulation of DNA sequences for preparing peptides or polypeptides having substitutions, additions, insertions or deletions, is described in detail in *Molecular Cloning: A Laboratory Manual*, 4th Edition, M.R. Green and J. Sambrook eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2012, for example. Furthermore, the peptides, polypeptides and amino acid variants described herein may be readily prepared with the aid of known peptide synthesis techniques such as, for example, by solid phase synthesis and similar methods.

In some embodiments, a fragment or variant of an amino acid sequence (peptide or polypeptide) is a "functional fragment" or "functional variant". The term "functional fragment" or "functional variant" of an amino acid sequence relates to any fragment or variant exhibiting one or more functional properties identical or similar to those of the amino acid sequence from which it is derived, *i.e.*, it is functionally equivalent. With respect to peptidoglycan hydrolases, one particular function is the activity to degrade peptidoglycan displayed by the amino acid sequence from which the fragment or variant is derived. The term "functional fragment" or "functional variant", as used herein, in particular refers to a variant molecule or sequence that comprises an amino acid sequence that is altered by one or more amino acids compared to the amino acid sequence of the parent molecule or sequence and that is still capable of fulfilling one or more of the functions of the parent molecule or sequence, *e.g.*, degrading peptidoglycan. In some embodiments, the modifications in the amino acid sequence of the parent molecule or sequence do not significantly affect or alter the characteristics of the molecule or sequence. In different embodiments, the function of the functional fragment or functional variant may be reduced but still significantly present, *e.g.*,

function of the functional fragment or functional variant may be at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the parent molecule or sequence. However, in other embodiments, function of the functional fragment or functional variant may be enhanced compared to the parent molecule or sequence.

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

15 In some embodiments, "isolated" means removed (e.g., purified) from the natural state or from an artificial composition, such as a composition from a production process. For example, a nucleic acid, peptide or polypeptide naturally present in a living animal is not "isolated", but the same nucleic acid, peptide or polypeptide partially or completely separated from the coexisting materials of its natural state is "isolated". An isolated nucleic acid, peptide or
20 polypeptide can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

The term "transfection" relates to the introduction of nucleic acids, in particular RNA, into a cell. For purposes of the present disclosure, the term "transfection" also includes the introduction of a nucleic acid into a cell or the uptake of a nucleic acid by such cell, wherein
25 the cell may be present in a subject, e.g., a patient, or the cell may be *in vitro*, e.g., outside of a patient. Thus, according to the present disclosure, a cell for transfection of a nucleic acid described herein can be present *in vitro* or *in vivo*, e.g. the cell can form part of an organ, a tissue and/or the body of a patient. According to the disclosure, transfection can be transient or stable. For some applications of transfection, it is sufficient if the transfected genetic
30 material is only transiently expressed. RNA can be transfected into cells to transiently express its coded protein. Since the nucleic acid introduced in the transfection process is usually not

integrated into the nuclear genome, the foreign nucleic acid will be diluted through mitosis or degraded. Cells allowing episomal amplification of nucleic acids greatly reduce the rate of dilution. If it is desired that the transfected nucleic acid actually remains in the genome of the cell and its daughter cells, a stable transfection must occur. Such stable transfection can be achieved by using virus-based systems or transposon-based systems for transfection, for example. Generally, nucleic acid encoding antigen is transiently transfected into cells. RNA can be transfected into cells to transiently express its coded protein.

The disclosure includes analogs of a peptide or polypeptide. According to the present disclosure, an analog of a peptide or polypeptide is a modified form of said peptide or polypeptide from which it has been derived and has at least one functional property of said peptide or polypeptide. *E.g.*, a pharmacological active analog of a peptide or polypeptide has at least one of the pharmacological activities of the peptide or polypeptide from which the analog has been derived. Such modifications include any chemical modification and comprise single or multiple substitutions, deletions and/or additions of any molecules associated with the peptide or polypeptide, such as carbohydrates, lipids and/or peptides or polypeptides. In some embodiments, "analogs" of peptides or polypeptides include those modified forms resulting from glycosylation, acetylation, phosphorylation, amidation, palmitoylation, myristoylation, isoprenylation, lipidation, alkylation, derivatization, introduction of protective/blocking groups, proteolytic cleavage or binding to an antibody or to another cellular ligand. The term "analog" also extends to all functional chemical equivalents of said peptides and polypeptides.

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

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

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

According to various embodiments of the present disclosure, a nucleic acid such as RNA encoding a peptide or polypeptide is taken up by or introduced, *i.e.* transfected or transduced, into a cell which cell may be present *in vitro* or in a subject, resulting in expression of said peptide or polypeptide. The cell may, *e.g.*, express the encoded peptide or polypeptide

intracellularly (*e.g.* in the cytoplasm and/or in the nucleus), may secrete the encoded peptide or polypeptide, and/or may express it on the surface. In some embodiments, the cell secretes the encoded peptide or polypeptide.

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

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

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

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

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

25 Prodrugs of a particular compound described herein are those compounds that upon administration to an individual undergo chemical conversion under physiological conditions to provide the particular compound. Additionally, prodrugs can be converted to the particular compound by chemical or biochemical methods in an *ex vivo* environment. For example, prodrugs can be slowly converted to the particular compound when, for example, placed in a transdermal patch reservoir with a suitable enzyme or chemical reagent. Exemplary prodrugs are esters (using an alcohol or a carboxy group contained in the particular compound) or

30

amides (using an amino or a carboxy group contained in the particular compound) which are hydrolyzable *in vivo*. Specifically, any amino group which is contained in the particular compound and which bears at least one hydrogen atom can be converted into a prodrug form. Typical N-prodrug forms include carbamates, Mannich bases, enamines, and enaminones.

5 In the present specification, a structural formula of a compound may represent a certain isomer of said compound. It is to be understood, however, that the present invention includes all isomers such as geometrical isomers, optical isomers based on an asymmetrical carbon, stereoisomers, tautomers and the like which occur structurally and isomer mixtures and is not limited to the description of the formula.

10 "Isomers" are compounds having the same molecular formula but differ in structure ("structural isomers") or in the geometrical (spatial) positioning of the functional groups and/or atoms ("stereoisomers"). "Enantiomers" are a pair of stereoisomers which are non-superimposable mirror-images of each other. A "racemic mixture" or "racemate" contains a pair of enantiomers in equal amounts and is denoted by the prefix ($\pm$). "Diastereomers" are
15 stereoisomers which are non-superimposable and which are not mirror-images of each other. "Tautomers" are structural isomers of the same chemical substance that spontaneously and reversibly interconvert into each other, even when pure, due to the migration of individual atoms or groups of atoms; *i.e.*, the tautomers are in a dynamic chemical equilibrium with each other. An example of tautomers are the isomers of the keto-enol-tautomerism. "Conformers"
20 are stereoisomers that can be interconverted just by rotations about formally single bonds, and include - in particular - those leading to different 3-dimensional forms of (hetero)cyclic rings, such as chair, half-chair, boat, and twist-boat forms of cyclohexane.

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

In some embodiments, the "polydispersity index" is may be calculated based on dynamic light
30 scattering measurements by the so-called cumulant analysis as mentioned in the definition of

the "average diameter". Under certain prerequisites, it can be taken as a measure of the size distribution of an ensemble of nanoparticles.

The "radius of gyration" (abbreviated herein as R_g) of a particle about an axis of rotation is the radial distance of a point from the axis of rotation at which, if the whole mass of the particle is assumed to be concentrated, its moment of inertia about the given axis would be the same as with its actual distribution of mass. Mathematically, R_g is the root mean square distance of the particle's components from either its center of mass or a given axis. For example, for a macromolecule composed of n mass elements, of masses m_i ($i = 1, 2, 3, \dots, n$), located at fixed distances s_i from the center of mass, R_g is the square-root of the mass average of s_i^2 over all mass elements and can be calculated as follows:

$$R_g = \left(\sum_{i=1}^n m_i \cdot s_i^2 / \sum_{i=1}^n m_i \right)^{1/2}$$

The radius of gyration can be determined or calculated experimentally, *e.g.*, by using light scattering. In particular, for small scattering vectors $\vec{q}$ the structure function S is defined as follows:

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

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

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

$$R_h = \frac{k_B \cdot T}{6 \cdot \pi \cdot \eta \cdot D}$$

wherein k_B is the Boltzmann constant; T is the temperature; η is the viscosity of the solvent; and D is the diffusion coefficient. The diffusion coefficient can be determined experimentally, *e.g.*, by using dynamic light scattering (DLS). Thus, one procedure to determine the

hydrodynamic radius of a particle or a population of particles (such as the hydrodynamic radius of particles contained in a sample or control composition as disclosed herein or the hydrodynamic radius of a particle peak obtained from subjecting such a sample or control composition to field-flow fractionation) is to measure the DLS signal of said particle or
 5 population of particles (such as DLS signal of particles contained in a sample or control composition as disclosed herein or the DLS signal of a particle peak obtained from subjecting such a sample or control composition to field-flow fractionation).

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

The term "UV" means ultraviolet and designates a band of the electromagnetic spectrum with a wavelength from 10 nm to 400 nm, *i.e.*, shorter than that of visible light but longer than X-rays.

The expression "multi-angle light scattering" or "MALS" as used herein relates to a technique
 15 for measuring the light scattered by a sample into a plurality of angles. "Multi-angle" means in this respect that scattered light can be detected at different discrete angles as measured, for example, by a single detector moved over a range including the specific angles selected or an array of detectors fixed at specific angular locations. In certain embodiments, the light source used in MALS is a laser source (MALLS: multi-angle laser light scattering). Based on the
 20 MALS signal of a composition comprising particles and by using an appropriate formalism (*e.g.*, Zimm plot, Berry plot, or Debye plot), it is possible to determine the radius of gyration (R_g) and, thus, the size of said particles. Preferably, the Zimm plot is a graphical presentation using the following equation:

$$\frac{R_\theta}{K^*c} = M_w P(\theta) - 2A_2 c M_w^2 P^2(\theta)$$

25 wherein c is the mass concentration of the particles in the solvent (g/mL); A_2 is the second virial coefficient (mol·mL/g²); $P(\theta)$ is a form factor relating to the dependence of scattered light intensity on angle; R_θ is the excess Rayleigh ratio (cm⁻¹); and K^* is an optical constant that is equal to $4\pi^2\eta_0 (dn/dc)^2\lambda_0^{-4}N_A^{-1}$, where η_0 is the refractive index of the solvent at the incident radiation (vacuum) wavelength, λ_0 is the incident radiation (vacuum) wavelength (nm), N_A is
 30 Avogadro's number (mol⁻¹), and dn/dc is the differential refractive index increment (mL/g) (*cf.*, *e.g.*, Buchholz et al. (Electrophoresis 22 (2001), 4118-4128); B.H. Zimm (J. Chem. Phys. 13

(1945), 141; P. Debye (J. Appl. Phys. 15 (1944): 338; and W. Burchard (Anal. Chem. 75 (2003), 4279-4291). Preferably, the Berry plot is calculated the following term:

$$\sqrt{\frac{R_{\theta}}{K^*c}}$$

wherein c , R_{θ} and K^* are as defined above. Preferably, the Debye plot is calculated the
 5 following term:

$$\frac{K^*c}{R_{\theta}}$$

wherein c , R_{θ} and K^* are as defined above.

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

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

Nucleic Acids

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
5 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
10 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
15 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
20 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
25 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, in some embodiments, modified by one or more alkyl groups, *e.g.*, one or more C₁₋₄ alkyl groups, *e.g.*, one or more methyl groups. Particular examples of modified purine or pyrimidine base moieties include
30 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.

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
5 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
10 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), trans-amplifying RNA (taRNA), single-stranded RNA (ssRNA), dsRNA,
15 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
20 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)).

RNA

25 According to the present disclosure, the term "RNA" includes "mRNA". According to the present disclosure, the term "mRNA" means "messenger-RNA" and includes a "transcript" which may be generated by using a DNA template. Generally, mRNA encodes a peptide or polypeptide.

mRNA is single-stranded but may contain self-complementary sequences that allow parts of
30 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.

- 5 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
10 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
15 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,
20 MEGAscript® T7 kit, MAXIsript®), 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
25 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, RNA is *in vitro* transcribed RNA (IVT-RNA) and may be obtained by *in vitro* transcription of an appropriate DNA template. The promoter for controlling transcription
30 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 RNA is "replicon RNA" or simply a "replicon", in particular "self-replicating RNA" or "self-amplifying RNA". In certain
5 embodiments, the replicon or self-replicating RNA 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
10 see José *et al.*, Future Microbiol., 2009, vol. 4, pp. 837–856). The total genome length of many alphaviruses typically ranges between 11,000 and 12,000 nucleotides, and the genomic RNA typically has a 5'-cap, and a 3' poly(A) tail. The genome of alphaviruses encodes non-structural proteins (involved in transcription, modification and replication of viral RNA and in protein modification) and structural proteins (forming the virus particle). There are typically two open
15 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
20 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
25 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
30 interest. Alphavirus-based trans-replication (trans-amplification) 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
5 certain alphaviral sequence elements to allow recognition and RNA synthesis by the alphaviral replicase.

In some embodiments of the present disclosure, the RNA 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
10 the RNA, it may be modified within the coding region, *i.e.*, the sequence encoding the expressed peptide or polypeptide, preferably without altering the sequence of the expressed peptide or polypeptide. Such modifications are described, for example, in WO 2007/036366 and PCT/EP2019/056502, and include the following: a 5'-cap structure; an extension or truncation of the naturally occurring poly(A) tail; an alteration of the 5'- and/or 3'-untranslated
15 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).

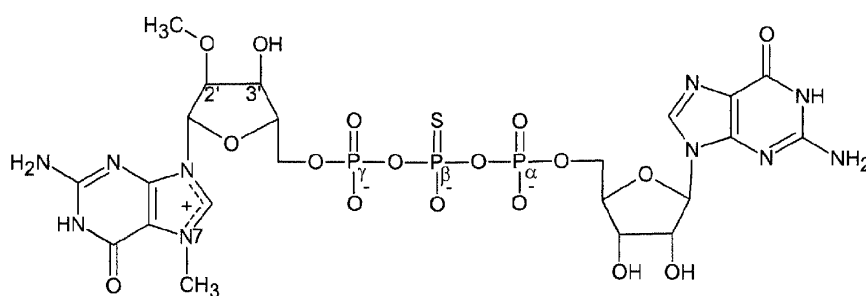
In some embodiments, the RNA comprises a 5'-cap structure. In some embodiments, the RNA
20 does not have uncapped 5'-triphosphates. In some embodiments, the RNA may comprise a conventional 5'-cap and/or a 5'-cap analog. The term "conventional 5'-cap" refers to a cap structure found on the 5'-end of an RNA molecule and generally comprises a guanosine 5'-triphosphate (Gppp) which is connected via its triphosphate moiety to the 5'-end of the next nucleotide of the RNA (*i.e.*, the guanosine is connected via a 5' to 5' triphosphate linkage to
25 the rest of the RNA). 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
30 preferred 5'-cap analogs are those having one or more substitutions at the bridging and non-bridging oxygen in the phosphate bridge, such as phosphorothioate modified 5'-cap analogs

at the β -phosphate (such as $m_2^{7,2'O}G(5')ppSp(5')G$ (referred to as beta-S-ARCA or β -S-ARCA)), as described in PCT/EP2019/056502. Providing an RNA 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 RNA strand, or the RNA may be generated, for example, by *in vitro* transcription, and the 5'-cap structure may be attached to the RNA post-transcriptionally using capping enzymes, for example, capping enzymes of vaccinia virus.

In some embodiments, the RNA comprises a 5'-cap structure selected from the group consisting of $m_2^{7,2'O}G(5')ppSp(5')G$ (in particular its D1 diastereomer), $m_2^{7,3'O}G(5')ppp(5')G$, and $m_2^{7,3'O}Gppp(m_1^{2'-O})ApG$.

In some embodiments, the RNA comprises a cap0, cap1, or cap2, preferably cap1 or cap2. According to the present disclosure, the term "cap0" means the structure " m^7GpppN ", wherein N is any nucleoside bearing an OH moiety at position 2'. According to the present disclosure, the term "cap1" means the structure " $m^7GpppNm$ ", wherein Nm is any nucleoside bearing an OCH_3 moiety at position 2'. According to the present disclosure, the term "cap2" means the structure " $m^7GpppNmNm$ ", wherein each Nm is independently any nucleoside bearing an OCH_3 moiety at position 2'.

The 5'-cap analog beta-S-ARCA (β -S-ARCA) has the following structure:



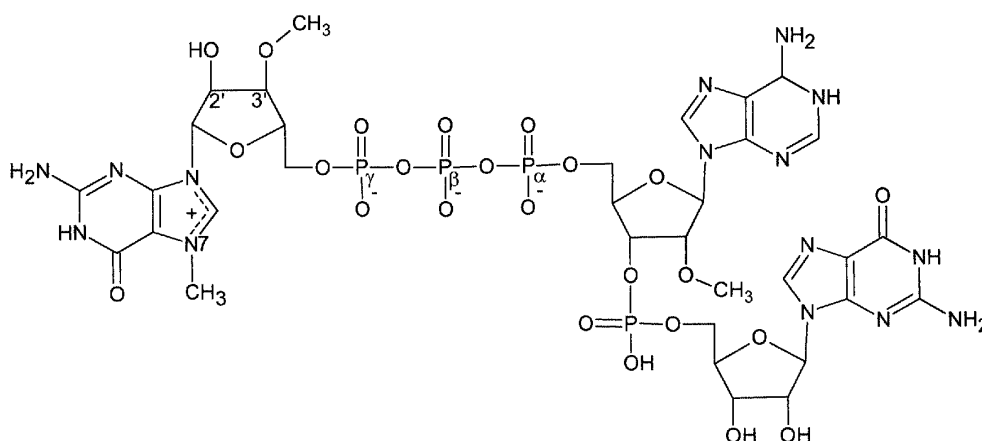
20

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 μ m, 4.6 x 250 mm is used for separation, whereby a flow rate of 1.3 ml/min can be

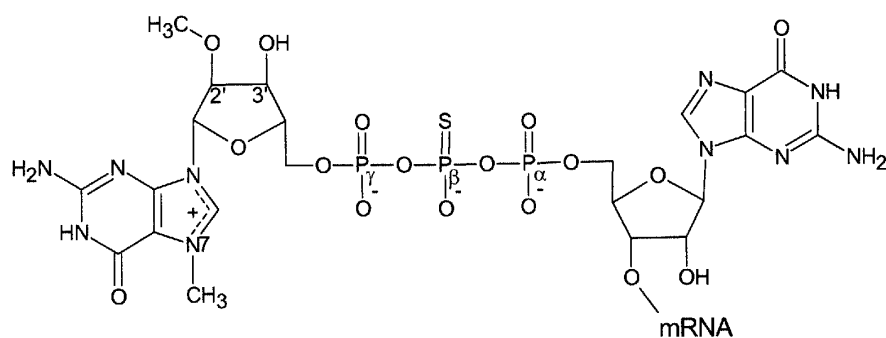
25

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.

5 The 5'-cap analog $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ (also referred to as $m_2^{7,3'-O}G(5')ppp(5')m^{2'-O}ApG$) which is a building block of a cap1 has the following structure:

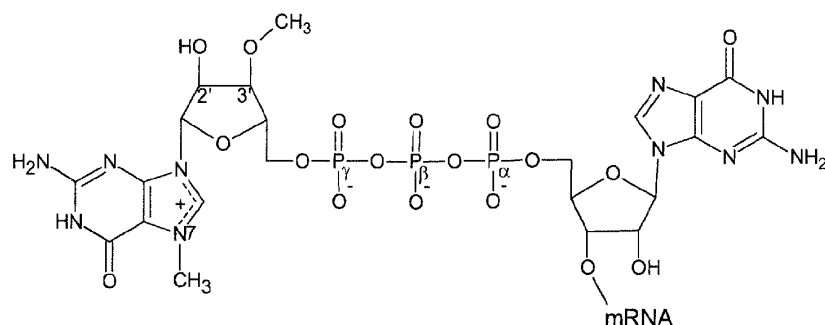


An exemplary cap0 mRNA comprising β -S-ARCA and mRNA has the following structure:

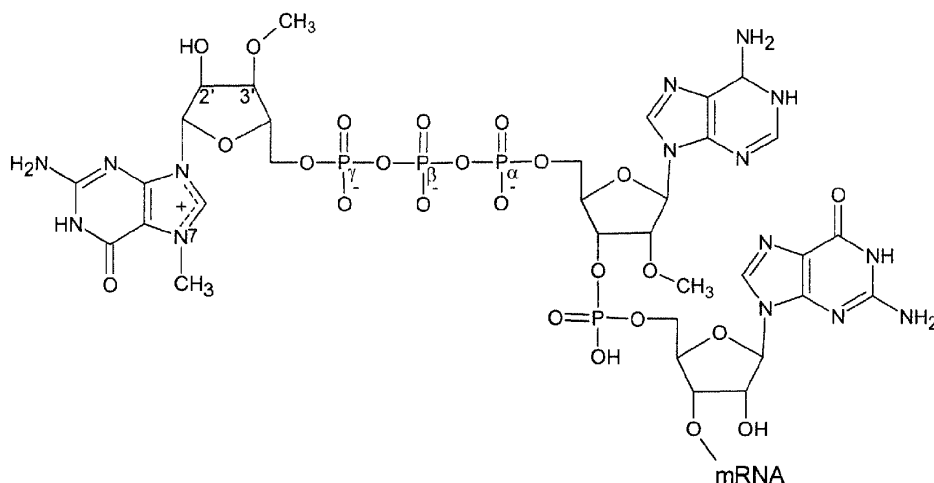


10

An exemplary cap0 mRNA comprising $m_2^{7,3^{10}}G(5')ppp(5')G$ and mRNA has the following structure:



An exemplary cap1 mRNA comprising $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ and mRNA has the following structure:



5

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 RNA 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 RNAs described herein. An uninterrupted poly-A tail is characterized by consecutive adenylate residues. In nature, an uninterrupted poly-A tail is typical. RNAs disclosed herein can have a poly-A tail attached to the free 3'-end of the RNA 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 RNA in transfected eukaryotic cells, as well as on the levels of protein

that is translated from an open reading frame that is present upstream (5') of the poly-A tail (Holtkamp *et al.*, 2006, Blood, vol. 108, pp. 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 (uridylylate), G nucleotides (guanylylate), or C nucleotides (cytidylylate). 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 (deoxythymidylylate) 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 RNA 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

5 some embodiments, the poly-A tail may essentially consist of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly-A tail may consist of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly-A tail comprises the poly-A tail shown in SEQ
10 ID NO: 13. 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, RNA 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
15 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
20 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
25 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,
30 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.

A particularly preferred 5'-UTR comprises the nucleotide sequence of SEQ ID NO: 11. A particularly preferred 3'-UTR comprises the nucleotide sequence of SEQ ID NO: 12.

- 5 In some embodiments, RNA comprises a 5'-UTR comprising the nucleotide sequence of SEQ ID NO: 11, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11.

- In some embodiments, RNA comprises a 3'-UTR comprising the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or
10 80% identity to the nucleotide sequence of SEQ ID NO: 12.

The RNA 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 RNA described herein is replaced (partially or completely, preferably completely) by a modified nucleoside. In some embodiments, the modified nucleoside is a modified uridine.

- 15 In some embodiments, the modified uridine replacing uridine is selected from the group consisting of pseudouridine (ψ), N1-methyl-pseudouridine (m1 ψ), 5-methyl-uridine (m5U), and combinations thereof.

- In some embodiments, the modified nucleoside replacing (partially or completely, preferably completely) uridine in the RNA may be any one or more of 3-methyl-uridine (m3U), 5-methoxy-uridine (mo5U), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s2U), 4-thio-uridine (s4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-oxyacetic acid (cmo5U), uridine 5-oxyacetic acid methyl ester (mcmo5U), 5-carboxymethyl-uridine (cm5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm5U),
25 5-carboxyhydroxymethyl-uridine methyl ester (mchm5U), 5-methoxycarbonylmethyl-uridine (mcm5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm5s2U), 5-aminomethyl-2-thio-uridine (nm5s2U), 5-methylaminomethyl-uridine (mnm5U), 1-ethyl-pseudouridine, 5-methylaminomethyl-2-thio-uridine (mnm5s2U), 5-methylaminomethyl-2-seleno-uridine (mnm5se2U), 5-carbamoylmethyl-uridine (ncm5U), 5-carboxymethylaminomethyl-uridine (cmnm5U), 5-carboxymethylaminomethyl-2-thio-uridine (cmnm5s2U), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (tm5U), 1-taurinomethyl-pseudouridine, 5-
- 30

taurinomethyl-2-thio-uridine (tm5s2U), 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),
 5 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-
 10 (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)-
 15 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
 20 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
 25 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 RNA 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
 30 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 RNA 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
5 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
10 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 RNA described herein. Codon-optimization is based on the finding that the translation efficiency is also determined by a different frequency in the occurrence of tRNAs in cells. Thus, the sequence of RNA may be modified such that codons for which
15 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 RNA described herein is increased compared to the G/C content of the corresponding coding sequence of the wild type RNA, wherein the amino acid sequence encoded by the RNA is preferably not modified compared to the amino acid sequence encoded by the wild type RNA.
20 This modification of the RNA sequence is based on the fact that the sequence of any RNA region to be translated is important for efficient translation of that RNA. 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
25 codons for the stability can be determined (so-called alternative codon usage). Depending on the amino acid to be encoded by the RNA, there are various possibilities for modification of the RNA sequence, compared to its wild type sequence. In particular, codons which contain A and/or U nucleotides can be modified by substituting these codons by other codons, which code for the same amino acids but contain no A and/or U or contain a lower content of A
30 and/or U nucleotides.

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

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 (m1 Ψ) 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 RNA 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 (m1 Ψ) or 5-methyluridine (m5U) for uridine), and (v) codon optimization.

Some aspects of the disclosure involve the targeted delivery of the RNA disclosed herein to certain cells or tissues.

Lipid-based RNA 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 RNA 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 RNA described herein, at least a portion of the RNA is delivered to a target cell or target organ. In some embodiments, at least a portion of the RNA is delivered to the cytosol of the target cell. In some embodiments, the RNA is RNA encoding a peptide or polypeptide and the RNA 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 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 cell in the skin. Thus, RNA particles described herein may be
5 used for delivering RNA to such target cell.

Peptidoglycan hydrolases/Endolysins and RNA encoding Peptidoglycan hydrolases/Endolysins

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

In some embodiments, RNA such as mRNA used in the present disclosure comprises a nucleic
20 acid sequence encoding a peptidoglycan hydrolase such as an endolysin.

In some embodiments, RNA such as mRNA used in the present disclosure comprises a nucleic
acid sequence encoding a peptidoglycan hydrolase such as an endolysin, and is capable of
expressing said peptidoglycan hydrolase such as endolysin, in particular if transferred into a
cell or subject, preferably a human cell or subject. Thus, in some embodiments, the RNA used
25 in the present disclosure contains a coding region (open reading frame (ORF)) encoding a
peptidoglycan hydrolase such as an endolysin. In this respect, an "open reading frame" or
"ORF" is a continuous stretch of codons beginning with a start codon and ending with a stop
codon. In some embodiments, RNA such as mRNA used in the present disclosure comprises a
nucleic acid sequence encoding more than one peptidoglycan hydrolase such as endolysin,
30 *e.g.*, two, three, four or more peptidoglycan hydrolases such as endolysins. In some
embodiments, the one or more peptidoglycan hydrolases such as endolysins encoded by the

RNA may comprise or consist of naturally occurring sequences, may comprise or consist of variants of naturally occurring sequences, or may comprise or consist of sequences which are not naturally occurring, e.g., recombinant sequences. In some embodiments, the peptide or polypeptide encoded by the RNA described herein may consist of the one or more
5 peptidoglycan hydrolases such as endolysins, or may comprise the one or more peptidoglycan hydrolases such as endolysins and may comprise additional sequences such as secretions signals, extended-PK groups, tags and any other sequences. In some embodiments, the additional sequences are fused to the one or more peptidoglycan hydrolases such as endolysins, in some embodiments, separated by a linker. In these embodiments, the one or
10 more peptidoglycan hydrolases such as endolysins may be considered the pharmaceutically active peptide or polypeptide even if additional sequences support the function or effect of the one or more peptidoglycan hydrolases such as endolysins.

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

Peptidoglycan hydrolases such as endolysins described herein can be prepared as fusion or chimeric polypeptides that include a peptidoglycan hydrolase portion such as endolysin portion and a heterologous polypeptide (i.e., a polypeptide that is not a peptidoglycan
30 hydrolase such as an endolysin). The peptidoglycan hydrolase such as endolysin may be fused to an extended-PK group, which increases circulation half-life. Non-limiting examples of

extended-PK groups are described *infra*. It should be understood that other PK groups that increase the circulation half-life of peptides or polypeptides such as peptidoglycan hydrolases such as endolysins, or variants thereof, are also applicable to the present disclosure. In certain embodiments, the extended-PK group is a serum albumin domain (e.g., mouse serum albumin, human serum albumin).

As used herein, the term "PK" is an acronym for "pharmacokinetic" and encompasses properties of a compound including, by way of example, absorption, distribution, metabolism, and elimination by a subject. As used herein, an "extended-PK group" refers to a protein, peptide, or moiety that increases the circulation half-life of a biologically active molecule when fused to or administered together with the biologically active molecule. Examples of an extended-PK group include serum albumin (e.g., HSA), Immunoglobulin Fc or Fc fragments and variants thereof, transferrin and variants thereof, and human serum albumin (HSA) binders (as disclosed in U.S. Publication Nos. 2005/0287153 and 2007/0003549). Other exemplary extended-PK groups are disclosed in Kontermann, *Expert Opin Biol Ther*, 2016 Jul;16(7):903-15 which is herein incorporated by reference in its entirety. As used herein, an "extended-PK" peptidoglycan hydrolase such as "extended-PK" endolysin refers to a peptidoglycan hydrolase moiety such as endolysin moiety in combination with an extended-PK group. In some embodiments, the extended-PK peptidoglycan hydrolase such as extended-PK endolysin is a fusion protein in which a peptidoglycan hydrolase moiety such as endolysin moiety is linked or fused to an extended-PK group.

In certain embodiments, the serum half-life of an extended-PK peptidoglycan hydrolase such as extended-PK endolysin is increased relative to the peptidoglycan hydrolase such as endolysin alone (i.e., the peptidoglycan hydrolase such as endolysin not fused to an extended-PK group). In certain embodiments, the serum half-life of the extended-PK peptidoglycan hydrolase such as extended-PK endolysin is at least 20, 40, 60, 80, 100, 120, 150, 180, 200, 400, 600, 800, or 1000% longer relative to the serum half-life of the peptidoglycan hydrolase such as endolysin alone. In certain embodiments, the serum half-life of the extended-PK peptidoglycan hydrolase such as extended-PK endolysin is at least 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5 fold, 4-fold, 4.5-fold, 5-fold, 6-fold, 7-fold, 8-fold, 10-fold, 12-fold, 13-fold, 15-fold, 17-fold, 20-fold, 22-fold, 25-fold, 27-fold, 30-fold, 35-fold, 40-fold, or 50-fold greater than the serum half-life of the peptidoglycan hydrolase such as endolysin alone. In certain

embodiments, the serum half-life of the extended-PK peptidoglycan hydrolase such as extended-PK endolysin is at least 10 hours, 15 hours, 20 hours, 25 hours, 30 hours, 35 hours, 40 hours, 50 hours, 60 hours, 70 hours, 80 hours, 90 hours, 100 hours, 110 hours, 120 hours, 130 hours, 135 hours, 140 hours, 150 hours, 160 hours, or 200 hours.

- 5 As used herein, "half-life" refers to the time taken for the serum or plasma concentration of a compound such as a peptide or polypeptide to reduce by 50%, *in vivo*, for example due to degradation and/or clearance or sequestration by natural mechanisms. An extended-PK peptidoglycan hydrolase such as extended-PK endolysin suitable for use herein is stabilized *in vivo* and its half-life increased by, e.g., fusion to serum albumin (e.g., HSA or MSA), which resist
- 10 degradation and/or clearance or sequestration. The half-life can be determined in any manner known per se, such as by pharmacokinetic analysis. Suitable techniques will be clear to the person skilled in the art, and may for example generally involve the steps of suitably administering a suitable dose of the amino acid sequence or compound to a subject; collecting blood samples or other samples from said subject at regular intervals; determining the level
- 15 or concentration of the amino acid sequence or compound in said blood sample; and calculating, from (a plot of) the data thus obtained, the time until the level or concentration of the amino acid sequence or compound has been reduced by 50% compared to the initial level upon dosing. Further details are provided in, e.g., standard handbooks, such as Kenneth, A. et al., Chemical Stability of Pharmaceuticals: A Handbook for Pharmacists and in Peters et
- 20 al., Pharmacokinetic Analysis: A Practical Approach (1996). Reference is also made to Gibaldi, M. et al., Pharmacokinetics, 2nd Rev. Edition, Marcel Dekker (1982).

In certain embodiments, the extended-PK group includes serum albumin, or fragments thereof or variants of the serum albumin or fragments thereof (all of which for the purpose of the present disclosure are comprised by the term "albumin"). Polypeptides described herein

25 may be fused to albumin (or a fragment or variant thereof) to form albumin fusion proteins. Such albumin fusion proteins are described in U.S. Publication No. 20070048282.

As used herein, "albumin fusion protein" refers to a protein formed by the fusion of at least one molecule of albumin (or a fragment or variant thereof) to at least one molecule of a protein such as a therapeutic protein, in particular a peptidoglycan hydrolase such as an

30 endolysin. The albumin fusion protein may be generated by translation of a nucleic acid in which a polynucleotide encoding a therapeutic protein is joined in-frame with a

polynucleotide encoding an albumin. The therapeutic protein and albumin, once part of the albumin fusion protein, may each be referred to as a "portion", "region" or "moiety" of the albumin fusion protein (e.g., a "therapeutic protein portion" or an "albumin protein portion"). In a highly preferred embodiment, an albumin fusion protein comprises at least one molecule of a therapeutic protein (including, but not limited to a mature form of the therapeutic protein) and at least one molecule of albumin (including but not limited to a mature form of albumin). In some embodiments, an albumin fusion protein is processed by a host cell such as a cell of the target organ for administered RNA, e.g. a liver cell, and secreted into the circulation. Processing of the nascent albumin fusion protein that occurs in the secretory pathways of the host cell used for expression of the RNA may include, but is not limited to signal peptide cleavage; formation of disulfide bonds; proper folding; addition and processing of carbohydrates (such as for example, N- and O-linked glycosylation); specific proteolytic cleavages; and/or assembly into multimeric proteins. An albumin fusion protein is preferably encoded by RNA in a non-processed form which in particular has a signal peptide at its N-terminus and following secretion by a cell is preferably present in the processed form wherein in particular the signal peptide has been cleaved off. In a most preferred embodiment, the "processed form of an albumin fusion protein" refers to an albumin fusion protein product which has undergone N-terminal signal peptide cleavage, herein also referred to as a "mature albumin fusion protein".

In preferred embodiments, albumin fusion proteins comprising a therapeutic protein have a higher plasma stability compared to the plasma stability of the same therapeutic protein when not fused to albumin. Plasma stability typically refers to the time period between when the therapeutic protein is administered *in vivo* and carried into the bloodstream and when the therapeutic protein is degraded and cleared from the bloodstream, into an organ, such as the kidney or liver, that ultimately clears the therapeutic protein from the body. Plasma stability is calculated in terms of the half-life of the therapeutic protein in the bloodstream. The half-life of the therapeutic protein in the bloodstream can be readily determined by common assays known in the art.

As used herein, "albumin" refers collectively to albumin protein or amino acid sequence, or an albumin fragment or variant, having one or more functional activities (e.g., biological activities) of albumin. In particular, "albumin" refers to human albumin or fragments or

variants thereof especially the mature form of human albumin, or albumin from other vertebrates or fragments thereof, or variants of these molecules. The albumin may be derived from any vertebrate, especially any mammal, for example human, cow, sheep, or pig. Non-mammalian albumins include, but are not limited to, hen and salmon. The albumin portion of
5 the albumin fusion protein may be from a different animal than the therapeutic protein portion.

In certain embodiments, the albumin is human serum albumin (HSA), or fragments or variants thereof, such as those disclosed in US 5,876,969, WO 2011/124718, WO 2013/075066, and WO 2011/0514789.

10 The terms, human serum albumin (HSA) and human albumin (HA) are used interchangeably herein. The terms, "albumin" and "serum albumin" are broader, and encompass human serum albumin (and fragments and variants thereof) as well as albumin from other species (and fragments and variants thereof).

As used herein, a fragment of albumin sufficient to prolong the therapeutic activity or plasma
15 stability of the therapeutic protein refers to a fragment of albumin sufficient in length or structure to stabilize or prolong the therapeutic activity or plasma stability of the protein so that the plasma stability of the therapeutic protein portion of the albumin fusion protein is prolonged or extended compared to the plasma stability in the non-fusion state.

The albumin portion of the albumin fusion proteins may comprise the full length of the
20 albumin sequence, or may include one or more fragments thereof that are capable of stabilizing or prolonging the therapeutic activity or plasma stability. Such fragments may be of 10 or more amino acids in length or may include about 15, 20, 25, 30, 50, or more contiguous amino acids from the albumin sequence or may include part or all of specific domains of albumin. For instance, one or more fragments of HSA spanning the first two immunoglobulin-
25 like domains may be used. In a preferred embodiment, the HSA fragment is the mature form of HSA.

Generally speaking, an albumin fragment or variant will be at least 100 amino acids long, preferably at least 150 amino acids long.

According to the disclosure, albumin may be naturally occurring albumin or a fragment or
30 variant thereof. Albumin may be human albumin and may be derived from any vertebrate, especially any mammal.

- Preferably, the albumin fusion protein comprises albumin as the N-terminal portion, and a therapeutic protein as the C-terminal portion. Alternatively, an albumin fusion protein comprising albumin as the C-terminal portion, and a therapeutic protein as the N-terminal portion may also be used. In other embodiments, the albumin fusion protein has a therapeutic protein fused to both the N-terminus and the C-terminus of albumin. In a preferred embodiment, the therapeutic proteins fused at the N- and C-termini are the same therapeutic proteins. In another preferred embodiment, the therapeutic proteins fused at the N- and C-termini are different therapeutic proteins. In some embodiments, the different therapeutic proteins are both peptidoglycan hydrolases such as endolysins.
- 10 In some embodiments, the therapeutic protein(s) is (are) joined to the albumin through (a) peptide linker(s). A linker peptide between the fused portions may provide greater physical separation between the moieties and thus maximize the accessibility of the therapeutic protein portion, for instance, for binding to its cognate receptor. The linker peptide may consist of amino acids such that it is flexible or more rigid. The linker sequence may be
- 15 cleavable by a protease or chemically.
- As used herein, the term "Fc region" refers to the portion of a native immunoglobulin formed by the respective Fc domains (or Fc moieties) of its two heavy chains. As used herein, the term "Fc domain" refers to a portion or fragment of a single immunoglobulin (Ig) heavy chain wherein the Fc domain does not comprise an Fv domain. In certain embodiments, an Fc domain begins in the hinge region just upstream of the papain cleavage site and ends at the
- 20 C-terminus of the antibody. Accordingly, a complete Fc domain comprises at least a hinge domain, a CH2 domain, and a CH3 domain. In certain embodiments, an Fc domain comprises at least one of: a hinge (e.g., upper, middle, and/or lower hinge region) domain, a CH2 domain, a CH3 domain, a CH4 domain, or a variant, portion, or fragment thereof. In certain
- 25 embodiments, an Fc domain comprises a complete Fc domain (i.e., a hinge domain, a CH2 domain, and a CH3 domain). In certain embodiments, an Fc domain comprises a hinge domain (or portion thereof) fused to a CH3 domain (or portion thereof). In certain embodiments, an Fc domain comprises a CH2 domain (or portion thereof) fused to a CH3 domain (or portion thereof). In certain embodiments, an Fc domain consists of a CH3 domain or portion thereof.
- 30 In certain embodiments, an Fc domain consists of a hinge domain (or portion thereof) and a CH3 domain (or portion thereof). In certain embodiments, an Fc domain consists of a CH2

domain (or portion thereof) and a CH3 domain. In certain embodiments, an Fc domain consists of a hinge domain (or portion thereof) and a CH2 domain (or portion thereof). In certain embodiments, an Fc domain lacks at least a portion of a CH2 domain (e.g., all or part of a CH2 domain). An Fc domain herein generally refers to a polypeptide comprising all or part of the

5 Fc domain of an immunoglobulin heavy-chain. This includes, but is not limited to, polypeptides comprising the entire CH1, hinge, CH2, and/or CH3 domains as well as fragments of such peptides comprising only, e.g., the hinge, CH2, and CH3 domain. The Fc domain may be derived from an immunoglobulin of any species and/or any subtype, including, but not limited to, a human IgG1, IgG2, IgG3, IgG4, IgD, IgA, IgE, or IgM antibody. The Fc domain encompasses

10 native Fc and Fc variant molecules. As set forth herein, it will be understood by one of ordinary skill in the art that any Fc domain may be modified such that it varies in amino acid sequence from the native Fc domain of a naturally occurring immunoglobulin molecule. In certain embodiments, the Fc domain has reduced effector function (e.g., FcγR binding).

The Fc domains of a polypeptide described herein may be derived from different

15 immunoglobulin molecules. For example, an Fc domain of a polypeptide may comprise a CH2 and/or CH3 domain derived from an IgG1 molecule and a hinge region derived from an IgG3 molecule. In another example, an Fc domain can comprise a chimeric hinge region derived, in part, from an IgG1 molecule and, in part, from an IgG3 molecule. In another example, an Fc domain can comprise a chimeric hinge derived, in part, from an IgG1 molecule and, in part,

20 from an IgG4 molecule.

In certain embodiments, an extended-PK group includes an Fc domain or fragments thereof or variants of the Fc domain or fragments thereof (all of which for the purpose of the present disclosure are comprised by the term "Fc domain"). The Fc domain does not contain a variable region that binds to antigen. Fc domains suitable for use in the present disclosure may be

25 obtained from a number of different sources. In certain embodiments, an Fc domain is derived from a human immunoglobulin. In certain embodiments, the Fc domain is from a human IgG1 constant region. It is understood, however, that the Fc domain may be derived from an immunoglobulin of another mammalian species, including for example, a rodent (e.g. a mouse, rat, rabbit, guinea pig) or non- human primate (e.g. chimpanzee, macaque) species.

Moreover, the Fc domain (or a fragment or variant thereof) may be derived from any immunoglobulin class, including IgM, IgG, IgD, IgA, and IgE, and any immunoglobulin isotype, including IgG1, IgG2, IgG3, and IgG4.

5 A variety of Fc domain gene sequences (e.g., mouse and human constant region gene sequences) are available in the form of publicly accessible deposits. Constant region domains comprising an Fc domain sequence can be selected lacking a particular effector function and/or with a particular modification to reduce immunogenicity. Many sequences of antibodies and antibody-encoding genes have been published and suitable Fc domain sequences (e.g. hinge, CH2, and/or CH3 sequences, or fragments or variants thereof) can be
10 derived from these sequences using art recognized techniques.

In certain embodiments, the extended-PK group is a serum albumin binding protein such as those described in US2005/0287153, US2007/0003549, US2007/0178082, US2007/0269422, US2010/0113339, WO2009/083804, and WO2009/133208, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is transferrin, as
15 disclosed in US 7,176,278 and US 8,158,579, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is a serum immunoglobulin binding protein such as those disclosed in US2007/0178082, US2014/0220017, and US2017/0145062, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is a fibronectin (Fn)-based scaffold domain protein that
20 binds to serum albumin, such as those disclosed in US2012/0094909, which is herein incorporated by reference in its entirety. Methods of making fibronectin-based scaffold domain proteins are also disclosed in US2012/0094909. A non-limiting example of a Fn3-based extended-PK group is Fn3(HSA), i.e., a Fn3 protein that binds to human serum albumin.

In certain aspects, the extended-PK peptidoglycan hydrolase such as extended-PK endolysin,
25 suitable for use according to the disclosure, can employ one or more peptide linkers. As used herein, the term "peptide linker" refers to a peptide or polypeptide sequence which connects two or more domains (e.g., the extended-PK moiety and a peptidoglycan hydrolase moiety such as an endolysin moiety) in a linear amino acid sequence of a polypeptide chain. For example, peptide linkers may be used to connect a peptidoglycan hydrolase moiety such as
30 an endolysin moiety to a HSA domain.

Linkers suitable for fusing the extended-PK group to e.g. a peptidoglycan hydrolase such as an endolysin are well known in the art. Exemplary linkers include glycine-serine-polypeptide linkers, glycine-proline-polypeptide linkers, and proline-alanine polypeptide linkers. In certain embodiments, the linker is a glycine-serine-polypeptide linker, i.e., a peptide that consists of
5 glycine and serine residues.

In some embodiments, the RNA encoding a peptidoglycan hydrolase such as an endolysin is a single-stranded, 5' capped RNA that is translated into the respective protein upon entering cells of a subject being administered the RNA, e.g., liver cells, or cells of a tissue affected by a bacterial infection. Preferably, the RNA contains structural elements optimized for maximal
10 efficacy of the RNA with respect to stability and translational efficiency (5' cap, 5' UTR, 3' UTR, poly(A) sequence).

In some embodiments, beta-S-ARCA(D1) is utilized as specific capping structure at the 5'-end of the RNA. In some embodiments, $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ is utilized as specific capping structure at the 5'-end of the RNA. In some embodiments, the 5'-UTR comprises the
15 nucleotide sequence of SEQ ID NO: 11, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11. In some embodiments, the 3'-UTR comprises the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12. In some embodiments, the poly(A) sequence is 110
20 nucleotides in length and consists of a stretch of 30 adenosine residues, followed by a 10 nucleotide linker sequence and another 70 adenosine residues. This poly(A) sequence was designed to enhance RNA stability and translational efficiency in dendritic cells. In some embodiments, the poly(A) sequence comprises the nucleotide sequence of SEQ ID NO: 13, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 13.
25

In some embodiments, RNA described herein has the structure: Cap-5' UTR-Peptidoglycan hydrolase-3' UTR-PolyA. In some embodiments, RNA described herein has the structure: Cap-5' UTR-SP-Peptidoglycan hydrolase-3' UTR-PolyA, wherein SP comprises a signal peptide encoding sequence.

30 In some embodiments, the RNA encoding a peptidoglycan hydrolase such as an endolysin is expressed in cells of the subject to provide the peptidoglycan hydrolase such as endolysin. In

some embodiments, the peptidoglycan hydrolase such as endolysin is secreted. In some embodiments, the RNA encoding the peptidoglycan hydrolase such as endolysin is transiently expressed in cells of the subject. In some embodiments, the RNA encoding the peptidoglycan hydrolase such as endolysin is administered systemically. In some embodiments, after systemic administration of the RNA encoding the peptidoglycan hydrolase such as endolysin, expression of the RNA encoding the peptidoglycan hydrolase such as endolysin in liver occurs. In some embodiments, after systemic administration of the RNA encoding the peptidoglycan hydrolase such as endolysin, expression of the RNA encoding the peptidoglycan hydrolase such as endolysin in cells of a tissue affected by a bacterial infection, e.g., lung cells occurs.

10 The term "peptidoglycan hydrolase", as used herein, refers to any polypeptide which is capable of hydrolyzing the peptidoglycan of bacteria, such as Gram-negative bacteria. The term is not restricted to a specific enzymatic cleavage mechanism. Three main classes of bacterial peptidoglycan hydrolases are glycosidases that cleave the backbone of glycan, the amidases that cleave the side-chain peptide and peptidases (endopeptidases and

15 carboxypeptidases) that cleave within the peptide side-chain. The term encompasses naturally occurring peptidoglycan hydrolases, such as peptidoglycan hydrolases of eukaryotic, prokaryotic or viral (in particular bacteriophage) origin. The term encompasses for example lysozymes, endolysins, tail-spike depolymerases, Virion-associated peptidoglycan hydrolases (VAPGH), bacteriocins (e.g. lysostaphin) and autolysins. The "peptidoglycan hydrolase" may

20 also be a synthetic or modified polypeptide (e.g., recombinant polypeptide) capable of hydrolyzing the peptidoglycan of bacteria. For example, enzymatically active shuffled endolysins in which domains of two or more endolysins have been swapped/exchanged qualify as "peptidoglycan hydrolase" just as truncated endolysins, in which only the enzymatic active domain remains. Peptidoglycan hydrolase activity can be measured by assays well

25 known in the art, e.g., by antibacterial assays, e.g., by assays which are e.g. described in Briers et al. (J. Biochem. Biophys Methods; 2007; 70: 531-533) or Donovan et al. (J. FEMS Microbiol Lett. 2006 December; 265(1).

Endolysins include bacteriophage-encoded enzymes, which act by hydrolyzing the host cell wall and subsequently allow the release of bacteriophage progenies. The lytic activity of

30 endolysins is classified into different types: (i) acetylmuramidases, (ii) transglycosylases, (iii) glucosaminidases, (iv) amidases, and (v) endopeptidases.

Generally, endolysins have a molecular weight ranging from 15 to 40 kDa. The majority of endolysins have a modular configuration. Modular endolysins are often characterized by the presence of one or two (multi-domain) N-terminal enzymatically active domains (EADs) linked by a short, flexible linker region to a C-terminal cell wall-binding domain (CBD). The enzymatically active domain (EAD) of modular endolysins functions to cleave various specific peptidoglycan bonds in the murein layer of the host bacterium, while the cell wall-binding domain (CBD) recognizes and binds to different epitopes in the cell wall for proper fixation of the catalytic effect of the EAD.

Generally, endolysins acting on Gram-negative bacteria have a simple globular configuration of the EAD without a CBD. Recent studies also show the occurrence of Gram-negative phage endolysins with globular configuration, having one or two CBDs at the N-terminus while the EAD module is at the C-terminus.

EADs may be categorized into three classes of six distinct enzymatic activities based on the mode of action and individual enzymatic specificities:

- (i) Glycosidases generally cleave the β -1,4 glycosidic bonds linking alternating polymeric structures of N-acetylmuramic acids (MurNAc) and N-acetylglucosamines (GlcNAc) in the peptidoglycan layer;
- (ii) Amidases catalyze the cleavage of amide bonds between the MurNAc and the first amino acid in the peptide stem moiety, L-alanine;
- (iii) Endopeptidases cleave bonds between two amino acids of the stem peptide. Bond cleavage can either occur within interpeptide bridge or stem peptide–interpeptide bridge.

As used herein, the term “endolysin” includes peptidoglycan hydrolases encoded by bacteriophages (or bacterial viruses) and includes any peptidoglycan hydrolase which is suitable to hydrolyse bacterial cell walls. Such peptidoglycan hydrolases may be naturally occurring enzymes of bacteriophages, modified enzymes of bacteriophages and non-naturally occurring enzymes, e.g., recombinant enzymes, such as enzymes having a non-naturally occurring arrangement of domains, which domains may be derived from naturally occurring enzymes. In some embodiments, endolysins are enzymes which have been engineered to alter their host range and efficiency. In some embodiments, endolysins comprise endolysins comprising naturally occurring domains and naturally occurring endolysins, e.g., endolysins comprising an N-terminal naturally occurring catalytic domain, and a C-terminal naturally

occurring cell wall binding domain, chimeric lysins characterized by shuffling of natural domains to create desirable properties, artilysins characterized by addition of an LPS disruptor to disrupt Gram-negative outer membrane, and truncated lysine, e.g., wherein the C-terminal cell wall binding domain has been removed to increase activity.

- 5 In some embodiments, an endolysin is derived from a bacteriophage (bacterial virus). In some embodiments, an endolysin may be a β -(1,4)-glycosylase (lysozyme), transglycosylase, amidase or endopeptidases. In some embodiments, "endolysins" comprise at least one "enzymatically active domain" (EAD) having at least one of the following activities: endopeptidase, N-acetyl-muramoyl-L-alanine-amidase (amidase), N-acetyl-muramidase, N-
10 acetyl-glucosaminidase (lysozyme) or transglycosylases. In some embodiments, endolysins may contain also regions which are enzymatically inactive, and bind to the cell wall of the host bacteria, the so-called CBDs (cell wall binding domains). Endolysins may contain one or more, e.g., two or more CBDs. However, the term "endolysin" as used herein also includes enzymes having at least one EAD but no CBDs. Generally, the cell wall binding domain is able to bind
15 different components on the surface of bacteria. Preferably, the cell wall binding domain is a peptidoglycan binding domain and binds to the bacteria's peptidoglycan structure. The different domains of an endolysin can be connected by a domain linker.

In gram-positive bacteria, the cytoplasmic membrane is surrounded by a peptidoglycan layer. The main purpose of the cell wall of Gram-positive bacteria is to maintain the shape of the
20 bacteria and counteract the pressure inside the bacterial cells. Peptidoglycan or murein is a polymer composed of sugar and amino acid. The sugar component is composed of N-acetylglucosamine residues and a N-acetylmuramic acid residues that are β -(1,4) linked. A peptide chain consisting of 3 to 5 amino acids is bound to N-acetylmuramic acid. Peptide chains can be cross-linked to peptide chains of other chains to form a 3D mesh-like layer. The
25 peptide chain can contain D- and L-amino acid residues, and its composition can vary depending on the type of bacteria.

In contrast to gram-positive bacteria, gram-negative bacteria have an outer membrane with a characteristic asymmetric bilayer. The outer membrane bilayer consists of an inner monolayer containing phospholipids (primarily phosphatidylethanolamine) and an outer monolayer
30 composed primarily of lipopolysaccharide (LPS).

The term "cell wall" as used herein refers to all components that form the outer cell enclosure of bacteria and thus guarantee their integrity. In particular, the term "cell wall" as used herein refers to peptidoglycan, the outer membrane of the Gram-negative bacteria with the lipopolysaccharide, the bacterial cell membrane, but also to additional layers deposited on

5 the peptidoglycan as e.g. capsules, outer protein layers or slimes.

The term "EAD" as used herein refers to the enzymatically active domain of an endolysin. The EAD is responsible for hydrolysing bacterial peptidoglycans. It exhibits at least one enzymatic activity of an endolysin. The EAD can also be composed of more than one enzymatically active module. The term "EAD" is used herein synonymously with the term "catalytic domain".

10 Most gram-negative bacteria and many gram-positive bacteria form bacterial biofilms. A biofilm is defined as an aggregate or aggregate of microorganisms attached to a surface. Adherent bacteria are often surrounded and protected by extracellular polymeric substances produced by Gram-negative and Gram-positive bacteria. Bacteria are more resistant to antibacterials such as antibiotics, disinfectants and cell wall degrading enzymes through

15 biofilms. Furthermore, currently, biofilm treatment is not feasible because extracellular polymeric substances protect themselves against degradation by antibacterials, disinfectants or biofilm degradants.

In some embodiments, the endolysin is or is derived from a Gram-negative phage endolysin. In some embodiments, the endolysin is or is derived from a *Pseudomonas* phage endolysin.

20 In some embodiments, the endolysin is or is derived from a *Pseudomonas aeruginosa* phage endolysin. In some embodiments, the endolysin is or is derived from endolysin LysPA26 having the following sequence:

MRTNNIDAIKEHEGLRLVAYLDSVGVWTIGYGDTPDVKGLTITKEEAEKRLRKRLVEFEG
YVNTYVKVPLKQHQFDALVSLVYNIGPTNFKTSTLLKKLNAGDYIGAADQFLVWNKGRVDGK
25 LVVIKGLANRRAKERKQFLGE
(SEQ ID NO: 1)

For purposes of the present disclosure, the above sequence is considered the wildtype LysPA26 amino acid sequence. Position numberings in LysPA26 protein given herein are in relation to the amino acid sequence according to SEQ ID NO: 1 and corresponding positions in

30 LysPA26 protein variants.

In some embodiments, the encoded amino acid sequence comprises, consists essentially of or consists of an endolysin such as LysPA26.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of amino acids 6 to 138 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 6 to 138 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 6 to 138 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 6 to 138 of SEQ ID NO: 1. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of amino acids 6 to 138 of SEQ ID NO: 1.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of amino acids 2 to 145 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 2 to 145 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 2 to 145 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 2 to 145 of SEQ ID NO: 1. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of amino acids 2 to 145 of SEQ ID NO: 1.

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

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 2, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2, or a fragment of the nucleotide sequence of SEQ ID NO: 2, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 2.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 2, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to

the nucleotide sequence of SEQ ID NO: 2, or a fragment of the nucleotide sequence of SEQ ID NO: 2, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 2; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 1.

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

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 3, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, or a fragment of the nucleotide sequence of SEQ ID NO: 3, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 3; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 1.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 6, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence of SEQ ID NO:

6, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 6.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 6, a
5 nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence of SEQ ID NO: 6, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having
10 at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 6; and/or (ii) encodes an amino acid sequence comprising
15 the amino acid sequence of SEQ ID NO: 1.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 8, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8, or a fragment of the nucleotide sequence of SEQ ID NO: 8, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80%
20 identity to the nucleotide sequence of SEQ ID NO: 8. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 8.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 8, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8, or a fragment of the nucleotide sequence of SEQ ID NO: 8, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having
25 at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1. In some embodiments, RNA (i) comprises the
30 the amino acid sequence of SEQ ID NO: 1. In some embodiments, RNA (i) comprises the

nucleotide sequence of SEQ ID NO: 8; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 1.

According to certain embodiments, a signal peptide (or signal sequence) is fused, either directly or through a linker, to a peptidoglycan hydrolase, e.g., an endolysin. Accordingly, in
 5 some embodiments, a signal peptide is fused to the above described amino acid sequences derived from LysPA26 endolysin or functional fragments thereof comprised by the encoded amino acid sequences described above.

In some embodiments, an open reading frame of the RNA described herein encodes a polypeptide that includes a signal sequence, e.g., that is functional in mammalian cells.

10 In some embodiments, a utilized signal sequence is "intrinsic" in that it is, in nature, associated with (e.g., linked to) the encoded polypeptide.

In some embodiments, a utilized signal sequence is heterologous to the encoded polypeptide – e.g., is not naturally part of a polypeptide (e.g., protein) whose sequences are included in the encoded polypeptide.

15 In some embodiments, signal peptides are sequences, which are typically characterized by a length of about 15 to 30 amino acids.

In many embodiments, signal peptides are positioned at the N-terminus of an encoded polypeptide as described herein, without being limited thereto. In some embodiments, signal peptides preferably allow the transport of the polypeptide encoded by RNAs of the present
 20 disclosure with which they are associated into a defined cellular compartment, preferably the cell surface, the endoplasmic reticulum (ER) or the endosomal-lysosomal compartment.

In some embodiments, an RNA sequence encodes a peptidoglycan hydrolase, e.g., an endolysin, that may comprise or otherwise be linked to a signal sequence (e.g., secretory sequence), such as those listed in Table 2, or a sequence having 1, 2, 3, 4, or 5 amino acid
 25 differences relative thereto. In some embodiments, a signal sequence such as MRVMAPRTLILLSGALALTETWAGS, or a sequence having 1, 2, 3, 4, or at the most 5 amino acid differences relative thereto is utilized. In some embodiments, a sequence such as MRVMAPRTLILLSGALALTETWAGS, or a sequence having 1, 2, 3, 4, or at the most 5 amino acid differences relative thereto, is utilized.

30 In some embodiments, a signal sequence is selected from those included in the Table 2 below:

Table 2: Signal sequences

Signal	Sequence (Amino Acid)
HSV-1 gD SP	MGGAAARLGAVILFVVIVGLHGVRISKY
HSV-2 gD SP	MGRLTSGVGTAALLVVAVGLRVVCA
HSV-2	MGRLTSGVGTAALLVVAVGLRVVCAKYA
SARS-CoV-2-S	MFVFLVLLPLVSSQCVNLT
human Ig heavy chain signal peptide	MDWIWRILFLVGAATGAHSQM
human Ig heavy chain signal peptide	MDWTWRVFCLLAVAPGAHS
HuIgGk signal peptide	METPAQLLFLLLWLPDTTG
IgE heavy chain epsilon-1 signal peptide	MDWTWILFLVAAATRVHS
Japanese encephalitis PRM signal sequence	MLGSNSGQRVVFTILLLVAPAYS
VSVg protein signal sequence	MKCLLYLAFLFIGVNCA
	MDWTWILFLVAAATRVHS
	ETPAQLLFLLLWLPDTTG
	MLGSNSGQRVVFTILLLVAPAYS
	MKCLLYLAFLFIGVNCA
	MWLVSLAIVTACAGA
	MFVFLVLLPLVSSQC

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a functional fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 4.

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

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 5, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, or a fragment of the nucleotide sequence of SEQ ID NO: 5, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a functional fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 5; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 4.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 7, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7, or a fragment of the nucleotide sequence of SEQ ID NO: 7, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80%

identity to the nucleotide sequence of SEQ ID NO: 7. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 7.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 7, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7, or a fragment of the nucleotide sequence of SEQ ID NO: 7, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a functional fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 7; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 4.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9, or a fragment of the nucleotide sequence of SEQ ID NO: 9, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 9.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9, or a fragment of the nucleotide sequence of SEQ ID NO: 9, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a functional fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, RNA (i) comprises the

nucleotide sequence of SEQ ID NO: 9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 4.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 14, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 14, or a functional fragment of the amino acid sequence of SEQ ID NO: 14, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 14. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 14.

10 In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 15 or 16, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 15 or 16, or a fragment of the nucleotide sequence of SEQ ID NO: 15 or 16, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 15 or 16. In some
15 embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 15 or 16.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 15 or 16, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 15 or 16, or a fragment of the nucleotide sequence of SEQ ID NO: 15 or 16, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%,
20 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 15 or 16; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 14, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 14, or a functional fragment of the amino acid sequence of SEQ ID NO: 14, or the amino acid sequence having at least 99%, 98%, 97%, 96%,
25 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 14. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 15 or 16; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 14.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 17, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 17, or a functional fragment of the amino acid sequence of SEQ ID NO: 17, or the amino acid sequence having at least 99%,

98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 17. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 17.

- In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 18 or 19, a
- 5 nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 18 or 19, or a fragment of the nucleotide sequence of SEQ ID NO: 18 or 19, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 18 or 19. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 18 or 19.
- 10 In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 18 or 19, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 18 or 19, or a fragment of the nucleotide sequence of SEQ ID NO: 18 or 19, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 18 or 19; and/or (ii)
- 15 encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 17, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 17, or a functional fragment of the amino acid sequence of SEQ ID NO: 17, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 17. In some
- 20 embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 18 or 19; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 17.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 20, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 20, or a functional fragment

25 of the amino acid sequence of SEQ ID NO: 20, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 20. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 20.

- In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 21 or 22, a
- 30 nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 21 or 22, or a fragment of the nucleotide sequence of

SEQ ID NO: 21 or 22, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 21 or 22. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 21 or 22.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 21 or 22, a
5 nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 21 or 22, or a fragment of the nucleotide sequence of SEQ ID NO: 21 or 22, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 21 or 22; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 20, an
10 amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 20, or a functional fragment of the amino acid sequence of SEQ ID NO: 20, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 20. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 21 or 22; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 20.
15

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 23, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 23, or a functional fragment of the amino acid sequence of SEQ ID NO: 23, or the amino acid sequence having at least 99%,
20 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 23. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 23.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 24 or 25, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 24 or 25, or a fragment of the nucleotide sequence of SEQ ID NO: 24 or 25, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 24 or 25. In some
25 embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 24 or 25.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 24 or 25, a
30 nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 24 or 25, or a fragment of the nucleotide sequence of

- SEQ ID NO: 24 or 25, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 24 or 25; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 23, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 23, or a functional fragment of the amino acid sequence of SEQ ID NO: 23, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 23. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 24 or 25; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 23.
- 10 In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 26, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 26, or a functional fragment of the amino acid sequence of SEQ ID NO: 26, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 26.
- 15 In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 26.
- In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 27 or 28, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 27 or 28, or a fragment of the nucleotide sequence of
- 20 SEQ ID NO: 27 or 28, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 27 or 28. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 27 or 28.
- In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 27 or 28, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 27 or 28, or a fragment of the nucleotide sequence of
- 25 SEQ ID NO: 27 or 28, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 27 or 28; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 26, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 26, or a functional fragment of the amino acid
- 30 sequence of SEQ ID NO: 26, or the amino acid sequence having at least 99%, 98%, 97%, 96%,

95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 26. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 27 or 28; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 26.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence
5 of SEQ ID NO: 29, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%,
85%, or 80% identity to the amino acid sequence of SEQ ID NO: 29, or a functional fragment
of the amino acid sequence of SEQ ID NO: 29, or the amino acid sequence having at least 99%,
98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 29.
In some embodiments, the encoded amino acid sequence comprises the amino acid sequence
10 of SEQ ID NO: 29.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 30 or 31, a
nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to
the nucleotide sequence of SEQ ID NO: 30 or 31, or a fragment of the nucleotide sequence of
SEQ ID NO: 30 or 31, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%,
15 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 30 or 31. In some
embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 30 or 31.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 30 or 31, a
nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to
the nucleotide sequence of SEQ ID NO: 30 or 31, or a fragment of the nucleotide sequence of
20 SEQ ID NO: 30 or 31, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%,
90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 30 or 31; and/or (ii)
encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 29, an
amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to
the amino acid sequence of SEQ ID NO: 29, or a functional fragment of the amino acid
25 sequence of SEQ ID NO: 29, or the amino acid sequence having at least 99%, 98%, 97%, 96%,
95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 29. In some
embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 30 or 31; and/or (ii)
encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 29.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence
30 of SEQ ID NO: 32, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%,
85%, or 80% identity to the amino acid sequence of SEQ ID NO: 32, or a functional fragment

of the amino acid sequence of SEQ ID NO: 32, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 32. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 32.

- 5 In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 33 or 34, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 33 or 34, or a fragment of the nucleotide sequence of SEQ ID NO: 33 or 34, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 33 or 34. In some
10 embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 33 or 34.

- In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 33 or 34, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 33 or 34, or a fragment of the nucleotide sequence of SEQ ID NO: 33 or 34, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%,
15 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 33 or 34; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 32, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 32, or a functional fragment of the amino acid sequence of SEQ ID NO: 32, or the amino acid sequence having at least 99%, 98%, 97%, 96%,
20 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 32. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 33 or 34; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 32.

- In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 35, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%,
25 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 35, or a functional fragment of the amino acid sequence of SEQ ID NO: 35, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 35. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 35.

- 30 In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 36 or 37, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to

the nucleotide sequence of SEQ ID NO: 36 or 37, or a fragment of the nucleotide sequence of SEQ ID NO: 36 or 37, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 36 or 37. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 36 or 37.

- 5 In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 36 or 37, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 36 or 37, or a fragment of the nucleotide sequence of SEQ ID NO: 36 or 37, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 36 or 37; and/or (ii)
- 10 encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 35, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 35, or a functional fragment of the amino acid sequence of SEQ ID NO: 35, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 35. In some
- 15 embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 36 or 37; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 35.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 38, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 38, or a functional fragment

20 of the amino acid sequence of SEQ ID NO: 38, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 38. In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 38.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 39 or 40, a

25 nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 39 or 40, or a fragment of the nucleotide sequence of SEQ ID NO: 39 or 40, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 39 or 40. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 39 or 40.

- 30 In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 39 or 40, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to

the nucleotide sequence of SEQ ID NO: 39 or 40, or a fragment of the nucleotide sequence of SEQ ID NO: 39 or 40, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 39 or 40; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 38, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 38, or a functional fragment of the amino acid sequence of SEQ ID NO: 38, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 38. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 39 or 40; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 38.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 41, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 41, or a functional fragment of the amino acid sequence of SEQ ID NO: 41, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 41.

In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 41.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 42 or 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 42 or 43, or a fragment of the nucleotide sequence of SEQ ID NO: 42 or 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 42 or 43. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 42 or 43.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 42 or 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 42 or 43, or a fragment of the nucleotide sequence of SEQ ID NO: 42 or 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 42 or 43; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 41, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 41, or a functional fragment of the amino acid

sequence of SEQ ID NO: 41, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 41. In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 42 or 43; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 41.

5 In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 44, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 44, or a functional fragment of the amino acid sequence of SEQ ID NO: 44, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 44.

10 In some embodiments, the encoded amino acid sequence comprises the amino acid sequence of SEQ ID NO: 44.

In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 45 or 46, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 45 or 46, or a fragment of the nucleotide sequence of

15 SEQ ID NO: 45 or 46, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 45 or 46. In some embodiments, RNA comprises the nucleotide sequence of SEQ ID NO: 45 or 46.

In some embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 45 or 46, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to

20 the nucleotide sequence of SEQ ID NO: 45 or 46, or a fragment of the nucleotide sequence of SEQ ID NO: 45 or 46, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 45 or 46; and/or (ii)

encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 44, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to

25 the amino acid sequence of SEQ ID NO: 44, or a functional fragment of the amino acid sequence of SEQ ID NO: 44, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 44. In some

embodiments, RNA (i) comprises the nucleotide sequence of SEQ ID NO: 45 or 46; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 44.

30

Embodiments of RNA described herein

In the following, embodiments of RNA for use herein are described, wherein certain terms used when describing elements thereof have the following meanings:

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

SP: Signal peptide.

Endolysin: Sequences encoding the respective endolysin or fragment.

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

A30L70: A poly(A)-tail measuring 110 nucleotides in length, consisting of a stretch of 30 adenosine residues, followed by a 10 nucleotide linker sequence and another 70 adenosine residues.

15 In some embodiments, RNA described herein has the structure:

Cap-hAg-Kozak-Endolysin-FI-A30L70

In some embodiments, RNA described herein has the structure:

Cap-hAg-Kozak-SP-Endolysin-FI-A30L70

In some embodiments, encoded amino acid sequences described herein have the structure:

20 SP-Endolysin

In some embodiments, hAg-Kozak comprises the nucleotide sequence of SEQ ID NO: 11. In some embodiments, SP comprises the amino acid sequence of SEQ ID NO: 10. In some embodiments, FI comprises the nucleotide sequence of SEQ ID NO: 12. In some embodiments, A30L70 comprises the nucleotide sequence of SEQ ID NO: 13.

25 In some embodiments, the RNA used in the present disclosure is non-immunogenic.

The term "non-immunogenic RNA" (such as "non-immunogenic mRNA") as used herein refers to RNA that does not induce a response by the immune system upon administration, *e.g.*, to a mammal, or induces a weaker response than would have been induced by the same RNA that differs only in that it has not been subjected to the modifications and treatments that
30 render the non-immunogenic RNA non-immunogenic, *i.e.*, than would have been induced by standard RNA (stdRNA). In certain embodiments, non-immunogenic RNA, which is also termed

modified RNA (modRNA) herein, is rendered non-immunogenic by incorporating modified nucleosides suppressing RNA-mediated activation of innate immune receptors into the RNA and/or limiting the amount of double-stranded RNA (dsRNA), e.g., by limiting the formation of double-stranded RNA (dsRNA), e.g., during *in vitro* transcription, and/or by removing
 5 double-stranded RNA (dsRNA), e.g., following *in vitro* transcription. In certain embodiments, non-immunogenic RNA is rendered non-immunogenic by incorporating modified nucleosides suppressing RNA-mediated activation of innate immune receptors into the RNA and/or by removing double-stranded RNA (dsRNA), e.g., following *in vitro* transcription.

For rendering the non-immunogenic RNA (especially mRNA) non-immunogenic by the
 10 incorporation of modified nucleosides, any modified nucleoside may be used as long as it lowers or suppresses immunogenicity of the RNA. Particularly preferred are modified nucleosides that suppress RNA-mediated activation of innate immune receptors. In some embodiments, the modified nucleosides comprise a replacement of one or more uridines with a nucleoside comprising a modified nucleobase. In some embodiments, the modified
 15 nucleobase is a modified uracil. In some embodiments, the nucleoside comprising a modified nucleobase is selected from the group consisting of 3-methyl-uridine (m^3U), 5-methoxy-uridine (mo^5U), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s^2U), 4-thio-uridine (s^4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho^5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-
 20 oxyacetic acid (cmo^5U), uridine 5-oxyacetic acid methyl ester ($mcmo^5U$), 5-carboxymethyl-uridine (cm^5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm^5U), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm^5U$), 5-methoxycarbonylmethyl-uridine (mcm^5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm^5s^2U), 5-aminomethyl-2-thio-uridine (nm^5s^2U), 5-methylaminomethyl-uridine (mnm^5U), 1-ethyl-pseudouridine, 5-
 25 methylaminomethyl-2-thio-uridine (mnm^5s^2U), 5-methylaminomethyl-2-seleno-uridine (mnm^5se^2U), 5-carbamoylmethyl-uridine (ncm^5U), 5-carboxymethylaminomethyl-uridine ($cmnm^5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s^2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-aurinomethyl-uridine (τm^5U), 1-aurinomethyl-pseudouridine, 5-aurinomethyl-2-thio-uridine (τm^5s^2U), 1-aurinomethyl-4-thio-pseudouridine, 5-methyl-2-
 30 thio-uridine (m^5s^2U), 1-methyl-4-thio-pseudouridine ($m^1s^4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-

pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m^5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp^3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine ($acp^3\psi$), 5-(isopentenylaminomethyl)uridine (inm^5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm^5s^2U), α -thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m^5Um), 2'-O-methyl-pseudouridine (ψm), 2-thio-2'-O-methyl-uridine (s^2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm^5Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm^5Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine ($cmnm^5Um$), 3,2'-O-dimethyl-uridine (m^3Um), 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm^5Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, and 5-[3-(1-E-propenylamino)uridine. In certain embodiments, the nucleoside comprising a modified nucleobase is pseudouridine (ψ), N1-methyl-pseudouridine ($m1\psi$) or 5-methyl-uridine ($m5U$), in particular N1-methyl-pseudouridine.

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

During synthesis of RNA by *in vitro* transcription (IVT) using T7 RNA polymerase significant amounts of aberrant products, including double-stranded RNA (dsRNA) are produced due to unconventional activity of the enzyme. dsRNA induces inflammatory cytokines and activates effector enzymes leading to protein synthesis inhibition. Formation of dsRNA can be limited during synthesis of RNA by *in vitro* transcription (IVT), for example, by limiting the amount of uridine triphosphate (UTP) during synthesis. Optionally, UTP may be added once or several times during synthesis of RNA. Also, dsRNA can be removed from RNA such as IVT RNA, for example, by ion-pair reversed phase HPLC using a non-porous or porous C-18 polystyrene-divinylbenzene (PS-DVB) matrix. Alternatively, an enzymatic based method using *E. coli* RNaseIII that specifically hydrolyzes dsRNA but not ssRNA, thereby eliminating dsRNA contaminants from IVT RNA preparations can be used. Furthermore, dsRNA can be separated

from ssRNA by using a cellulose material. In some embodiments, an RNA preparation is contacted with a cellulose material and the ssRNA is separated from the cellulose material under conditions which allow binding of dsRNA to the cellulose material and do not allow binding of ssRNA to the cellulose material. Suitable methods for providing ssRNA are disclosed, for example, in WO 2017/182524.

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

In some embodiments, the amount of double-stranded RNA (dsRNA) is limited, e.g., dsRNA (especially mRNA) is removed from non-immunogenic RNA, such that less than 10%, less than 5%, less than 4%, less than 3%, less than 2%, less than 1%, less than 0.5%, less than 0.3%, less than 0.1%, less than 0.05%, less than 0.03%, less than 0.01%, less than 0.005%, less than 0.004%, less than 0.003%, less than 0.002%, less than 0.001%, or less than 0.0005% of the RNA in the non-immunogenic RNA composition is dsRNA. In some embodiments, the non-immunogenic RNA (especially mRNA) is free or essentially free of dsRNA. In some embodiments, the non-immunogenic RNA (especially mRNA) composition comprises a purified preparation of single-stranded nucleoside modified RNA. In some embodiments, the non-immunogenic RNA (especially mRNA) composition comprises single-stranded nucleoside modified RNA (especially mRNA) and is substantially free of double stranded RNA (dsRNA). In some embodiments, the non-immunogenic RNA (especially mRNA) composition comprises at least 90%, at least 91%, at least 92%, at least 93 %, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.9%, at least 99.99%, at least 99.991%, at least 99.992%, at least 99.993%, at least 99.994%, at least 99.995%, at least 99.996%, at least 99.997%, or at least 99.998% single stranded nucleoside modified RNA, relative to all other nucleic acid molecules (DNA, dsRNA, etc.).

Various methods can be used to determine the amount of dsRNA. For example, a sample may be contacted with dsRNA-specific antibody and the amount of antibody binding to RNA may be taken as a measure for the amount of dsRNA in the sample. A sample containing a known amount of dsRNA may be used as a reference.

- 5 For example, RNA may be spotted onto a membrane, e.g., nylon blotting membrane. The membrane may be blocked, e.g., in TBS-T buffer (20 mM TRIS pH 7.4, 137 mM NaCl, 0.1% (v/v) TWEEEN-20) containing 5% (w/v) skim milk powder. For detection of dsRNA, the membrane may be incubated with dsRNA-specific antibody, e.g., dsRNA-specific mouse mAb (English & Scientific Consulting, Szirák, Hungary). After washing, e.g., with TBS-T, the membrane may be
- 10 incubated with a secondary antibody, e.g., HRP-conjugated donkey anti-mouse IgG (Jackson ImmunoResearch, Cat #715-035-150), and the signal provided by the secondary antibody may be detected.

- In some embodiments, the non-immunogenic RNA (especially mRNA) is translated in a cell more efficiently than standard RNA with the same sequence. In some embodiments,
- 15 translation is enhanced by a factor of 2-fold relative to its unmodified counterpart. In some embodiments, translation is enhanced by a 3-fold factor. In some embodiments, translation is enhanced by a 4-fold factor. In some embodiments, translation is enhanced by a 5-fold factor. In some embodiments, translation is enhanced by a 6-fold factor. In some embodiments, translation is enhanced by a 7-fold factor. In some embodiments, translation is enhanced by
- 20 an 8-fold factor. In some embodiments, translation is enhanced by a 9-fold factor. In some embodiments, translation is enhanced by a 10-fold factor. In some embodiments, translation is enhanced by a 15-fold factor. In some embodiments, translation is enhanced by a 20-fold factor. In some embodiments, translation is enhanced by a 50-fold factor. In some embodiments, translation is enhanced by a 100-fold factor. In some embodiments, translation
- 25 is enhanced by a 200-fold factor. In some embodiments, translation is enhanced by a 500-fold factor. In some embodiments, translation is enhanced by a 1000-fold factor. In some embodiments, translation is enhanced by a 2000-fold factor. In some embodiments, the factor is 10-1000-fold. In some embodiments, the factor is 10-100-fold. In some embodiments, the factor is 10-200-fold. In some embodiments, the factor is 10-300-fold. In some embodiments,
- 30 the factor is 10-500-fold. In some embodiments, the factor is 20-1000-fold. In some embodiments, the factor is 30-1000-fold. In some embodiments, the factor is 50-1000-fold. In

some embodiments, the factor is 100-1000-fold. In some embodiments, the factor is 200-1000-fold. In some embodiments, translation is enhanced by any other significant amount or range of amounts.

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

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

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

5 In some embodiments, an RNA described herein is a single-stranded RNA. In some embodiments, an RNA described herein is a linear RNA. In some embodiments, a single-stranded RNA is a coding RNA in that its nucleotide sequence includes an open reading frame. In some embodiments, a single-stranded RNA has a nucleotide sequence that encodes a polypeptide or a plurality of polypeptides of the present disclosure.

10 In many embodiments, a relevant RNA is an mRNA.

In some embodiments, an RNA includes unmodified uridine residues; an RNA that includes only unmodified uridine residues may be referred to as a "uRNA". In some embodiments, an RNA includes one or more modified uridine residues; in some embodiments, such an RNA (e.g., an RNA including entirely modified uridine residues) is referred to as a "modRNA".

15 In some embodiments, an RNA may be a self-amplifying RNA (saRNA). In some embodiments, an RNA may be a trans-amplifying RNA (see, for example, WO2017/162461).

In some embodiments, RNA (e.g., a single stranded RNA) described herein has a length of at least 500 ribonucleotides (such as, e.g., at least 600 ribonucleotides, at least 700 ribonucleotides, at least 800 ribonucleotides, at least 900 ribonucleotides, at least 1000 ribonucleotides, at least 1250 ribonucleotides, at least 1500 ribonucleotides, at least 1750 ribonucleotides, at least 2000 ribonucleotides, at least 2500 ribonucleotides, at least 3000 ribonucleotides, at least 3500 ribonucleotides, at least 4000 ribonucleotides, at least 4500 ribonucleotides, at least 5000 ribonucleotides, or longer). In some embodiments, RNA described herein is single-stranded RNA having a length of about 800 ribonucleotides to 5000 ribonucleotides.

25 In some embodiments, a relevant RNA includes a polypeptide-encoding portion or a plurality of polypeptide-encoding portions. In some particular embodiments, such a portion or portions encode at least polypeptide or polypeptides that is or comprises a peptidoglycan hydrolase, e.g., endolysin. In certain embodiments, an encoded peptidoglycan hydrolase, e.g., endolysin, may be a variant of a wild type polypeptide.

30

In some embodiments, a single-stranded RNA (*e.g.*, mRNA) may comprise a secretion signal-encoding region (*e.g.*, a secretion signal-encoding region that allows an encoded target entity or entities (comprising a peptidoglycan hydrolase, *e.g.*, endolysin) to be secreted upon translation by cells). In some embodiments, such a secretion signal-encoding region may be or comprise a non-human secretion signal. In some embodiments, such a secretion signal-encoding region may be or comprise a human secretion signal.

In some embodiments, a single-stranded RNA (*e.g.*, mRNA) may comprise at least one non-coding sequence element (*e.g.*, to enhance RNA stability and/or translation efficiency). Examples of non-coding sequence elements include but are not limited to a 3' untranslated region (UTR), a 5' UTR, a cap structure, a poly adenine (polyA) tail, and any combinations thereof.

At least four formats useful for RNA pharmaceutical compositions may be used herein, namely non-modified uridine containing mRNA (uRNA), nucleoside modified mRNA (modRNA), self-amplifying mRNA (saRNA), and trans-amplifying RNAs.

Features of modified uridine (*e.g.*, pseudouridine) platform may include reduced adjuvant effect, blunted immune innate immune sensor activating capacity and thus augmented polypeptide (*e.g.*, protein) expression.

Features of self-amplifying platform may include, for example, long duration of polypeptide (*e.g.*, protein) expression, good tolerability and safety, higher likelihood for efficacy with very low RNA dose.

In some embodiments, a self-amplifying platform (*e.g.*, RNA) comprises two nucleic acid molecules, wherein one nucleic acid molecule encodes a replicase (*e.g.*, a viral replicase) and the other nucleic acid molecule is capable of being replicated (*e.g.*, a replicon) by said replicase *in trans* (*trans*-replication system). In some embodiments, a self-amplifying platform (*e.g.*, RNA) comprises a plurality of nucleic acid molecules, wherein said nucleic acids encode a plurality of replicases and/or replicons.

In some embodiments, a *trans*-replication system comprises the presence of both nucleic acid molecules in a single host cell.

In some such embodiments, a nucleic acid encoding a replicase (*e.g.*, a viral replicase) is not capable of self-replication in a target cell and/or target organism. In some such embodiments, a nucleic acid encoding a replicase (*e.g.*, a viral replicase) lacks at least one conserved

sequence element important for (-) strand synthesis based on a (+) strand template and/or for (+) strand synthesis based on a (-) strand template.

In some embodiments, a self-amplifying RNA comprises a 3' untranslated region (UTR), a 5' UTR, a cap structure, a poly adenine (polyA) tail, and any combinations thereof.

- 5 In some embodiments, a self-amplifying platform does not require propagation of virus particles (*e.g.*, is not associated with undesired virus-particle formation). In some embodiments, a self-amplifying platform is not capable of forming virus particles.

Particles

- 10 RNA, in particular mRNA, described herein 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 RNA are involved in particle formation. This results in complexation and spontaneous formation of RNA particles.
- 15 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.
- 20 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
- 25 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
- 30 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.

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 RNA 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 RNA 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 RNAs. In some embodiments, formed lipoplexes possess distinct internal arrangements of molecules that arise due to the transformation from liposomal structure into compact RNA-lipoplexes. In some embodiments, these formulations are characterized by their poor encapsulation of the RNA and incomplete entrapment of the RNA.

- In some embodiments, an LPX particle comprises an amphiphilic lipid, in particular cationic or cationically ionizable amphiphilic lipid, and RNA (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 RNA (especially mRNA) results in complexation and spontaneous formation of RNA 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, an RNA (especially mRNA) lipoplex particle is a nanoparticle.
- 10 In general, a lipid nanoparticle (LNP) is obtainable from direct mixing of RNA 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 RNA LNPs described herein the RNA 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.
- 15 In some embodiments, RNA (*e.g.*, mRNA) may be noncovalently associated with a particle as described herein. In embodiments, the RNA (especially mRNA) may be adhered to the outer surface of the particle (surface RNA (especially surface mRNA)) and/or may be contained in the particle (encapsulated RNA (especially encapsulated mRNA)).
- 25 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
- 30

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.

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

RNA particles (especially 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 RNA to obtain RNA 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
5 milky gel that turns subsequently into a liposomal suspension.

The term "ethanol injection technique" refers to a process, in which an ethanol solution comprising lipids is rapidly injected into an aqueous solution through a needle. This action disperses the lipids throughout the solution and promotes lipid structure formation, for example lipid vesicle formation such as liposome formation. Generally, the RNA (especially
10 mRNA) lipoplex particles described herein are obtainable by adding RNA (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 RNA
15 (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.

20 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 RNA in
25 an aqueous buffer. While RNA particles described herein may comprise polymer conjugated lipids such as PEG lipids, provided herein are also RNA particles which do not comprise polymer conjugated lipids such as PEG lipids.

In some embodiments, the LNPs comprising RNA and at least one cationic or cationically ionizable lipid described herein are prepared by (a) preparing an RNA solution containing
30 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 RNA

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 RNA 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 an RNA solution containing water and a buffering system; and (c') mixing the liposomes or colloidal preparation prepared under (a') with the RNA solution prepared under (b'). After step (c') one or more steps selected from
5
10 diluting and filtrating, such as tangential flow filtrating, can follow.

The present disclosure describes particles comprising RNA (especially mRNA) and at least one cationic or cationically ionizable lipid which associates with the RNA to form RNA particles and compositions comprising such particles. The RNA particles may comprise RNA which is complexed in different forms by non-covalent interactions to the particle. The particles
15 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 RNA 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
20 associate with RNA to form RNA particles. Such components include any component which can be part of RNA particles.

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

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

contain exclusively the specific RNA 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.

- 5 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 RNA-containing particles is obtainable. Individual particulate populations may be together in one container, comprising a mixed population of individual particulate formulations.
- 10 Alternatively, it is possible that all RNA 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 RNA species together with a particle-forming agent, thereby allowing the formation of particles. As opposed to a mixed particulate formulation, a combined particulate formulation will typically comprise
- 15 particles which comprise more than one RNA species. In a combined particulate composition different RNA species are typically present together in a single particle.

Polymers

- Given their high degree of chemical flexibility, polymers are commonly used materials for
- 20 nanoparticle-based delivery. Typically, cationic polymers are used to electrostatically condense the negatively charged RNA 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
- 25 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.
- 30 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.

- 5 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
10 (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
15 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
20 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
25 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
30 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
5 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
10 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,
15 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,
20 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
25 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
30 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 includes 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 ellipsoid 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.

5 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
10 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
15 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
20 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.

25

Cationic/Cationically ionizable lipids

The RNA 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
30 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.

5 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
10 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
15 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
20 protonated, e.g., under physiological conditions.

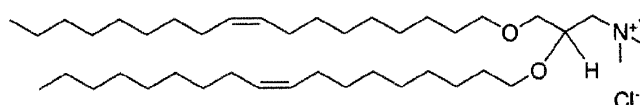
Examples of cationic or cationically ionizable lipids include, but are not limited to N,N-dimethyl-2,3-dioleoyloxypropylamine (DODMA), 1,2-dioleoyl-3-trimethylammonium propane (DOTAP); 1,2-di-O-octadecenyl-3-trimethylammonium propane (DOTMA), 3-(N-(N',N'-dimethylaminoethane)-carbamoyl)cholesterol (DC-Chol), dimethyldioctadecylammonium
25 (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-distearoyloxy-N,N-dimethyl-3-aminopropane (DSDMA), 2,3-di(tetradecoxy)propyl-(2-hydroxyethyl)-dimethylazanium (DMRIE), 1,2-dimyristoyl-sn-glycero-3-ethylphosphocholine (DMEPC), 1,2-dimyristoyl-3-trimethylammonium propane (DMTAP), 1,2-dioleoyloxypropyl-3-dimethyl-hydroxyethyl
30 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-dilinenyloxy-N,N-dimethylaminopropane (DLenDMA), dioctadecylamidoglycyl spermine (DOGS), 3-dimethylamino-2-(cholest-5-en-3-beta-oxybutan-4-oxy)-1-(cis,cis-9,12-octadecadienoxy)propane (CLinDMA), 2-[5'-(cholest-5-en-3-beta-oxy)-3'-oxapentoxy]-3-dimethyl-1-(cis,cis-9',12'-octadecadienoxy)propane (CpLinDMA), N,N-dimethyl-3,4-dioleyloxybenzylamine (DMOBA), 1,2-N,N'-dioleylearbamyl-3-dimethylaminopropane (DOcarbDAP), 2,3-Dilinoleoyloxy-N,N-dimethylpropylamine (DLinDAP), 1,2-N,N'-Dilinoleylearbamyl-3-dimethylaminopropane (DLincarbDAP), 1,2-Dilinoleoylcarbaryl-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), ($\pm$)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis(cis-9-tetradecenyoxy)-1-propanaminium bromide (GAP-DMORIE), ($\pm$)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis(dodecyloxy)-1-propanaminium bromide (GAP-DLRIE), ($\pm$)-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 (β AE-DMRIE), N-(4-carboxybenzyl)-N,N-dimethyl-2,3-bis(oleoyloxy)propan-1-aminiun (DOBAQ), 2-{[8-[(3 β)-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-propene (DMDAP), 1,2-dipalmitoyl-3-dimethylammonium-propene (DPDAP), N1-[2-[(1S)-1-[(3-aminopropyl)amino]-4-[di(3-amino-propyl)amino]butylcarboxamido]ethyl]-3,4-di[oyleyloxy]-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-aminiun 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-dodecylcarbamoylethyl)-{2-[(2-dodecylcarbamoylethyl)-2-[(2-dodecylcarbamoylethyl)-{2-[(2-

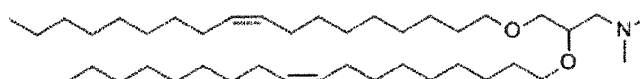
dodecylcarbamoyl-ethylamino)-ethyl]-amino)-ethylamino)propionamide (lipidoid 98N₁₂-5), 1-[2-[bis(2-hydroxydodecyl)amino]ethyl]-[2-[4-[2-[bis(2 hydroxydodecyl)amino]ethyl]piperazin-1-yl]ethyl]amino]dodecan-2-ol (lipidoid C12-200).

In some embodiments, the cationic or cationically ionizable lipid is DOTMA. In some
5 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:



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



15 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.

20 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
25 as non-cationic lipids. Optimizing the formulation of RNA 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 RNA delivery.

One or more additional lipids may or may not affect the overall charge of the RNA 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.

5 In some embodiments, the RNA 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 RNA particle characteristics, such as charge, particle size, stability, tissue selectivity, and bioactivity
10 of the RNA. 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 RNA particles (especially in the particles comprising mRNA) described herein comprise one or more of the
15 following: neutral lipids, steroids, and combinations thereof.

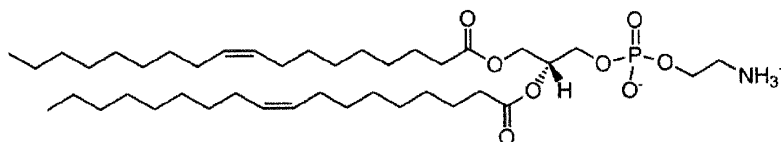
In some embodiments, the one or more additional lipids comprise a neutral lipid which is a phospholipid. In some embodiments, the phospholipid is selected from the group consisting of phosphatidylcholines, phosphatidylethanolamines, phosphatidylglycerols, phosphatidic acids, phosphatidylserines and sphingomyelins. Specific phospholipids that can be used
20 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),
25 diarachidoylphosphatidylcholine (DAPC), dibehenoylphosphatidylcholine (DBPC), ditricosanoylphosphatidylcholine (DTPC), dilignoceroylphatidylcholine (DLPC), palmitoyl-oleoyl-phosphatidylcholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC) and
30 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.

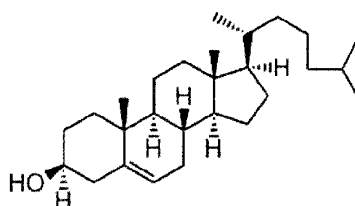
In some embodiments, the additional lipid comprises one of the following: (1) a phospholipid, (2) cholesterol or a derivative thereof; or (3) a mixture of a phospholipid and cholesterol or a derivative thereof. Examples of cholesterol derivatives include, but are not limited to, cholestanol, cholestanone, cholestenone, coprostanol, cholesteryl-2'-hydroxyethyl ether, cholesteryl-4'-hydroxybutyl ether, tocopherol and derivatives thereof, and mixtures thereof. Thus, in some embodiments, the RNA 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 RNA 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.

- 5 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*.

Polyethyleneglycol (PEG)-conjugated lipids

- 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 phosphatidylethanolamine (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 dialkoxypopylcarbamate such as ω -methoxy(polyethoxy)ethyl-N-(2,3-di(tetradecanoxo)propyl)carbamate or 2,3-di(tetradecanoxo)propyl-N-(ω -methoxy(polyethoxy)ethyl)carbamate, and the like.
- 10 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.

Embodiments of Lipoplex Particles

- 15 In some embodiments of the present disclosure, the RNA described herein may be present in RNA lipoplex particles.
- Lipoplexes (LPX) are electrostatic complexes which are generally formed by mixing preformed cationic lipid liposomes with anionic RNA. Formed lipoplexes possess distinct internal arrangements of molecules that arise due to the transformation from liposomal structure into compact RNA-lipoplexes. These formulations are generally characterized by their poor encapsulation of the RNA and incomplete entrapment of the RNA.
- 20 In certain embodiments, the RNA lipoplex particles include both a cationic lipid and an additional lipid. In an exemplary embodiment, the cationic lipid is DOTMA and the additional lipid is DOPE.
- 25 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.
- 30

- RNA 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 RNA lipoplex particles have an average diameter that ranges from about 250 nm to about 700 nm. In another embodiment, the RNA lipoplex particles have an average diameter that ranges from about 300 nm to about 500 nm. In an exemplary embodiment, the RNA lipoplex particles have an average diameter of about 400 nm.
- The RNA lipoplex particles and compositions comprising RNA lipoplex particles described herein are useful for delivery of RNA to a target tissue after parenteral administration, in particular after intravenous administration.

Embodiments of Lipid nanoparticles (LNPs)

- In some embodiments, RNA 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 RNA molecules are attached, or in which the one or more RNA 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 RNA in an aqueous buffer. In some embodiments, in the RNA LNPs described herein the RNA 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 RNA, encapsulated within or associated with the lipid nanoparticle.

5 In some embodiments, the LNP comprises from 40 to 60 mol percent, 40 to 55 mol percent, from 45 to 55 mol percent, or from 45 to 50 mol percent of the cationic lipid.

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.

10 In some embodiments, the steroid is present in a concentration ranging from 30 to 50 mol percent, from 30 to 45 mol percent, from 35 to 45 mol percent or from 35 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.

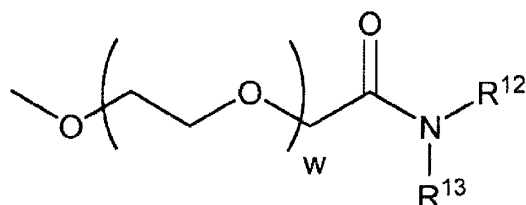
15 In some embodiments, the LNP comprises from 45 to 55 mol percent a cationic lipid; from 5 to 15 mol percent of a neutral lipid; from 30 to 45 mol percent of a steroid; from 1 to 5 mol percent of a polymer-conjugated lipid; and the RNA, encapsulated within or associated with the lipid nanoparticle.

20 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.

25 In some embodiments, the steroid is cholesterol.

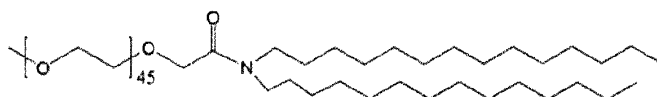
In some embodiments, the polymer conjugated lipid is a pegylated lipid. In some embodiments, the pegylated lipid has the following structure:



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

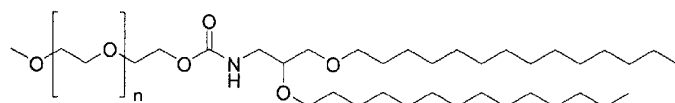
- 5 R^{12} and R^{13} are each independently a straight or branched, saturated or unsaturated alkyl chain containing from 10 to 30 carbon atoms, wherein the alkyl chain is optionally interrupted by one or more ester bonds; and w has a mean value ranging from 30 to 60. In some embodiments, R^{12} and 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
- 10 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, e.g., having the following structure:



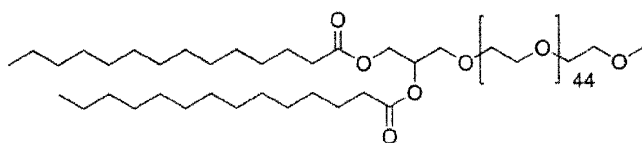
15

In some embodiments, a pegylated lipid is or comprises PEG₂₀₀₀-C-DMA, e.g., having the following structure:

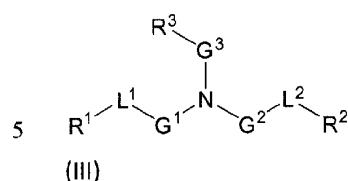


- 20 wherein n may have a mean value ranging from 30 to 60, such as about 50.

In some embodiments, the pegylated lipid is or comprises DMG-PEG 2000, e.g., having the following structure:



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



or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:
one of L^1 or L^2 is $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{O})-$, $-\text{O}-$, $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{O})\text{S}-$, $\text{SC}(\text{O})-$, $-\text{NR}^a\text{C}(\text{O})-$,
10 $-\text{C}(\text{O})\text{NR}^a-$, $\text{NR}^a\text{C}(\text{O})\text{NR}^a-$, $-\text{OC}(\text{O})\text{NR}^a-$ or $-\text{NR}^a\text{C}(\text{O})\text{O}-$, and the other of L^1 or L^2 is $-\text{O}(\text{C}=\text{O})-$,
 $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{O})-$, $-\text{O}-$, $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{O})\text{S}-$, $\text{SC}(\text{O})-$, $-\text{NR}^a\text{C}(\text{O})-$, $-\text{C}(\text{O})\text{NR}^a-$, $\text{NR}^a\text{C}(\text{O})\text{NR}^a-$,
 $-\text{OC}(\text{O})\text{NR}^a-$ or $-\text{NR}^a\text{C}(\text{O})\text{O}-$ 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;

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

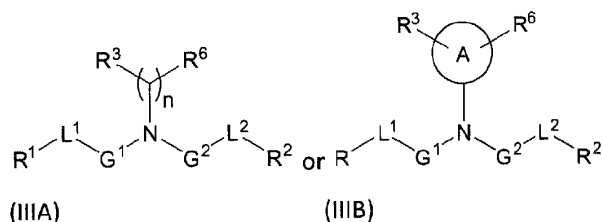
R^3 is H, OR^5 , CN, $-\text{C}(\text{O})\text{OR}^4$, $-\text{OC}(\text{O})\text{R}^4$ or $-\text{NR}^5\text{C}(\text{O})\text{R}^4$;

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.

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



wherein:

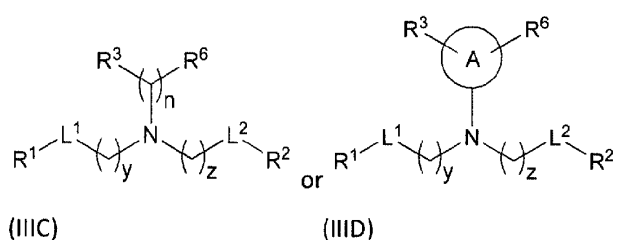
25 A is a 3 to 8-membered cycloalkyl or cycloalkylene ring;

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).

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

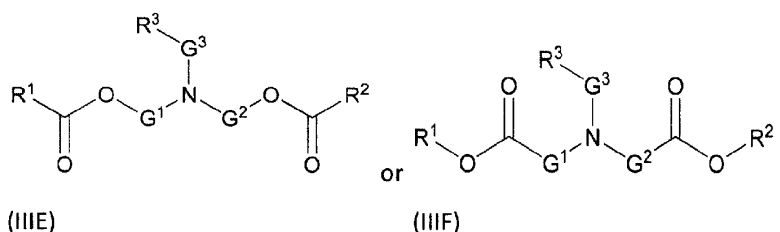


wherein y and z are each independently integers ranging from 1 to 12.

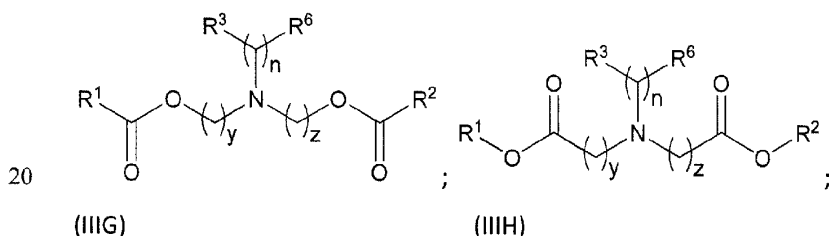
- 10 In any of the foregoing embodiments of Formula (III), one of L^1 or L^2 is $-O(C=O)-$. For example, in some embodiments each of L^1 and L^2 are $-O(C=O)-$. In some different embodiments of any of the foregoing, L^1 and L^2 are each independently $-(C=O)O-$ or $-O(C=O)-$. For example, in some embodiments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

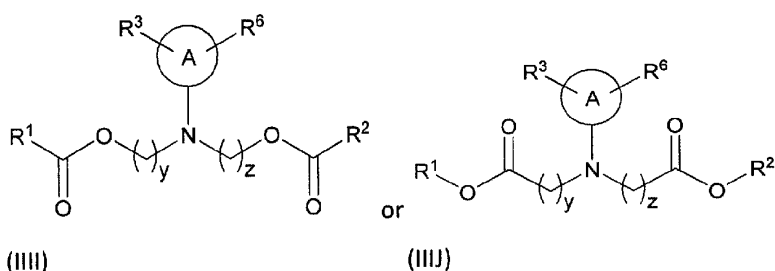
In some different embodiments of Formula (III), the lipid has one of the following structures

- 15 (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):





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.

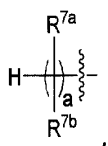
5 In some embodiments, n is 3. In some embodiments, n is 4. In some embodiments, n is 5. In some embodiments, n is 6.

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.

10 In some of the foregoing embodiments of Formula (III), R⁶ is H. In other of the foregoing embodiments, R⁶ is C₁-C₂₄ alkyl. In other embodiments, R⁶ is OH.

In some embodiments of Formula (III), G³ is unsubstituted. In other embodiments, G³ is substituted. In various different embodiments, G³ is linear C₁-C₂₄ alkylene or linear C₁-C₂₄ alkenylene.

15 In some other foregoing embodiments of Formula (III), R¹ or R², or both, is C₆-C₂₄ alkenyl. For example, in some embodiments, R¹ and R² each, independently have the following structure:



wherein:

R^{7a} and R^{7b} are, at each occurrence, independently H or C₁-C₁₂ alkyl; and

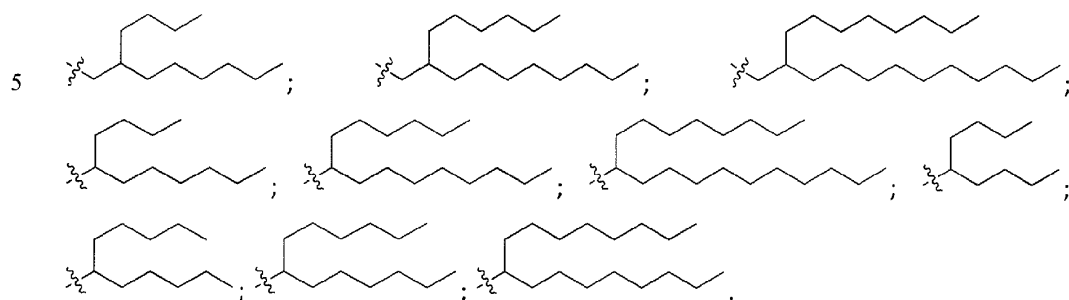
20 a is an integer from 2 to 12,

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
25 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:



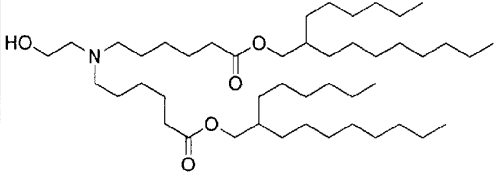
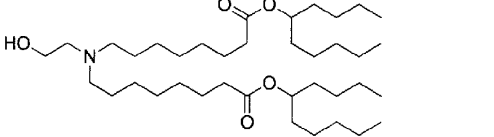
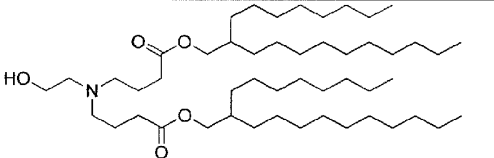
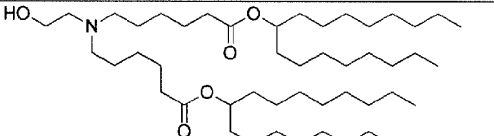
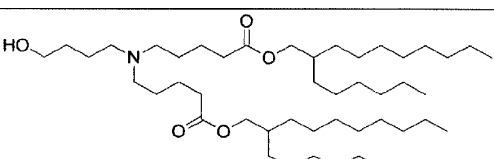
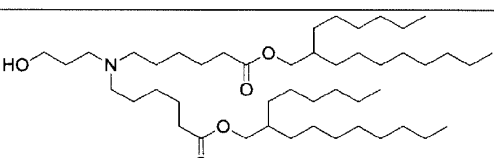
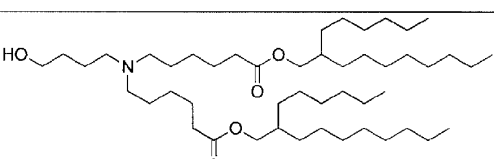
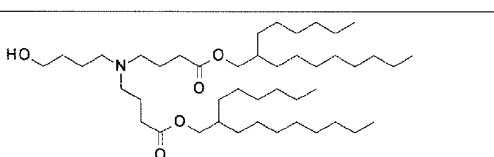
In some of the foregoing embodiments of Formula (III), R^3 is OH, CN, $-C(=O)OR^4$, $-OC(=O)R^4$ or $-NHC(=O)R^4$. In some embodiments, R^4 is methyl or ethyl.

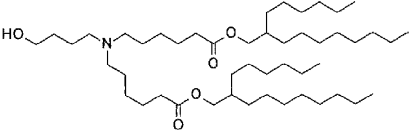
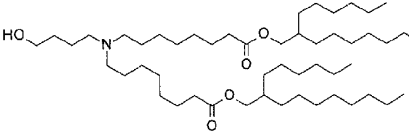
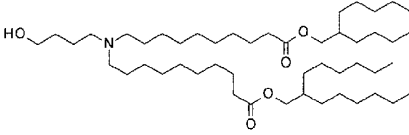
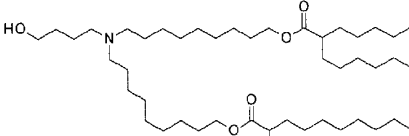
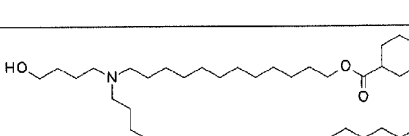
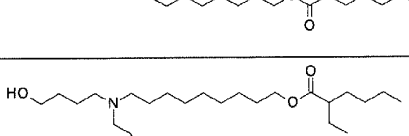
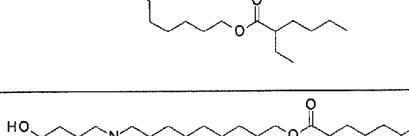
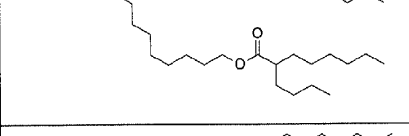
- 10 In various different embodiments, the cationic lipid of Formula (III) has one of the structures set forth in the table below.

Representative Compounds of Formula (III).

No.	Structure
III-1	
III-2	
III-3	

No.	Structure
III-4	
III-5	
III-6	
III-7	
III-8	
III-9	
III-10	

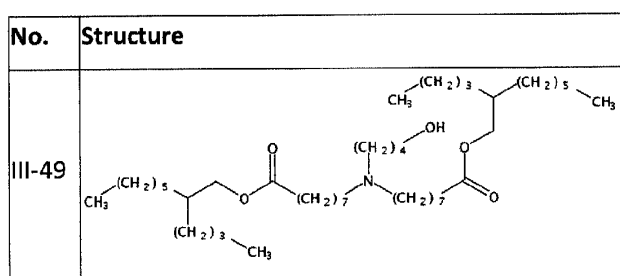
No.	Structure
III-11	
III-12	
III-13	
III-14	
III-15	
III-16	
III-17	
III-18	

No.	Structure
III-19	
III-20	
III-21	
III-22	
III-23	
III-24	
III-25	
III-26	

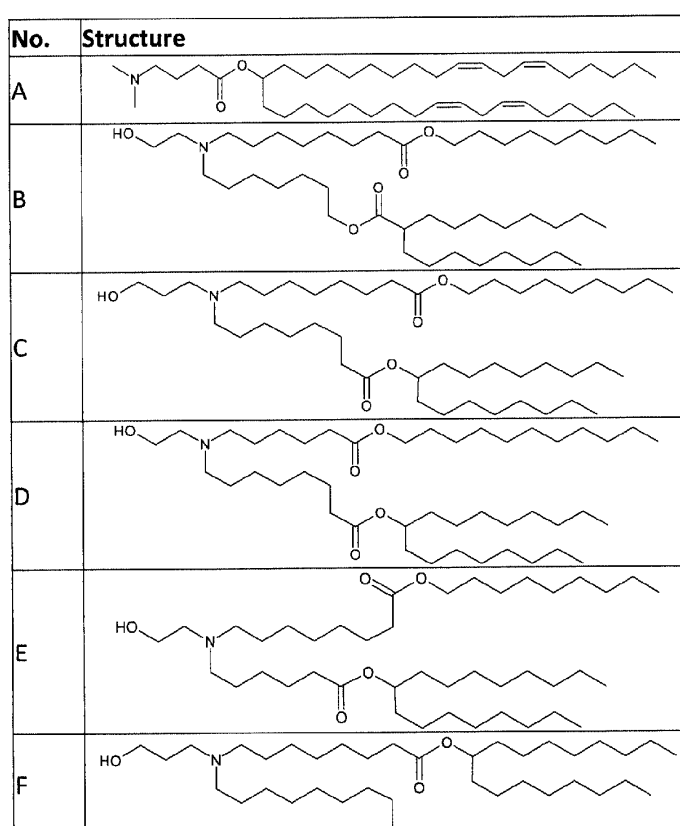
No.	Structure
III-27	
III-28	
III-29	
III-30	
III-31	
III-32	
III-33	

No.	Structure
III-34	
III-35	
III-36	
III-37	
III-38	
III-39	
III-40	

No.	Structure
III-41	
III-42	
III-43	
III-44	
III-45	
III-46	
III-47	
III-48	



Further representative cationic lipids are as follows:

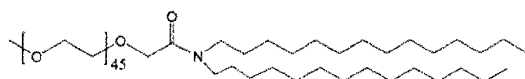


- 5 In some embodiments, the LNP comprises a cationic lipid shown in the above table, e.g., a cationic lipid of Formula (B) or Formula (D), in particular a cationic lipid of Formula (D), RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the neutral lipid is DSPC. In some embodiments,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

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 (DSPC)) 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), RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the pegylated lipid is ALC-0159.

ALC-0159:

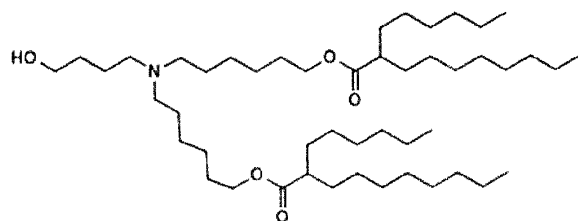


In some embodiments, RNA described herein is formulated in a composition comprising a cationic lipid, a neutral lipid, a steroid, a polymer conjugated lipid; and the RNA.

In some embodiments, the composition comprises a lipid of Formula (III), RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the pegylated lipid is ALC-0159. In some embodiments, the cationic lipid is a lipid of Formula (III), the neutral lipid is DSPC, the steroid is cholesterol, and the pegylated lipid is ALC-0159.

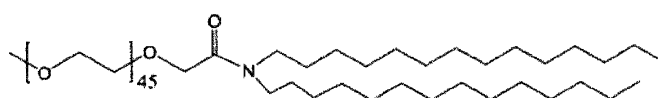
In some embodiments, the composition comprises ALC-0315, RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the pegylated lipid is ALC-0159. In some embodiments, the cationic lipid is ALC-0315, the neutral lipid is DSPC, the steroid is cholesterol, and the pegylated lipid is ALC-0159.

ALC-0315 = ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate) / 6-[N-6-(2-hexyldecanoyloxy)hexyl-N-(4-hydroxybutyl)amino]hexyl 2-hexyldecanoate

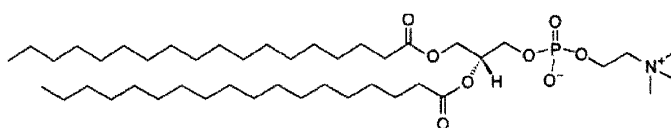


ALC-0159 = 2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide / 2-[2-(ω-methoxy (polyethyleneglycol2000) ethoxy]-N,N-ditetradecylacetamide

5

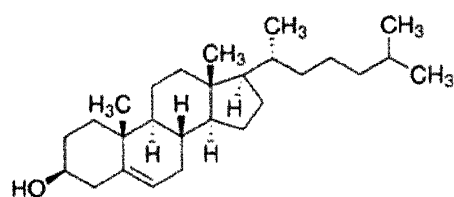


DSPC = 1,2-Distearoyl-*sn*-glycero-3-phosphocholine



10

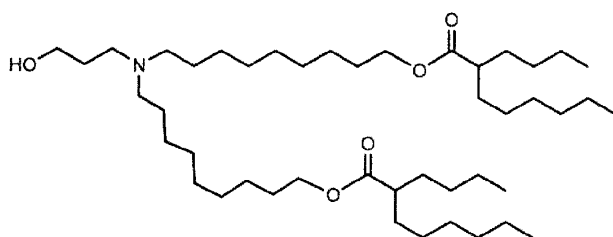
Cholesterol:



15 In some embodiments, the composition comprises ALC-0366, RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the pegylated lipid is ALC-0159. In some embodiments, the cationic lipid is ALC-0366, the neutral lipid is DSPC, the steroid is cholesterol, and the pegylated lipid is ALC-0159.

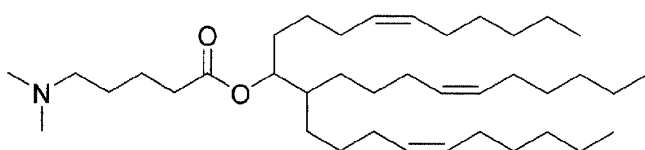
20

ALC-0366 = ((3-hydroxypropyl)azanediyl)bis(nonane-9,1-diyl) bis(2-butyloctanoate)

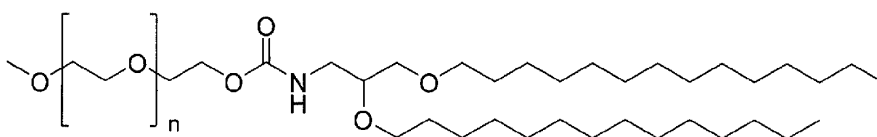


In some embodiments, the composition comprises 3D-P-DMA, RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the pegylated lipid is PEG₂₀₀₀-C-DMA. In some embodiments, the cationic lipid is 3D-P-DMA, the neutral lipid is DSPC, the steroid is cholesterol, and the pegylated lipid is PEG₂₀₀₀-C-DMA.

3D-P-DMA = (6Z,16Z)-12-((Z)-dec-4-en-1-yl)docosa-6,16-dien-11-yl
(dimethylamino)pentanoate



"PEG₂₀₀₀-C-DMA": 3-N-[(ω-Methoxy poly(ethylene glycol)2000) carbamoyl]-1,2-dimyristyloxy-propylamine (MPEG-(2 kDa)-C-DMA or Methoxy-polyethylene glycol-2,3-bis(tetradecyloxy)propylcarbamate (2000))



wherein n has a mean value ranging from 30 to 60, such as about 50.

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.

Compositions comprising RNA

A composition comprising one or more RNAs described herein, e.g., in the form of RNA particles, may comprise salts, buffers, or other components as further described below.

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 RNA 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 RNA particles such as for freezing RNA 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 RNA particle compositions described herein have a pH suitable for the stability of the RNA particles and, in particular, for the stability of the RNA. 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 RNA activity during storage, freezing, and/or lyophilization, for example to reduce or prevent aggregation, particle collapse, RNA 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
5 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
10 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.

15 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,
20 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.
25 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,
30 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
5 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
10 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.

15 In some embodiments, RNA particle compositions may include sucrose. Without wishing to be bound by theory, sucrose functions to promote cryoprotection of the compositions, thereby preventing RNA (especially mRNA) particle aggregation and maintaining chemical and physical stability of the composition. In some embodiments, RNA particle compositions may include alternative cryoprotectants to sucrose. Alternative stabilizers include, without
20 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.

25 Some embodiments of the present disclosure contemplate the use of a chelating agent in an RNA 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
30 accelerated RNA 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)glycoether-N,N,N',N'-tetraacetic acid. In some embodiments, the chelating agent is EDTA or a salt of EDTA. In an exemplary embodiment, the chelating agent is EDTA disodium dihydrate. In some embodiments, the EDTA is at a concentration from about 0.05 mM to about 5 mM, from about 0.1 mM to about 2.5 mM or from about 0.25 mM to about 1 mM.

In an alternative embodiment, the RNA particle compositions described herein do not comprise a chelating agent.

Pharmaceutical compositions

The agents described herein may be administered in pharmaceutical compositions or medicaments and may be administered in the form of any suitable pharmaceutical composition. In some embodiments, the pharmaceutical composition is for therapeutic or prophylactic treatments, e.g., for use in treating or preventing a disease such as bacterial infection.

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 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.

- 5 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 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
- 10 desired reaction in a treatment of a disease may also be delay of the onset or a prevention of the onset of said disease or said condition. An effective amount of the pharmaceutical compositions described herein will depend on the condition to be treated, the severeness of the disease, the individual parameters of the patient, including age, physiological condition, size and weight, the duration of treatment, the type of an accompanying therapy (if present),
- 15 the specific route of administration and similar factors. Accordingly, the doses administered of the pharmaceutical compositions described herein may depend on various of such parameters. In the case that a reaction in a patient is insufficient with an initial dose, higher doses (or effectively higher doses achieved by a different, more localized route of administration) may be used.

- 20 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
- 25 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
- 30 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
5 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,
10 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
15 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.

20 **Routes of administration of pharmaceutical compositions**

In some embodiments, the pharmaceutical compositions described herein may be
administered intravenously, intraarterially, subcutaneously, intradermally, dermally,
intranodally, or intramuscularly. In some embodiments, the pharmaceutical composition is
formulated for local administration or systemic administration. Systemic administration may
25 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
30 by intravenous administration. In some embodiments, the pharmaceutical compositions are
formulated for respiratory/pulmonary administration route and/or administered by

respiratory/pulmonary administration route, e.g., by inhalation. In this manner, the RNA described herein may be locally/regionally or systemically delivered to lungs and/or the respiratory tract. The lungs may also be used as a portal of entry to the body, enabling delivery of the RNA via the airways into the bloodstream. Thus, inhaled formulations may be used for
5 systemic delivery.

Use of compositions

Compositions described herein may be used in the therapeutic or prophylactic treatment of various diseases, in particular diseases caused by or related to bacterial infection.

10 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
15 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.
20 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.

The term "infectious disease" or "infection" refers to any disease which can be transmitted from individual to individual or from organism to organism, and is caused by a microbial agent. Infectious diseases are known in the art and include, for example, a viral disease, a bacterial
25 disease, or a parasitic disease, which diseases are caused by a virus, a bacterium, and a parasite, respectively.

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
30 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.

The agents and methods described herein find application in the treatment of a bacterial infection in a subject. In some embodiments, the agents and methods described herein are useful for treating a bacterial infection caused by Gram-positive bacteria and/or Gram-negative bacteria and/or atypical bacteria.

The term "bacterial infection" refers to the invasion of a subject by pathogenic bacteria. This includes the excessive growth of bacteria which are normally present in or on the body of the subject, but more generally, a bacterial infection can be any situation in which the presence of a bacterial population(s) is damaging to a host organism. Thus, for example, a subject suffers from a bacterial population when excessive numbers of a bacterial population are present in or on the subject's body, or when the effects of the presence of a bacterial population(s) is damaging to the cells, tissue, or organs of the subject.

The agents and methods described herein are applicable to a variety of bacterial infections caused by different bacterial strains, in particular bacteria which are a major cause of morbidity and mortality in hospital-based infections, e.g., *Staphylococcus aureus*, *Streptococcus pneumoniae*, various Enterococci, and *Pseudomonas aeruginosa*. However, the approach described below is clearly applicable to any human bacterial pathogens including but not restricted to *Mycobacterium tuberculosis*, *Neisseria gonorrhoeae*, *Haemophilus influenza*, *Acinobacter*, *Escherichia coli*, *Shigella dysenteriae*, *Streptococcus pyogenes*, *Helicobacter pylori*, and *Mycoplasma species*. In some embodiments, the peptidoglycan hydrolase, e.g., endolysin, used is effective in breaking down peptidoglycan in bacterial cell wall of the bacterium causing the infection to be treated. In some embodiments, the peptidoglycan hydrolase, e.g., endolysin, used is derived from a bacteriophage of the bacterium causing the infection to be treated.

In some embodiments, the bacterial infection is caused by at least one bacterium selected from the genera Enterococcus, Staphylococcus, Streptococcus, Bacillus, Acinetobacter, Burkholderia, Coxiella, Francisella, Yersinia, Klebsiella, Escherichia, Enterobacter and Pseudomonas.

In some embodiments, the bacterial infection is caused by at least one bacterium selected from the genera Enterococcus, Staphylococcus, Acinetobacter, Burkholderia, Klebsiella, Escherichia, Enterobacter and Pseudomonas.

In some embodiments, the bacterial infection is caused by at least one bacterium selected from *Enterococcus faeculis*, *Enterococcus faecium*, *Staphylococcus aureus*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Bacillus cereus*, *Bacillus subtilis*, *Haemophilus influenzae*, *Acinetobacter baumannii*, *Burkholderia multivorans*, *Burkholderia cenocepacia*, *Burkholderia cepacia*, *Burkholderia mallei*,

Burkholderia pseudomallei, *Coxiella burnetii*, *Citrobacter freundii*, *Escherichia coli*, *Enterobacter cloacae*, *Enterobacter aerogenes*, *Francisella tularensis*, *Yersina pestis*, *Klebsiella pneumoniae*, *Serratia marcesens*, *Salmonella typhi*, *Salmonella typhimurum*, *Stenotrophomonas maltophilia*, *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae*. In
5 some embodiments, the bacterial infection is caused by at least one bacterium selected from *Enterococcus faeculis*, *Enterococcus faecium*, *Staphylococcus aureus*, *Acinetobacter baumannii*, *Burkholderia multivorans*, *Burkholderia cenocepacia*, *Burkholderia cepacia*, *Klebsiella pneumonia* and *Pseudomonas aeruginosa*.

In some embodiments, the bacterial infection is caused by Gram-positive bacteria selected
10 from *Enterococcus faeculis*, *Enterococcus faecium*, *Staphylococcus aureus*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Bacillus cereus* and *Bacillus subtilis*.

In some embodiments, the infection is caused by Gram-negative bacteria, such as
15 *Haemophilus influenzae*, *Acinetobacter baumannii*, *Burkholderia multivorans*, *Burkholderia cenocepacia*, *Burkholderia cepacia*, *Burkholderia mallei*, *Burkholderia pseudomallei*, *Coxiella burnetii*, *Citrobacter freundii*, *Escherichia coli*, *Enterobacter cloacae*, *Enterobacter aerogenes*, *Francisella tularensis*, *Yersina pestis*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae*.

In some embodiments, the bacterial infection is caused by drug-resistant bacteria. Such drug-
20 resistant bacteria are bacteria that are resistant to one or more antibacterials other than the agents described herein. The term "resistance", "antibacterial resistance", or "drug-resistant" refers to bacteria that are able to survive exposure to one or more antibacterial drugs. In some embodiments, the drug-resistant bacteria include *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Streptococcus pneumoniae* (including penicillin-resistant *Streptococcus pneumoniae*), *Staphylococcus aureus* (including vancomycin-resistant *Staphylococcus aureus* (VISA)), methicillin-resistant *Staphylococcus aureus* (MRSA) (including hospital-acquired MRSA, community acquired MRSA and coagulase negative staphylococci), *Acinetobacter baumannii*, *Burkholderia multivorans*, *Burkholderia cenocepacia*, *Burkholderia cepacia*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* (including
25 penicillin-resistant *Neisseria gonorrhoeae*).
30

In some embodiments, the drug-resistant bacteria is a multiple drug resistant bacteria. The language "multiple drug resistant bacteria" includes bacteria that is resistant to two or more drugs, e.g., antibiotics, typically used for the treatment of such bacterial infections, for example, tetracycline, penicillin, cephalosporins (e.g., ceftriazone or cefixime), glycopeptides
5 (e.g. vancomycin), quinolones (e.g., norfloxacin, ciprofloxacin or ofloxacin), co-trimoxazole, sulfonamides, aminoglycosides (e.g., kanamycin or gentamicin) and macrolides (e.g., azithromycin).

In some embodiments, the invention provides a method for treating bacterial skin infections in a subject using the agents and methods described herein.

10 In some embodiments, the bacterial skin infections are caused by *Streptococcus pyogenes*, *Streptococcus agalactiae*, or *Staphylococcus aureus*, including MRSA and/or VRSA.

In some embodiments, the invention provides a method for treating pneumonia in a subject using the agents and methods described herein.

The term "pneumonia" refers to an inflammatory condition of the lungs caused by a bacterial
15 infection. In some embodiments, the pneumonia is caused by a *Klebsiella pneumoniae*, *Streptococcus pneumoniae* or *Staphylococcus aureus* infection. In some embodiments, the pneumonia is nosocomial pneumonia (e.g., hospital-acquired pneumonia) or community-acquired pneumonia. In some embodiments, the pneumonia is caused by penicillin-resistant *Klebsiella pneumoniae*.

20 In some embodiments, the invention provides a method for treating a condition selected from urinary tract infections, sepsis and bacteremia in a subject using the agents and methods described herein.

In some embodiments, administration of RNA may be performed by single administration or by multiple administrations.

25 In some embodiments, an amount the RNA described herein from 0.1 µg to 300 µg, 0.5 µg to 200 µg, or 1 µg to 100 µg, such as about 1 µg, about 3 µg, about 10 µg, about 30 µg, about 50 µg, or about 100 µg may be administered per dose.

In some embodiments, a regimen described herein includes at least one dose. In some
30 embodiments, a regimen includes a first dose and at least one subsequent dose. In some embodiments, the first dose is the same amount as at least one subsequent dose. In some embodiments, the first dose is the same amount as all subsequent doses. In some

embodiments, the first dose is a different amount as at least one subsequent dose. In some embodiments, the first dose is a different amount than all subsequent doses. In some embodiments, a regimen comprises two doses. In some embodiments, a provided regimen consists of two doses.

- 5 In some embodiments, a regimen administered to a subject may comprise a plurality of doses (e.g., at least two doses, at least three doses, or more). In some embodiments, a regimen administered to a subject may comprise a first dose and a second dose, which are given at least 2 weeks apart, at least 3 weeks apart, at least 4 weeks apart, or more. In some
10 embodiments, such doses may be at least 1 month, at least 2 months, at least 3 months, at least 4 months, at least 5 months, at least 6 months, at least 7 months, at least 8 months, at least 9 months, at least 10 months, at least 11 months, at least 12 months, or more apart. In some embodiments, doses may be administered days apart, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60 or
15 more days apart. In some embodiments, doses may be administered about 1 to about 3 weeks apart, or about 1 to about 4 weeks apart, or about 1 to about 5 weeks apart, or about 1 to about 6 weeks apart, or about 1 to more than 6 weeks apart. In some embodiments, doses may be separated by a period of about 7 to about 60 days, such as for example about 14 to about 48 days, etc. In some embodiments, a minimum number of days between doses may
20 be about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or more. In some embodiments, a maximum number of days between doses may be about 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, or fewer. In some embodiments, doses may be about 21 to about 28 days apart. In some embodiments, doses may be about 19 to about 42 days apart.
25 In some embodiments, doses may be about 7 to about 28 days apart. In some embodiments, doses may be about 14 to about 24 days. In some embodiments, doses may be about 21 to about 42 days.

Citation of documents and studies referenced herein is not intended as an admission that any
30 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: Bactericidal activity of lysins secreted from *P.pastoris* cells

P. pastoris is a well established eukaryotic expression system for the high level production of recombinant proteins. Proteins in *P.pastoris* are typically translated, folded and processed by a similar set of factors as in higher eukaryotes, such as human cells. In particular, expression of proteins containing N-terminal signal sequences results in their targeting to the secretory pathway and secretion into the extracellular space. The steps occurring during this process are highly conserved between eukaryotes and involve the translocation of the protein across the membrane of the endoplasmic reticulum (ER), its glycosylation and its trafficking to the plasma membrane. The use of human cells for the delivery of lysins by mRNA technology necessitates that lysins are able to pass through the secretory pathway in an active state. Since the majority of lysins are derived from phages and have evolved to be synthesized in a bacterial cytoplasm, it has been unclear if different lysins can be expressed and secreted in a eukaryotic cell. Using *P.pastoris* as an expression system, the inventors established here that a set of 11 model lysins can be trafficked through the eukaryotic secretory pathway and at the same time maintain their bactericidal activity.

For this purpose, the amino acid sequences (excluding the start methionine) of 11 lysins with predicted activity against *S.aureus* were modified with the N-terminal pre- and pro sequences of *S.cerevisiae* alpha mating factor, followed by a Kex2/Ste13 signal peptidase site, a hexa-histidine tag and a 3C protease cleavage site (see Appendix 1). The signal sequences target the lysin constructs to the secretory pathway and are cleaved off at the signal peptidase sites after translocation into the ER. The hexa-histidine tag and 3C sites were further introduced to facilitate lysin purification by affinity capture and removal of the hexa-histidine tag after proteolytic cleavage, respectively. The constructs were codon optimized for expression in *P. pastoris* and ordered from Genescript in plasmids conferring antibiotic resistance and control of lysin expression by an AOX1 methanol-inducible promoter. Plasmids were linearized by restriction and transformed into a wild type (mut+) *P. pastoris* strain, followed by selection of transformants on YPD plates containing 100 µg/ml Zeocin. Multiple colonies per construct were transferred each into 3 ml BMGY medium in 24 well plates and grown shaking at 30 °C. After 24 hours, protein expression was induced by exchanging the medium to BMMY

(containing 1% methanol). The medium was supplemented with 1% methanol every 24 hours and culture supernatants were harvested after 72 hours of expression, flash frozen in liquid nitrogen and stored at -80 °C.

5 For each transformant, lysin secretion into culture supernatants was estimated by SDS-PAGE and Coomassie staining. Supernatants with highest apparent levels of secretion were serially diluted with lysin buffer (20mM NaPO₄ pH 6.0, 100mM NaCl) and tested for bactericidal activity. For this purpose, 30 µl of supernatants (and dilutions thereof) were mixed with *S.aureus* cells in 70 µl BHI medium such that the reaction contained 5x10⁵ cfu/ml. The optical
10 density of these solutions (OD) at 620nm was monitored in a plate reader over the course of 24 hours to quantify bacterial growth. In these experiments, *S.aureus* cells were able to grow in the presence of supernatants from a parental *P.pastoris* strain, showing that the culture medium itself does not have lytic activity. In contrast, *S.aureus* failed to grow in presence of supernatants from strains expressing each of the lysin constructs. These experiments show
15 that the level of active lysins in undiluted supernatants was sufficiently high to lyse all *S.aureus* cells in the culture (Figure 1A). Depending on the construct and clone, mixing of supernatants and buffer resulted in bacterial growth at different folds of dilution. For example, bacterial growth started to appear after a one-fold dilution of the L0475 supernatant, whereas growth only appeared at an eight-fold dilution of the L0502 supernatant, suggesting that the L0502
20 supernatant had a higher concentration and/or lytic activity compared to the L0475 supernatant. Similarly, different clones of one lysin construct varied in the fold dilutions at which bacterial growth started to appear (Figure 1B), suggesting differences in the levels of expression and secretion between different transformants. Taken together, the data demonstrate that lysins can be channelled through the secretory pathway and maintain their
25 activity after secretion. Surprisingly, all of the tested lysins were expressible in an active form in *P. pastoris*.

Example 2: Comparison of lysin activity after expression in *E.coli* and *P.pastoris*

Expression of the lysins in *P.pastoris* was compared to their expression in *E.coli*. Constructs
30 containing N-terminal hexa-histidine tags and 3C protease cleavage sites (but lacking the export signal sequences) were codon optimized for expression in *E.coli* followed by DNA

synthesis at TWIST Biosciences. The constructs were subsequently cloned into an IPTG-inducible pET21 vector conferring resistance to Kanamycin. After transformation into BL21 cells, lysins were expressed in 2 ml cultures of TB + 1.5% lactose for 24 hours at 25 °C. Bacterial pellets were lysed in 1.5 ml lysis buffer (50 mM Hepes pH 7.5, 300 mM NaCl, 20 mM Imidazole, 0.5 mM TCEP, 2 µl Benzonase per ml, 1x FastBreak) and centrifuged at 21130 rcf at 4°C for 15 min. Cleared lysates were applied to 100 µl Nickel affinity matrix, washed with 15 column volumes each of wash buffer I (50 mM Hepes pH 7.5, 300 mM NaCl, 20 mM Imidazole, 0.5 mM TCEP) and wash buffer II (50 mM Hepes pH 7.5, 300 mM NaCl, 40 mM Imidazole, 0.5 mM TCEP) and eluted in 4 column volumes elution buffer (50 mM Hepes pH 7.5, 300 mM NaCl, 500 mM Imidazole, 0.5 mM TCEP). Eluates were buffer exchanged into 20 mM NaH₂PO₄/Na₂HPO₄ pH 6, 100 mM NaCl using a CentriPure P96 Gelfiltration column array, snap frozen in liquid nitrogen and stored at -80 °C.

SDS-PAGE of eluates followed by Coomassie staining showed that three of the 11 constructs could not be purified in a soluble form. Surprisingly, *S.aureus* killing assays (as performed in Example 1) showed that several of these eluates had no lytic activity. Of the constructs that could be purified in a soluble form, a large fraction was also inactive (7 out of 11). In contrast, all lysin constructs showed lytic activity against *S.aureus* cells after expression and secretion in *P.pastoris* (Table 3, Figure 1A). Taken together, this data indicates that the use of *P.pastoris* might facilitate the expression of active lysins, possibly due to the increased folding capacity of eukaryotic cells.

Table 3: Comparison of the bactericidal activity of lysins purified in *E.coli* and lysins secreted in *P.pastoris*. Domain architecture, internal name and NCBI identifiers are indicated. Lysin constructs that were found active in *P.pastoris* supernatants but could not be purified from *E.coli* in an active form are shaded in grey. soluble(Ec): classification if protein could be purified in a soluble form from *E.coli*. active(Ec): classification if the lysin purified from *E.coli* was active in *S.aureus* killing assays. active(Pp): classification if the lysin secreted in *P.pastoris* was active in preventing *S.aureus* growth. y: yes. n: no.

domain architecture	Internal name	NCBI ID	soluble (Ec)	active (Ec)	active (Pp)
Ami2	L0487	WP_115921155	n	n	y
CHAP-Ami2-SH3	L0489	ABF71642	y	n	y
	L0479	AWH13065	y	n	y
	L0480	YP_024461	y	y	y
	L0473	WP_195722665	y	y	y

LysM-CHAP	L0475	YP_009268702	y	n	y
	L0499	WP_070648624	n	n	y
	L0482	SWP_143119924	y	y	y
PepM23-SH3	L0483	Lysostaphin	y	y	y
phage lysozyme	L0484	PTF44644	y	n	y
SH3-Ami3	L0502	WP_053025743	n	n	y

Appendix 1 – Amino acid and DNA sequences of 11 lysins for expression in *P.pastoris*

>L0473_Protein

5 MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLGDFDVAVLFPFSNSTNNGLLFINTTIIASIAAKEEGV
SLEKREAEAHHHHHHLEVLFGQGPMAKTKKQIEQRLQAYAKGTVDSPYRITKPTSYDPSFGVMEVGAIDADRYHAQ
CQDLITDYVLWLTNDKVRPWGNAKDQINQSYGTGFKIHKNLPSTVAKTGWIAVFTSGTYSQYGHIGIVYSPGDTNK
FHIIEQNWNGWANKKPLTRWDNYYGLTHFIEVPYAKETITTKKTAPKAATSKPAKKISVSHNNINLQSNFAKRGYNP
10 KGVVIHNDAGRSTAQQYENSLKGANQNRLNNGIAHAYADSSHIWYAIERIAHTGDGVKKGTYGNYDYFGIEVC
QSMSASDKEFMANEQVVLQEVARLLKKWKLPANRNTVRLHLEFSATSCPHRSLTMRGTGFNPVTQGRPSDAIKKLK
DIWIKEIRAYMEGKTPSIKTVANKPGSASTPATKSNSDGWKVNAVNTYYKTERASFTVGSTRIAVRSTGPFTTCTPFTG
WLNPGVTITYDEV MVQNGFVWLGYDSYNGRRYVPIRTAQGTPGSSNYVLGPLWGTIS*

>L0473_DNA

15 ATGAGATTCCATCTATTTTCACTGCTGTTTTGTTTCTGCTCTCTCTGCTTTGGCTGCTCCAGTTAACTACTACTAC
TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTACTCTGATTTGGAGGGAGATTTTGATGTTG
CTGTTTTGCCTTTCTCTAACTCTACTAACAACGGTTTGTGTTTATTAACACTACTATCGCTTCTATTGCTGCTAAA
GAAGAGGGTGTTTCTTTGGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAGGTTTGTTC
AAGGTCCAATGGCTAAGACTAAGAAACAAATCGAGCAAGATTGCAAGCTTATGCTAAAGGTACTGTTGATTC
20 TCCTTACAGAATTACTAAGCCAATCTTATGATCCTTCTTCGGTGTTATGGAAGTTGGTGCTATTGATGCTGA
TAGATACTATCATGCTCAATGTCAAGATTTGATCACTGATTACGTTTTGTGGTTGACTGATAACAAAGTTAGACC
ATGGGGTAACGCTAAGGATCAAATTAATCAATCTACGGTACTGGTTTTAAGATCCACAAGAATTTGCCTTCTA
CTGTTGCTAAAACCTGGTTGGATTGCTGTTTTCACTTCTGGTACTTACTCTCAATATGGTCATATTGGTATTGTTTA
CTCTCCAGGAGATACTAACAAGTTCCACATCATCGAACAAAACCTGGAACGGTTGGGCTAACAAGAAACCATTTG
25 ACTAGATGGGATACTACTACGGTTTGACTCATTTTCATCGAAGTTCCTTACGCTAAGGAGACTATCACTACTAA
GAAAACCTGCTCCAAAGGCTGCTACTTCTAAACCTGCTAAGAAAATTTCTGTTTCTCATAACAACATTAACTTGCA
ATCTAATTTTGCTAAAAGAGGTTACAACCCAAAGGTTGTTGTTATTCACAATGATGCTGGTAGATCTACTGCTC
ACAATACGAAAATCTTTGAAGGGTGCTAACCAAAATAGATTGAACAATGGTATTGCTCATGCTTACGCTGAT
TCTTCTCACATTTGGTATGCTATTGAACCTGAGAGAATTGCTTACCATACTGGAGATGGTGTTAAGAAAGGTAC
30 TGGTAACTACGATTACTTCGGTATCGAAGTTTGTCAATCTATGTCTGCTTCTGATAAAGAGTTTATGGCTAATGA
GCAAGTTGTTTTGCAAGAGGTTGCTAGATTGTTGAAGAAATGGAAGTTGCCAGCTAACAGAAACACTGTTAGA
TTGCATTTGGAATTTCTGCTACTTCTGTCTCACAGATCTTTGACTATGAGAACTGGTTTCAATCCAGTTACTC
AAGGTAGACCTTCTGATGCTATTATTAAGAAATGAAGGATATCTGGATCAAGGAAATCAGAGCTTACATGGA
GGGTAAACCTCCATCTATTAAACCTGTTGCTAACAAGCCAGGTTCTGCTTCTACTCCTGCTACTAAATCTAATTC
35 TGATGGTTGGAAGGTTAACGCTTACAACACTTACTACAAGACTGAGAGAGCTTCTTTACTGTTGGTTCTACTA
GAATTGCTGTTAGATCTACTGGTCCATTCACTACTGTCTTCTACTGGTTGGTTGAACCCAGGTGTTACTATCA

CTTACGATGAAGTTATGGTTCAAAACGGTTTCGTTTGGTTGGGTTACGATTCTTATAATGGTAGAAGATACGTT
CCAATTAGAACTGCTCAAGGTACTCCTGGTTCTCTAATTATGTTTTGGGTCCTTGTGGGGTACTATTCTTAA

>L0475_Protein

5 MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSdleGDFDVAVLpFSNSTNNGLLfINTTIAIAAKEEGV
SLEKREAEahhhhhHLEVLfQGPMQAKLTKKEFIEWLKTSEGKQFNADLWYGFQCFDYANAawKVLfGLLLKGLG
AKDIPFANNFDGLATVYQNTPDfLAQPGDMVVFgSNYGAGYGHVAWVIEATLDYIIVYEQNWLGgGWTDGIEQP
GWGWEKVTrrQHAYDFPMWFIRPNfKSEIAPRSVQSPTQAPKKETAKPQPKAVELKIKDVVKGYDLpKRGSNPK
GIVIHNDAGSKGATAEAYRNGLVNAPLSRLEAGIAHSYVSGNTVWQALDESQVGWHTANQLGNkYYYGIEVCQS
10 MGADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEFTSTSCPHRSSVLHTGFDpVTRGLLPEDKRLQLKDY
FIKQIRAYMDGKIPVATVSNDSSASSNTVKPVASAWKRnKYGYTYYMEESARFTNGNQpITVRKVGPFLYCPVGYQF
QPGGYCDYTEVMLQDGHVWVGyTWEGQRYYPiRTWNGSAPPNQILGDLWGEIS*

>L0475_DNA

15 ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTAC
TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTATTCTGATTGGAGGGAGATTTTGATGTTG
CTGTTTTGCCATTCTCTAACTCTACTAACAACGGTTTGTGTTTATTAACACTACTATCGTTCTATTGCTGCTAA
GGAAGAGGGTGTTTCTTTGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAGGTTTTGTTT
CAAGGTCCTATGCAGGCTAAGTTGACTAAGAAAGAGTTTATTGAGTGGTTGAAGACTTCTGAAGGTAACAAT
20 TCAACGCTGATTGTGGTATGGTTTTCAATGTTTCGATTACGCTAATGCTGCTTGGAAAGTTTTGTTGGTTTGT
TGTTGAAGGGTTTGGGTGCTAAGGATATCCCTTCGCTAACAACTTCGATGGTTGGCTACTGTTACCAAAAC
ACTCCAGATTTCTTGGCTCAACCTGGAGATATGGTTGTTTTCGGTTCTAATTATGGTGTGTTACGGTCATGTT
GCTTGGGTATTGAAGCTACTTTGGATTACATCATCGTTTACGAGCAAACTGGTTGGGTGGTGGTGGACTG
ATGGTATTGAACAACAGGTTGGGGTGGGAGAAGGTTACTAGAAGACAACACGCTTATGATTTCCAATGTG
25 GTTCATCAGACCTAACTTCAAGTCTGAAATCGCTCCTAGATCTGTTCAATCTCCAACCTCAAGCTCCTAAGAAAGA
AACTGCTAAGCCACAACCTAAAGCTGTTGAGTTGAAGATTATTAAGGATGTTGTTAAGGGTTACGATTTGCCAA
AGAGAGGTTCTAACCCTAAGGGTATCGTTATCCATAACGATGCTGGTTCTAAGGGTGTCTACTGCTGAAGCTTAT
AGAAACGGTTTGGTTAATGCTCCATTGTCCAGATTGGAAGCTGGTATTGCTCATTCTTACGTTTCTGGTAACACT
GTTTGCAAGCTTTGGATGAATCTCAAGTTGGTTGGCACACTGCTAACCAATTGGGTAATAAGTACTATTACGG
30 TATTGAGGTTTGTCAATCTATGGGTGCTGATAACGCTACTTCTTGAAGAACGAACAAGCTACTTTCCAAGAGT
GTGCTAGATTGTTGAAGAAATGGGGTTTGGCAGCTAACAGAAACACTATCAGATTGCACAACGAGTTTACTTCT
ACTTCTGTCTCATAGATCTTCTGTTTGCACACTGGTTTCGATCCAGTTACTAGAGGTTTGTTCCTGAGGAT
AAGAGATTGCAATTGAAGGATTACTTCATCAAGCAAATCAGAGCTTACATGGATGGTAAAATTCAGTTGCTA
CTGTTTCTAACGATTCTTCTGCTTCTTCTAATACTGTTAAACCTGTTGCTTCTGCTTGAAGAGAAACAAGTATG
35 GTACTTATTACATGGAAGAGTCTGCTAGATTCACTAACGGTAATCAACCAATTACTGTTAGAAAGGTTGGTCCA
TTTTTGATTGCTCTGTTGGTTACCAATTCACCTGGTGGTTATTGTGATTACACTGAAGTTATGTTGCAAGAT
GGTCATGTTTGGGTGGTTACACTTGGGAGGGTCAAAGATATTACTTGCCAATTAGAACTTGAACGAGTTCTG
CTCCACCTAATCAAATTTGGGAGATTGTGGGGTGAAATTTCTTAA

40 >L0479_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSdleGDFDVAVLpFSNSTNNGLLfINTTIAIAAKEEGV
SLEKREAEahhhhhHLEVLfQGPMQAKLTKKEFIEWLKTSEGKQFNVDLWYGFQCFDYANAGWVKVLfGLLLKGLG
AKDIPFANNFDGLATVYQNTPDfLAKPGDMVVFgSNYGAGYGHVAWVIEATLDYIIVYEQNWLGgGWTDRIEQP

GWGWEKVTRRQHALRFYVYVPSYPKSETAPRSIQSPTQASKKETAKPQPKAVELKIIKDVVKGVDLPKRGGNPKGI
VIHNDAGSKGATAEAYRNLVNPAPLSRLEAGIAHSYVSGNTVWQALDESQVGWHTANQLGNKYVYIEVCQSM
GADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEFTSTSCPHRSSVLHTGFDPVTRGLLPEDKRLQLKDYFIK
QIRAYMDGKIPVATVSNESASSNTVKPVASAWKRNYGTYMEESARFTNGNQPIVRKVGPFLLSCPVGYYQFQP
5 GGYCDYTEVMLQDGHVWVGTYWEGQRYLPIRTWNGSAPPNQILGDLWGEIS*

>L0479_DNA

ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTACTAC
TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTACTCTGATTTGGAGGGAGATTTTGATGTTG
10 CTGTTTTGCCATTCTCTAACTCTACTAACAACGGTTTGTGTTTATTAACACTACTATCGCTTCTATTGCTGCTAA
GGAAGAGGGTGTTCCTTTGGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAGGTTTTGTTC
CAAGGCTCTATGCAGGCTAAGTTGACTAAGAAAGAGTTTATTGAGTGGTTGAAGACTTCTGAAGGTAAACAAT
TCAACGTTGATTGTGGTACGGTTTTCAATGTTTCGATTATGCTAATGCCGGTTGGAAAGTTTTGTTGGTTGT
TGTTGAAGGGTTTGGGTGCTAAGGATATCCCTTCGCTAACAACCTCGATGGTTTGGCTACTGTTTACCAAAAC
15 ACTCCAGATTTCTTGGCTAAGCCTGGAGATATGGTTGTTTTCGGTTCTAATTACGGTGCTGGTTATGGTCATGTT
GCTTGGGTATTGAAGCTACTTTGGATTACATCATCGTTTACGAGCAAAATTGGTTGGGTGGTGGTTGGACTG
ATAGAATTGAACAACAGGTTGGGGTTGGGAGAAGGTTACTAGAAGACAACACGCTTTGAGATTCCCTTATGT
TGTTTACCCATCTTATCCTAAATCTGAACTGCTCCAAGATCTATTCAATCTCCTACTCAAGCTTCTAAGAAAGA
AACTGCTAAGCCACAACCTAAAGCTGTTGAGTTGAAGATTATTAAGGATGTTGTTAAGGGTTACGATTTGCCAA
20 AGAGAGGTGGTAACCCTAAGGGTATCGTTATCCATAACGATGCTGGTTCTAAAGGTGCTACTGCTGAAGCTTA
CAGAAACGGTTTGGTTAATGCTCCATTGTCAGATTGGAAGCTGGTATTGCTCATTCTTATGTTTCTGGTAACAC
TGTTTGGCAAGCTTTGGATGAATCTCAAGTTGGTTGGCACACTGCTAACCAATTGGGTAATAAGTACTATTACG
GTATTGAGGTTTGTCAATCTATGGGTGCTGATAACGCTACTTCTTGAAGAACGAACAAGCTACTTTCCAAGAG
TGTGCTAGATTGTTGAAGAAATGGGGTTGCCAGCTAACAGAAACACTATCAGATTGCACAACGAGTTTACTTC
25 TACTTCTTGTCTCATAGATCTTCTGTTTTGCACACTGGTTTCGATCCAGTTACTAGAGGTTTGTGCTGAGGA
TAAGAGATTGCAATTGAAGGATTACTTCATCAAGCAATCAGAGCTTACATGGATGGTAAATCCCAGTTGCT
ACTGTTTCTAACGAATCTTCTGCTTCTTCTAATACTGTTAAACCTGTTGCTTCTGCTTGGAGAGAAACAAATAC
GGTACTTACTACATGGAAGAGTCTGCTAGATTCACTAACGGTAATCAACCAATTACTGTTAGAAAGGTTGGTCC
ATTTTTGCTTGTCTGTTGGTTACCAATCCAACCTGGTGGTACTGTGATTATACTGAAGTTATGTTGCAAGA
30 TGGTCATGTTTGGGTGGTTATACTGGGAGGGTCAAAGATATTACTTGCCAATTAGAACTGGAACGGTCTG
CTCCACCTAATCAAATTTGGGAGATTTGTGGGGTGAAATTTCTTAA

>L0480_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTNNGLLFINTTIIASIAAKEEGV
35 SLEKREAEAHHHHHHLEVLFGQPMMAKTQAEINKRLDAYAKGTVDSPYRVKKATSYDPSFGVMEAGAIADAGYYHA
QCQDLITDYVLWLTDNKVRTWGNKADQIKQSYGTGFKIHENKPSTVPKKGWIAVFTSGSYEQWGHIGIVYDGGNT
STFTILEQNWNGYANKKPTKRVNDNYGLTHFIEIPVKAGTTVKKETAKKSASKTPAPKKKATLKVSKNHINYTMDDR
GKKPEGMVIHNDAGRSSGQYENSLANAGYARYANGIAHYGSEGYVWEAIDAKNQIAWHTGDGTGANSNGNFR
FAGIEVCQSMSASDAQFLKNEQAVFQFTAEEKFEWGLTPNRKTVRLHMEFVPTACPHRSMVLHTGFNPVTQGRP
40 SQAIMNKLKDIFYKIQIKNYMDKGTSSSTVVKDGKTSSASTPATRPVTSWKKNQYGTWYKPENATFVNGNQPIVT
RIGSPFLNAPVGGNLPAGATIVYDEVCIQAGHIWIGYNAYNGNRVYCPVRTCQGVPPNQIPGVAWGVFK*

>L0480_DNA

ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTACTAC
TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTACTCTGATTTGGAGGGAGATTTTGATGTTG
CTGTTTTGCCATTCTCTAACTCTACTAACAACGGTTTGTTGTTTATTAACACTACTATCGCTTCTATTGCTGCTAA
GGAAGAGGGTGTTCCTTTGGAGAAAAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAAAGTTTTGTTT
5 CAAGGTCCTATGGCTAAGACTCAAGCTGAGATTAATAAGAGATTGGATGCTTACGCTAAGGGTACTGTTGATT
CTCCATACAGAGTTAAGAAAGCTACTTCTTATGATCCTTCTTTCGGTGTTATGGAAGCTGGTGCTATTGATGCTG
ATGGTTACTATCATGCTCAATGTCAAGATTTGATCACTGATTACGTTTTGTGGTTGACTGATAACAAGGTTAGA
ACTTGGGGTAACGCTAAGGATCAAATTAAGCAATCTTACGGTACTGGTTTTAAATCCATGAAAAACAAGCCATC
TACTGTTCTAAGAAAGGTTGGATTGCTGTTTTCACTTCTGGTTCTTACGAGCAATGGGGTCACATTGGTATTGT
10 TTACGATGGTGGTAACACTTCTACTTTTACTATTTTGAACAAAACCTGGAATGGTTACGCTAACAAGAAACCAA
CTAAGAGAGTTGATAACTACTACGGTTTGACTCACTTCATTGAAATTCCTGTTAAGGCTGGTACTACTGTTAAG
AAAGAGACTGCTAAGAAATCTGCTTCTAAAACCTCCAGCTCCTAAGAAAAAGGCTACTTTGAAAGTTTCTAAGAA
CCATATCAACTACACTATGGATAAGAGAGGTAAAAAGCCAGAGGGTATGGTTATTACAACGATGCTGGTAGA
TCTTCTGGTCAACAATACGAGAACTCTTGGCTAATGCTGGTTACGCTAGATATGCTAACGGTATTGCTCATTAC
15 TATGGTTCTGAAGGTTATGTTTGGGAGGCTATTGATGCTAAAAATCAAATTGCTTGGCACACTGGAGATGGTA
CTGGTGCTAACTCTGGTAATTTTAGATTCTGCTGGTATTGAAGTTGTCAATCTATGTCTGCTTCTGATGCTCAAT
TTTTGAAGAACGAGCAAGCTGTTTTCCAATTCAGTCTGAAAAATTCAGGAGTGGGGTTTGACTCCTAACAGA
AAGACTGTTAGATTGCATATGGAATTTGTTCCAACCTGCTTGTCTCATAGATCTATGGTTTTGCACACTGGTTTC
AATCCAGTTACTCAAGGTAGACCTTCTCAAGCTATCATGAACAAATTGAAGGATTACTTCATCAAGCAAATTA
20 GAATTACATGGATAAGGGTACTTCTTCTTCTACTGTTGTTAAAGATGGTAAACTTCTTCTGCTTCTACTCCAGC
TACTAGACCTGTTACTGGTTCTTGGAAAAAGAACCAATACGGTACTTGGTATAAACCAGAGAATGCTACTTTTG
TTAACGGTAATCAACCTATTGTTACTAGAATTGGTTCTCATTCTTGAACGCTCCTGTTGGTGGTAATTTGCCAG
CTGGTGCTACTATCGTTTACGATGAAGTTTGTATCCAAGCTGGTCATATTTGGATTGGTTACAACGCTTATAAC
GGTAATAGAGTTTATTGTCCAGTTAGAACTGTCAAGGTGTTCCACCTAATCAAATTCCTGGTGTGCTTGGGG
25 TGTTCCTCAAGTAA

>L0482_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSLEGLDFDVAVLPSNSTNNGLLFINTTASIAAKEEGV
SLEKREAEAHHHHHHLEVLFGQPREAPKTQIYTVKKGDTLSAIALKYKTTVSNIQNTNNIANPNLIFIGQKLKVPMT
30 LVEPKPKTVSSNNKSNNSSTLNYLKTLENRGWDFDGSYGWQCFDLNVVYWNHLYGHGLKGYGAKDIPYANNFN
SEAKIYHNTPTFKAEPGDLVVFSGRFGGGYGHTAIVLNGDYDGKLMKFQSLDQNWNNGGWRKAIEVAHKVHNY
ENDMIFIRPFKKA*

>L0482_DNA

ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTACTAC
TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTACTCTGATTTGGAGGGAGATTTTGATGTTG
CTGTTTTGCCTTTCTCTAACTCTACTAACAACGGTTTGTTGTTTATTAACACTACTATCGCTTCTATTGCTGCTAA
GGAAGAGGGTGTTCCTTTGGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAAAGTTTTGTTT
CAAGGTCCAAGAGAGGCTCCTAAGACTCAAATCTACACTGTTAAGAAAGGAGATACTTTGTCTGCTATCGCTTT
35 GAAGTACAAGACTACTGTTTCTAACATCCAAAACACTAACAACATCGCTAACCCTAAATTTGATTTTCATTGGTCA
AAAGTTGAAAGTTCCAATGACTCCTTTGGTTGAACCAAAGCCTAAGACTGTTTCTTCTAACAATAAGTCTAACT
TAATCTTCTACTTTGAAGTACTTGAAGACTTTGGAGAATAGAGTTGGGATTTTGATGGTTCTTATGGTTGGC
AATGTTTCGATTTGGTTAACGTTTACTGGAATCATTTGTATGGTCACGGTTTGAAGGGTTACGGTGCTAAGGAT
ATCCCTTACGCTAACAACCTCAACTCTGAAGCTAAGATCTACCATAATACTCCAACCTTCAAAGCTGAGCCTGGA
40

GATTTGGTTGTTTTCTCTGGTAGATTCCGGTGGTGGTTACGGTCACACTGCTATTGTTTTGAACGGAGATTACGA
TGGTAAATTGATGAAGTTCCAATCTTTGGATCAAAATTGGAACAATGGTGGTTGGAGAAAGGCTGAAGTTGCT
CATAAGGTTGTTCACAACTACGAGAACGATATGATTTTCATTAGACCATTCAAGAAAGCTTAA

5 >L0483_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVL PFSNSTNNGLLFINTTIIASIAAKEEGV
SLEKREAEAHHHHHHLEVL FQGPAATHEHSAQWLNNYKKGYGYPYPLGINGGMHYGVDFMNI GTPVKAIS SG
KIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVVKVDYVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFS
NSTAQDPM PFLKSAGYGKAGGTVPTPTNTGWKTNKYGTLYKSESASFPTNTDIITRTTGPF RSMPQSGVLKAGQTI
10 HYDEV MKQDGHVWVG YTGNSGQRIYLPVRTWNKSTNTLGV LWGTIK*

>L0483_DNA

ATGAGATTCCATCTATTTTCACTGCTGTTTTGTTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTACTAC
TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTACTCTGATTTGGAGGGAGATTTTGATGTTG
15 CTGTTTTGCCTTTCTCTAACTCTACTAACAACGGTTTGTTGTTTATTAACACTACTATCGCTTCTATTGCTGCTAA
GGAAGAGGGTGTTTCTTTGGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAAGTTTTGTTT
CAAGGTCCAGCTGCTACTCATGAGCACTCTGCTCAATGGTTGAACAATTACAAGAAAGGTTACGGTTATGGTCC
ATATCCTTTGGGTATTAACGGTGGTATGCATTACGGTGTTGATTTCTTTATGAACATCGGTACTCCTGTTAAGGC
TATTTCTTCTGGTAAAAATTGTTGAAGCTGGTTGGTCTAACTATGGTGGTGGTAATCAAATTGGTTTGATTGAGA
20 ACGATGGTGTT CATAGACAATGGTACATGCATTGTCTAAGTACAACGTTAAGGTTGGAGATTACGTTAAGGC
TGGTCAAATTATTGGTTGGTCTGTTCTACTGGTTATTCTACTGCTCCACATTTGCACTTCCAAAGAATGGTTAA
CTCTTTCTTAATTCTACTGCTCAAGATCCAATGCCTTTCTTGAAGTCTGCTGGTTACGGTAAAGCTGGTGGTAC
TGTTACTCCAACCTCTAACAACCTGGTTGGAAGACTAACAAGTACGGTACTTTGTACAAGTCTGAATCTGCTTCTT
TACTCCAAACACTGATATCATCACTAGAACTACTGGTCCATTAGATCTATGCCTCAATCTGGTGTGTTTGAAGGC
25 TGGTCAAAC TATCCATTACGATGAGGTTATGAAGCAAGATGGTCACGTTTGGGTTGGTTACTGGTAAATTCTG
GTCAAAGAATCTACTGCTGTAGAACTTGAACAAGTCTACTAATACTTTGGGTGTTTTGTGGGGTACTATT
AAATAA

>L0484_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVL PFSNSTNNGLLFINTTIIASIAAKEEGV
30 SLEKREAEAHHHHHHLEVL FQGP DQFLDEKEGNHTTAYRDGSGIWTICRGATMVDGKPVIPGMRLSKEKCDQVN
AIERDKALAWVERN IKVPLTEPQKAGIASFCPYNIGPGKCFPSTFYKRLNAGDRKGACEAIRWWIKDGGDRDCRIRSN
NCYGQVIRRDQESALACWGIDQ*

35 >L0484_DNA

ATGAGATTTCCTTCTATTTTCACTGCTGTTTTGTTTGCTGCTTCTTCTGCTTTGGCTGCTCCTGTTAACTACTACTAC
TGAAGATGAGACTGCTCAAATTCAGCTGAAGCTGTTATTGGTTACTCTGATTTGGAGGGAGATTTTGATGTTG
CTGTTTTGCCTTTCTCTAACTCTACTAACAACGGTTTGTTGTTTATTAACTACTATCGCTTCTATCGCTGCTAA
GGAAGAGGGTGTTTCTTTGGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAAGTTTTGTTT
40 CAAGGTCCAGATCAATTCTTGGATGAAAAAGAGGGTAATCATACTACTGCTTATAGAGATGGTTCTGGTATTTG
GACTATTTGTAGAGGTGCTACTATGGTTGATGGTAAACCAGTTATTCTGGTATGAGATTGTCTAAGGAAAAGT

GTGATCAAGTTAACGCTATTGAGAGAGATAAGGCTTTGGCTTGGGTTGAAAGAAATATTAAGGTTCCATTGAC
TGAGCCTCAAAAAGCTGGTATTGCTTCTTTTGTCCATACAACATTGGTCCTGGTAAATGTTTCCCATCTACTTTC
TACAAGAGATTGAATGCTGGAGATAGAAAGGGTGCTTGTGAAGCTATTAGATGGTGGATTAAGGATGGTGGT
AGAGATTGTAGAATTAGATCTAACAACTGTTACGGTCAAGTTATTAGAAGAGATCAAGAATCTGCTTTGGCTTG
5 TTGGGGTATTGATCAATAA

>L0487_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYS DLEGDFDVA VLPFSNSTNNGLLFINTTASIAAKEEGV
SLEKREAEAHHHHHHLEVL FQGPMRYIPAAHHSAGSNKPVNRVVIHATCPDVGFPSASRKGRAVSTANYFASPSSG
10 GSAHYVCDIGETVQCLSESTIGWHAPPNPHSLGIEICADGGSHASFRVPGHAYTREQWLDPRVWPAVERAAVLCR
QLCDKHGVPKRKLSVADLKAGRRGICGHTDVTDAWHQSDHDDPGPWFPWDFMMAVNVNGHGGGSSSEELTVAD
VKALHNQIKLSAQLSGSVNKLHHDVGVVQVQNGDLSKRVDALSWVKNPVTGKLWRTKDALWSVWYVYVLECRS
RIDRLES AVNGLKK*

15 >L0487_DNA

ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTAC
TGAAGATGAGACTGCTCAAATCCTGCTGAAGCTGTTATTGGTTACTCTGATTGGAGGGAGATTTTGATGTTG
CTGTTTTGCCTTTCTCTAACTCTACTAACAACGGTTTGTTGTTTATTAACACTACTATCGCTTCTATTGCTGCTAA
GGAAGAGGGTGTTTCTTTGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAGGTTTTGTTT
20 CAAGGTCCAATGAGATATATTCCTGCTGCTCATCACTCTGCTGGTTCTAACAAGCCAGTTAATAGAGTTGTTATT
CATGCTACTTGTCCAGATGTTGGTTTTCTTCTGCTTCCAGAAAAGGTAGAGCTGTTTCTACTGCTAACTACTTC
GCTTCTCCTTCTTCTGGTGTTCTGCTCATTATGTTTGTGATATTGGTGAACTGTTCAATGTTTGTCTGAGTCTA
CTATTGGTTGGCATGCTCCACCTAATCCACACTCTTTGGGTATTGAAATTTGTGCTGATGGTGGTTCTCATGCTT
CTTTCAGAGTTCTGGTCACGCTTACACTAGAGAACAATGGTTGGATCCAAGAGTTTGGCCTGCTGTTGAGAG
25 AGCTGCTGTTTTGTGTAGACAATTGTGTGATAAGCATGGTGTTCAAAGAGAAAATTGTCTGTTGCTGATTGA
AAGCTGGTAGAAGAGGTATTTGTGGTCACACTGATGTTACTGATGCTTGGCATCAATCTGATCAGCATGATCCA
GGTCTTGGTTTCTTGGGATAAGTTCATGGCTGTTGTTAACGGTCATGGTGGTGGTTCTTCTTCTGAAGAGTT
GACTGTTGCTGATGTTAAGGCTTTCACAACCAAAATTAACAATTGTCTGCTCAATTGTCTGGTCCGTTAAACAA
ATTGCATCAGATGTTGGTGTGTTCAAGTTCAAAATGGAGATTTGTCTAAGAGAGTTGATGCTTTGTCTTGGG
30 TTA AAAACCCAGTTACTGGTAAATTGTGGAGAACTAAAGATGCTTTGTGGTCTGTTGGTACTATGTTTTGGA
TGTAGATCCAGAATTGATAGATTGGAGTCTGCTGTTAATGGTTTGAAGAAATAA

>L0489_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYS DLEGDFDVA VLPFSNSTNNGLLFINTTASIAAKEEGV
35 SLEKREAEAHHHHHHLEVL FQGPMMLMTKNQAEKWFNDSL GKQFNPDGWYGFQCYDYANMFFMLATGERLQGL
YAYNIPFDNKAKIEKYGQIIKNYDSFLPQKLDIVVFPSKYGGGAGHVEIVESANLNTFTSFGQNWNGKGWTVNGVAQ
PGWGPETVTRHVHYDNP MYFIRLNFPNNLSVGNKAKGIKQATTKEAVIKPKIMLVAGHGYNDPGAVGNGTN
ERDFIRKYITPNI AKYLRHAGHEVALYGGSSQSQDMYQDTAYGVNVGNKKDYGLYWVKSQGYDIVLEIHLDAAGES
ASGGHVISSQFNADTIDKSIQDVIKNNLGQIRGVTPRNDLLNVNVS AEININYRLSELGFITNKNDMDWIKKNYDLY
40 SKLIAGAIHGKPIGGLVAGNVKTS AKNKKNPVPAGYTL DKNNVYPYKEQGN YTVANVKGNVNRDGYSTNSRITGV
LPNNTTITYD GAYCINGYRWITYIANSQRRYIATGEVDKAGNRIS SFKFTI*

>L0489_DNA

ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTA
CTGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTACTCTGATTTGGAGGGAGATTTTGATGTT
GCTGTTTTGCCTTTCTCTAACTCTACTAACAACGGTTTGTTGTTTCATTAACACTACTATCGCTTCTATCGCTGCTA
5 AGGAAGAGGGTGTTTTCTTTGGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAGGTTTTGTT
TCAAGGTCCAATGTTGATGACTAAGAACCAAGCTGAAAAATGGTTCGATAACTCTTTGGGTAAACAATTCAATC
CTGATGGTTGGTACGGTTTCCAATGTTACGATTACGCTAACATGTTTTTCATGTTGGCTACTGGTGAAAGATTG
CAAGGTTTGATCGCTTACAACATCCCATTTCGATAATAAGGCTAAGATCGAGAAGTACGGTCAAATTATTAAGAA
CTACGATTCCTTCTTCCACAAAAGTTGGATATCGTTGTTTTCCCTTCTAAATACGGTGGTGGTGCTGGTCATGT
10 TGAAATTGTTGAGTCTGCTAACTTGAATACTTTTACTTCTTTCCGGTCAAACTGGAATGGTAAAGGTTGGACTA
ACGGTGTTGCTCAACCAGGTTGGGGTCTGAAACTGTTACTAGACATGTTCACTACTACGATAACCCAATGTAC
TTCATCAGATTGAATTTCCCTAACAATTTGTCTGTTGGTAATAAGGCTAAGGGTATTATTAAGCAAGCTACTACT
AAGAAAGAGGGCTGTTATTAACCAAAGAAAATTATGTTGGTTGCTGGTCACGGTTATAACGATCCTGGTGCTG
TTGGTAACGGTACTAACGAAAGAGATTTTCATTAGAAAGTACATCACTCCTAACATCGCTAAGTATTTGAGACAT
15 GCTGGTCACGAGGTTGCTTGTACGGTGGTCTTCTCAATCTCAAGATATGTACCAAGATACTGCTTATGGTGT
TAACGTTGGTAATAAGAAAGATTACGGTTTGTACTGGGTTAAGTCTCAAGGTTACGATATCGTTTTGGAAATCC
ATTTGGATGCTGCTGGTGAATCTGCTTCTGGTGGTCACGTTATCATCTCTTCTCAATTCAACGCTGATACTATCG
ATAAGTCTATCCAAGATGTTATTAACCAATTTGGGTCAAATTAGAGGTGTTACTCCAAGAAACGATTTGTTG
AACGTTAACGTTTCTGCTGAAATTAACATCAACTACAGATTGCTGAGTTGGGTTTTATTACTAACAAAAACGAT
20 ATGGATTGGATTAAGAAAACTACGATTTGTATTCTAAGTTGATTGCTGGTGCTATTCATGGTAAACCTATTGG
TGGTTTGGTTGCTGGTAACGTTAAGACTTCTGCTAAAAACAAGAAAAATCCACCTGTTCCAGCTGGTTACACTT
TGGATAAGAACAATGTTCTTACAAGAAAGAACAGGGTAACACTGTTGCTAACGTTAAGGGTAACAACGT
TAGAGATGGTTACTCTACTAACTCCAGAATTACTGGTGTGTTTGGCAAACAACACTACTATCACTTACGATGGTGC
TTACTGTATCAACGGTTACAGATGGATCACTTACATCGCTAATTCTGGTCAAAGAAGATATATTGCTACTGGTG
25 AAGTTGATAAGGCTGGTAATAGAATTTCTTCTTTGGTAAATTCTCTACTATCTAA

>L0499_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSLEGDVAVLPFSNSTNNGLLFINTTASIAAKEEGV
SLEKREAEAHHHHHHLEVLFGQPTQNYVVKKG DYLGAIKKFNNTVSKIQLNNIKNPNFILIGQVLKVPFISSNNSST
30 NTNRRNNSSTLNYLNSLAYKGWDFD GAYGNQCFDLNVVYWNHLYGHGLRGYGAKDIPFANNFRGEATVYRNT
TFKAELGDVAVFSGRYGGGFGHTAIVLNGNYDGKLMKFQSLDQNWYGGDSRRADIAHKVVHNYDYDMYFIRPYK
KA*

>L0499_DNA

ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTAC
TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTATTCTGATTTGGAGGGAGATTTTGATGTTG
CTGTTTTGCCATTCTCTAACTCTACTAACAACGGTTTGTTGTTTCATTAACACTACTATCGCTTCTATCGCTGCTAA
GGAAGAGGGTGTTTTCTTTGGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAGGTTTTGTTT
CAAGGTCCTACTCAAAACTACGTTGTTAAGAAAGGAGATTATTTGGGTGCTATCGCTAAGAAATTCAATACTAC
40 TGTTTCTAAGATCCAAAAGTTGAACAACATTAAGAACCCTCACTTCTGATCGGTCAAGTTTTGAAAGTTCC
TTTCATTTCTTCTAACAACCTTCTACTAACAATAACAAGAACAACCTTCTTCTACTTTGAACTACTTGAAC
TCTTTGGCTTATAAGGGTTGGGATTTTGATGGTGCTTACGGTAACCAATGTTTCGATTTGGTTAACGTTTACTG
GAATCATTTGTATGGTCACGGTTTGAGAGGTTATGGTGCTAAGGATATTCATTTGCTAACAATTCAGAGGTG
AAGCTACTGTTTACAGAAACACTCCTACTTTTAAAGCTGAGTTGGGAGATGTTGCTGTTTTCTCTGGTAGATAC

GGTGGTGGTTTCGGTCATACTGCTATCGTTTTGAACGGTAACTACGATGGTAAATTGATGAAGTTCCAATCTTT
 GGATCAAACTGGTACGGTGGAGATTCTCGTAGAGCTGATATCGCTCATAAGGTTGTTCACTACGATTAC
 GATATGTACTTCATCAGACCTTACAAGAAAGCTTAA

5 >L0502_Protein

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLGDFDVAVL PFSNSTNNGLLFINTTIASIAAKEEGV
 SLEKREAEAHHHHHHLEVL FQGPMNKVDAWLSKHDLKNLTLLIVVGAFILFLLFMLMNHSDQGSNRITVSENAEL
 RTDPNAAYPVIYQVEKGDFTFLRSKSGKWIEVESRDGSEKGIAGWHTSLNIEADSNKKVNPLKNKVIVLDPGHGG
 NDQGASSNTSKSLEKNYLTAKELRRALEKAGATVKLTRSDDTYVSLDDRKLNGDAYISIHNSLSPNANGATV
 10 YWYKNNQEALAE LNTNIQKKAMLSNRGTRQENFQVLRQTNKPAVLLELGYISNPTDEV MIRNQIHRQVVEEAVV
 DGLKQYFAS*

>L0502_DNA

ATGAGATTTCCATCTATTTTCACTGCTGTTTTGTTGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACTACTAC
 15 TGAAGATGAGACTGCTCAAATTCCTGCTGAAGCTGTTATTGGTTACTCTGATTGGAGGGAGATTTTGATGTTG
 CTGTTTTGCCATTCTCTAACTCTACTAACAACGGTTTGTTGTTTATTAACACTACTATCGCTTCTATTGCTGCTAA
 GGAAGAGGGTGTTTCTTTGAAAAGAGAGAAGCTGAGGCTCATCACCATCACCATCACTTGGAGGTTTTGTTT
 CAAGGCTCATGAACAAGGTTGATGCTTGTTGTCTAAGCATGATTTGAAAAATTTGACTACTTTGATCGTTGT
 TGGTGCTTTTATTTTGTCTTGATCTTGTTGTTTATGTTGATGAACCACTCTGATCAAGGTTCTAACAGAATTACT
 20 GTTCTGAAAATGCTGAGTTGAGAACTGATCCAAATGCTGCTTACCCAGTTATCTACCAAGTTGAAAAGGGAG
 ATACTTTCACTAGATTGTCTAAGTCTGGTAAATGGATTGAAGTTGAGTCCAGAGATGGTTCTGAAAAGGTTG
 GATTGCTGGTTGGCATACTTCTTTGAACATTGAGGCTGATTCTAACAAGAAAGTTAACCCATTGAAGAACAAG
 TTATTGTTTTGGATCCTGGTCACGGTGGTAACGATCAAGGTGCTTCTTCTAATACTTCTAAGAAGTCTTTGAAA
 AGAACTACACTTTGAAGACTGCTAAAGAATTGCGTAGAGCTTTGAAAAGGCTGGTGCTACTGTTAAGTTGAC
 25 TAGATCTGATGATACTTACGTTTCTTTGGATGATAGAAAGTTGAACGGAGATGCTTACATCTCTATCCATAACG
 ATTCTTTGGATTCTCCTAACGCTAATGGTGCTACTGTTTACTGGTACAAGAACAACCAAGAAGCTTTGGCTGAG
 ACTTTGAACACTAACATCCAAAAGAAAGCTATGTTGTCTAACAGAGGTACTAGACAAGAAAACCTTCCAAGTTTT
 GAGACAACTAACAGCCAGCTGTTTTGTTGAATTGGGTTACATTTCTAACCTACTGATGAGGTTATGATTA
 GAAATCAAATCCACAGACAAGTTGTTGAAGAGGCTGTTGTTGATGGTTTGAAACAATATTTGCTTCTTAA
 30

Example 3: Construct design and mRNA production

In vitro transcription of lysin encoding mRNAs is based on the pST1-T7-AGA-dEarl-hAg-MCS-
 FI-A30LA70 plasmid-backbone and derivative DNA-constructs. These plasmid constructs
 contain a 5' UTR (untranslated region, a derivate of the 5'-UTR of *homo sapiens* hemoglobin
 35 subunit alpha 1 (hAg)), a 3' FI element (where F is a 136 nucleotide long 3'-UTR fragment of
 amino-terminal enhancer of split, mRNA and I is a 142 nucleotide long fragment of
 mitochondrially encoded 12S RNA both identified in *Homo sapiens*; WO 2017/060314) and a
 poly(A) tail of 100 nucleotides, with a linker after 70 nucleotides.

Lysins with a predicted activity against *Staphylococcus aureus* were selected and the encoding sequences originate from *Staphylococcus* species or their phages/viruses (Table 4). The husec signal sequence (derived from Ig heavy chain V-I region HG3; UniProtKB entry HV102_HUMAN) as well as a hexa-histidine tag followed by a 3C protease cleavage site is added N-terminally. The signal sequence targets the lysin constructs to the secretory pathway and is cleaved off at the signal peptidase site after translocation into the ER. The hexa-histidine tag and 3C sites are introduced to facilitate the quantification of translated lysin and its purification by affinity capture and removal of the hexa-histidine tag after proteolytic cleavage, respectively. The lysin sequences are codon-optimized for expression in *Pichia pastoris* or *Homo sapiens*.

Table 4: Origin of lysin sequences.

Domain architecture	Internal name	NCBI ID	Species of origin
Ami2	L0487	WP_115921155	<i>Staphylococcus epidermidis</i>
CHAP-Ami2-SH3	L0489	ABF71642	<i>Staphylococcus virus 80alpha</i>
	L0479	AWH13065	<i>Staphylococcus phage trsa10</i>
	L0480	YP_024461	<i>Staphylococcus phage K</i>
	L0473	WP_195722665	<i>Staphylococcus kloosii</i>
	L0475	YP_009268702	<i>Staphylococcus virus 80</i>
LysM-CHAP	L0499	WP_070648624	<i>Staphylococcus</i>
	L0482	WP_143119924	<i>Staphylococcus aureus</i>
PepM23-SH3	L0483	Lysostaphin	<i>Staphylococcus simulans</i>
		WP_013012297	
Phage lysozyme	L0484	PTF44644	<i>Staphylococcus epidermidis</i>
SH3-Ami3	L0502	WP_053025743	<i>Staphylococcus haemolyticus</i>

mRNA is generated by *in vitro* transcription as described by Kreiter et al. (Kreiter, S. et al. Cancer Immunol. Immunother. 56, 1577–87 (2007)) with substitution of the normal nucleoside uridine by 1-methyl-pseudouridine. Resulting mRNAs are equipped with a Cap1-structure and double-stranded (dsRNA) molecules is depleted. Purified mRNA is eluted in H₂O and stored at -80 °C until further use. *In vitro* transcription of all described mRNA constructs

is carried out at BioNTech SE. A list of all lysin constructs, which are used in subsequent experiments is shown in Table 5.

Table 5: Nucleotide and amino acid sequences of mRNA encoded lysins.

L0473	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGQPMKTKKQIEQRLQAYAKGT DSPYRITKPTSYDPSFGVMEVGAIADRYYHAQCQDLITDYVLWLTDNKRVPWGNKQDI NQSYGTGFKIHKNLSTVAKTGWIAVFTSGTYSQYGHIGIVYSPGDTNKFHIEQNWNGWA NKKPLTRWDNYYGLTHFIEVPYAKETITTKTAPKAATSKPAKKISVSHNNINLQSNFAKRGY NPKGVIHNDAGRSTAQQYENSLKGANQNLNNGIAHAYADSSHIWYAIEPERIAYHTGD GVKKGTGNYDYFGIEVCQSMSASDKEFMANEQVVLQEVARLLKKWKL PANRNTVRLHLEF SATSCPHRSLTMRGTGFPVTQGRPSDAIIKKLKDIWIKIRAYMEGKTPSIKTIVANKPGSAST PATKSNSDGWKVNAYNTYYKTERASFTVGSTRIAVRSTGPFTTCPTGWLNPGVITITYDEV MVQNGFVWLGYDSYNGRRYVPIRTAQGTPGSSNYVLGPLWGTIS*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGATTGGACGT GGCGAGTTTTCTGTCTGCTGGCTGTTGCCCCAGGCGCGCATTCTATGAAGCATCACCAC ATCACCCTCAGCAGGGCTCGAAGTCCTGTTCCAAGGACCAATGGCCAAAACCAAGAA/ CAGATAGAACAGAGACTGCAAGCTTACGCAAAGGGGACCGTGGATTCCCTTACCGAA TCACCAAGCCAACCAGCTACGACCCCTCCTTCGGTGTCTATGGAGGTTGGCGCTATAGAT GCGGATCGCTACTATCATGCCAGTGTCAGATCTGATAACTGACTATGTACTTTGGCTC ACTGACAACAAAGTAAGACCTTGGGGTAATGCCAAGGACCAGATTAACCAGAGCTATG GGACTGGTTTCAAAATCCACAAGAATCTGCCAGTACCGTGGCTAAAACAGGATGGATT GCCGTGTTTACATCCGGGACCTACTCACAGTACGGCCACATCGGCATCGTGTACAGTCC AGGAGATACCAACAAATTCCATATCATCGAGCAGAATTGGAATGGATGGGCAAACAAG AAACCTCTGACACGTTGGGACAACCTACTACGGGCTGACCCACTTTATCGAGGTCCCAT TGCGAAAGAGACAATCACAACGAAGAAAACCTGCACCTAAAGCAGCTACCTCAAAACCG GCTAAGAAGATCTCAGTTTCTCACAACAACATCAACCTTCAGAGCAACTTTGCAAAGAG AGGGTACAACCCAAAGGGAGTGGTGATTACAATGACGCCGGACGCTCTACTGCGCAA CAGTATGAAAACAGCCTGAAGGGCGCTAATCAGAACAGGCTGAATAACGGCATTGCC ATGCATATGCTGACAGCAGCCACATTTGGTATGCTATCGAACCCGAGAGAATCGCCTAT CACACAGGCGATGGCGTTAAGAAGGGGAACAGGCAATTACGACTACTTCGGTATTGAAG TATGCCAGAGCATGAGCGCAAGTGACAAAGAGTTTATGGCCAACGAACAGGTGGTACT GCAGGAGGTGGCTAGGTTGCTCAAGAAATGGAAGCTTCCCGCCAATCGGAATACCGTG CGGCTCCATCTGGAATTCTCTGCCACATCTTGCCCTCATAGGAGCCTCACAATGCGGACC GGTTTCAATCCGGTGACTCAGGGTAGACCCTCAGACGCCATCATCAAGAAGTTGAAGG/ TATCTGGATAAAAAGAGATTCGCGCATATATGGAGGGTAAAACCCCATCCATTAAGACTG TCGCCAATAAGCCCGGATCAGCCAGTACCCCTGCAACCAAATCCAACCTCCGATGGGTGG AAGGTCAATGCCTACAATACCTACTACAAAACGGAACGGGCCTCCTTTACAGTGGGCTC TACTCGGATTGCTGTGCGTAGTACAGGACCCTTTACTACATGCCCGTTTACAGGGTGGC GAACCCCGGAGTCACGATAACCTATGACGAGGTTATGGTGCAGAACGGGTTTGTGTGG TTGGGCTATGATAGCTACAATGGGAGGAGGTATGTCCCATTCGCACTGCCAAGGAA TCCTGGCAGCTCCAATTATGTGCTTGGCCCTCTGTGGGGCAGATTTCTTGActcgagCTG GTACTGCATGCACGAATGCTAGCTGCCCTTTCCCGTCTGGGTACCCGAGTCTCCCC CGACCTCGGGTCCAGGTATGCTCCACCTCCACCTGCCCACTACACACCTCTGCTAGT TCCAGACACCTCCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCC CCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATAC

	TAACCCAGGGTTGGTCAATTTCTGTGCCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAAAAA AA
L0475	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGGPMQAKLTKEFIEWLKTSEK QFNADLWYGFQCFDYANAAWKVLFGLLKGLGAKDIPFANNFDGLATVYQNTPDFLAQPG DMVVFGSNYGAGYGHVAWVIEATLDYIIVYEQNWLGSGWTDGIEQPGWGWKEVTRRQ HAYDFPMWFIRPNFKSEIAPRSVQSPTQAPKKETAKPQPKAVELKIKDVVKGYDLPKGRSN PKGIVIHNDAGSKGATAEAYRNLVNAPLSRLEAGIAHSYVSGNTVWQALDESQVGWHTA NQLGNKYYYGIEVCQSMGADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEFTSTS CPHRSSVLHTGFDPVTRGLLPEDKRLQLKDYFIKQIRAYMDGKIPVATVNDSSASSNTVKPV ASAWKRNKYGTYYMEESARFTNGNQIPITVRKVGPFYCPVGYQFQPGGYCDYTEVMLQD GHVWVGYTWEGQRYLPIRTWNGSAPPNQILGDLWGEIS*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGATTGGACAT GGCGAGTGTCTGCCTTCTTGTCTGTTGCACCTGGTGGCCATTCCATGAAACATCACCACC ATCACCATTAGCTGGCCTTGAAGTGCTTTCCAGGGACCTATGCAAGCGAAACTGACC AAGAAGGAGTTTATAGAGTGGCTGAAAACCTCTGAGGGTAAACAGTTCAACGCCGATC TGTGGTATGGTTTCCAGTGCTTTGATTACGCAAACGCTGCTTGAAAAGTGCTGTTTGCC TGCTCCTGAAGGGACTCGGTGCCAAGACATCCCCTTTGCCAACAATTCGACGGGCTG GCAACCGTCTACCAGAATACTCCTGACTTTCTCGCTCAGCCAGGGGACATGGTCGTGTT GGCTCCAATTATGGGGCAGGCTATGGACACGTAGCATGGGTGATAGAAGCCACTCTGG ACTACATAATCGTCTACGAGCAAAATTGGCTTGGCGGTGGGTGGACCGATGGCATTGA GCAGCCTGGCTGGGGATGGGAGAAGGTTACTCGTAGACAGCATGCTTACGACTTTCCC ATGTGGTTCATCCGTCCCACTTCAAGTCCGAGATTGCGCCAGAAGCGTCCAGTCTCCA ACCCAAGCTCCGAAGAAGGAAACAGCCAAGCCACAGCCTAAGGCCGTTGAACTGAAGA TTATTAAGGATGTGGTGAAGGGATATGACCTCCCGAAAAGAGGGAGCAATCCCAAAGG GATTGTGATCCACAACGATGCCGGGTCTAAAGGGGCAACAGCAGAAGCGTATAGGAAC GGGTTGGTGAATGCCCCTCTCAGTCGCTTGGAGGCGGGAATTGCACACTCATATGTGTC CGGAAATACGGTTTGGCAAGCTCTCGACGAAAGCCAAGTAGGCTGGCATAACAGCCAAC CAGCTGGGCAACAAATACTACTATGGCATCGAGGTGTGCCAGAGTATGGGCGCAGATA ATGCCACGTTCTGAAGAACGAGCAGGCTACCTTTCAAGAGTGTGCGCGGTTGCTGAA GAAGTGGGGACTGCCAGCTAATCGCAACACAATCAGGTTGCACAACGAGTTTACGAGT ACTAGCTGTCCACACAGAAGTAGCGTGCTGCACACAGGATTTCGATCCTGTGACAAGGG GACTGCTGCCCCGAAGATAAGCGGCTGCAGCTGAAAAGACTACTTCATCAACAGATTCCG GCCTATATGGACGGAAGATTCCCGTCGCTACTGTTAGCAATGACAGCTCTGCCTCATC AAATACCGTGAAACCGGTAGCAAGCGCCTGGAAAAGGAACAAGTATGGCACTTATTAC ATGGAGGAATCCGCCCCGATTACCAACGGTAATCAGCCCATCACAGTCCGCAAGGTCCG GCCATTCTGTACTGTCCCGTAGGCTACCAAGTTTCAGCCTGGTGGTACTGCGACTATAC CGAGGTGATGCTGCAAGATGGCCATGTCTGGGTTGGCTATACCTGGGAAGGGCAGCGC TACTACTTGCCAATACGGACTTGGAATGGTTCTGCCCCACCAACAGATCCTCGGAGAT CTGTGGGGCGAAATCTCCTGActcgagCTGGTACTGCATGCACGCAATGCTAGCTGCCCC TTTCCCGTCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCT CCACCTGCCCCACTACCACTCTGCTAGTTCCAGACACCTCCAAGCACGCAGCAATGC AGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTAG CAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTCGTGCCAG CCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACT AA AAAAAAAAAAAAAA

L0479	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGQPMQAKLTKEFIEWLKTSEKQFNVDLWYGFQCFDYANAGWKVLFGLLLKGLGAKDIPFANNFDGLATVYQNTPDFLAKPGDMVVFGSNYGAGYGHVAWVIEATLDYIIVYEQNLWGGGWTDRIEQPGWGWKEVTRRQHALRFPYVVYPSYPKSETAPRSIQSPTQASKKETAKPQPKAVELKIIDVVKGYDLPKRGGNPKGIVIHNDAGSKGATAEAYRNLVNAPLSRLEAGIAHSYVSGNTVWQALDESQVGWHTANQLGNKYYYGIEVCQSMGADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEFTSTSCPHRSSVLHTGFDPVTRGLLPEDKRLQLKDYFIKQIRAYMDGKIPVATVSNASSNTVKPVASAWKRNKYGTYYMEESARFTNGNQIPITVRKVGPFLLSCPVGYYQFQPGGYCDYTEV/MLQDGHVWVGTYWEGQRYLPIRTWNGSAPPNQILGDLWGEIS*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGACTGGACCTGGAGAGTCTTTGCCTCCTTGCACTTGCTCCTGGTGCCACAGTATGAAGCATCACCATCATCACCACAGTGCAGGACTGGAGGTGTTGTTTCAGGGACCAATGCAGGCCAAACTGACCAAGAAGGAGTTCATCGAATGGCTCAAGACTTCCGAGGGCAAGCAGTTCATGTGGACCTCTGGTATGGCTTTCAATGCTTTGATTATGCCAATGCTGGGTGGAAGGTCCTGTTGGCCTTCTCCTGAAAGGACTGGGTGCAAAGACATCCCTTCGCCAACAACTTCGATGGACTGGCTACCGTTTACCAGAATACCCCGGACTTTCTGGCCAAGCCTGGAGATATGGTGGTGTTCGGAAGCAACTATGGGGCAGGATATGGCCATGTAGCCTGGGTGATAGAAGCTACGTTGGACTACATCATCGTGTACGAGCAAAATTGGCTGGGCGGCGGCTGGACTGACCGAATTGAGCAACCTGGGTGGGGATGGGAGAAAGTGACCAGACGTCAGCACGCTCTGCGCTTCCCTATGTAGTGTACCCCTCTTACCCGAAGTCTGAGACAGCTCCCGGTCCATTGAGAGCCCAACACAGGCGAGCAAGAAAGAGACTGCCAAACCACAACCCAAAGCGGTTGAACTGAAGATCATCAAAGACGTCGTGAAAGGGTATGATCTGCCAAGAGAGGCGGAAATCCAAAGGGGATTGTGATCCACAATGACGCCGGGAGCAAAGGTGCAACTGCTGAAGCTTACCGCATGGGCTGGTCAACGCACCTTTGTCCGACTGGAAGCCGGCATTGCTCATAGCTATGTGTCTGGCAACACTGTTTGGCAAGCACTGGACGAATCACAGGTCGGCTGGCATACTGCCAATCAGCTCGGAAACAAGTACTACTACGGGATAGAGGTCTGTCACTCCATGGGCGCCGATAACGCGACTTTCCTGAAGAATGAGCAGGCCACATTCCAGGAATGCGCTCGGTTGTTGAAGAAATGGGGCCTTCTGCAAACAGGAACACAATCAGGCTGCACAATGAGTTCACATCCACGTCTTGCCACACAGGTCTAGCGTGCTGCACACAGGTTTTGATCCCGTGACGAGAGGGCTTCTCCCCGAAGATAAACGTCTGCAGCTCAAGGATTACTTCATCAAGCAGATAAGGGCCTATATGGACGGGAAAAATCCGGTAGCTACCGTCAGCAACGAGAGTAGCGCATCATCCAACACCGTTAAGCCAGTGGCCTCAGCCTGGAAGCGCAACAAATACGGTACCTATTACATGAGGAATCTGCAAGATTCACAAATGGTAATCAGCCTATTACCGTGCGGAAAGTCGGCCCCTTTCTCAGCTGTCTGTTGGATATCAGTTTCAGCCAGGCGGTTACTGCGACTATACCGAAGTAATGCTGCAAGATGGGCATGTGTGGGTTGGGTACACATGGGAAGGCCAACGGTACTATCTGCCCATTGCGACCTGGAACGGCAGTGCCCCACCCAACCAGATCCTTGGAGATCTGTGGGGTGAGATTTTCATGActcgagCTGGTACTGCATGCACGCAATGCTAGCTGCCCCCTTCCCGTCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTTCTGTCCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAA AAA AAAAAAAAA
L0480	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGQPMQAKTQAEINKRLDAYAKGTVDSPYRVKKATSYDPSFGVMEAGAIDADGYHAQCQDLITDYVLWLTDNKVRTWGNKADQIKQSYGTGFKIHENKPSTVPKKGWIAVFTSGSYEQWGHIGIVYDGGNTSTFTILEQNWNGYA

	NKKPTKRVDNYGLTHFIEIPVKAGTTVKKETAKKSASKTPAPKKKATLKVSKNHINYTMDKR GKKPEGMVIHNDAGRSSGQYENSLANAGYARYANGIAHYYGSEGYVWEAIDAKNQIAW HTGDGTGANSNFRFAGIEVCQSMSASDAQFLKNEQAVFQFTAIEKFKEWGLTPNRKTVRL HMEFVPTACPHRSMVLHTGFNPVTQGRPSQAIMNKLKDYFIKQIKNYMDKGTSSSTVVKD GKTSSASTPATRPVTGSWKKNQYGTWYKPENATFVNGNQPIVTRIGSPFLNAPVGGNLPA GATIVYDEVCIQAGHIWIGYNAYNGNRVYCPVRTCQGVPPNQIPGVAWGVFK*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGACTGGACTT GGCGGGTCTTCTGTTTGCTTGCCGTGCCCCCGGAGCACACTCTATGAAGCACCACCAC CATCACCCTCTGCCGGCTGGAAGTGCTGTTTCAGGGCCCTATGGCCAAGACACAGGC CGAGATCAACAAGCGGCTGGACGCCTATGCCAAGGGCACAGTGGATAGCCCTACAGA GTGAAGAAGGCCACCAGCTACGACCCAGCTTCGGCGTTATGGAAGCCGGCGCTATTG ACGCCGACGGCTATTATCATGCCAGTGCCAGGACCTGATCACCAGCTATGTGCTGTGG CTGACCGACAACAAAGTGCGGACATGGGGCAACGCCAAGGACCAGATCAAGCAGAGC TACGGCACC GGCTTCAAGATCCACGAGAACAAGCCTAGCACC GTGCTAAGAAAGGCT GGATCGCCGTGTTTACCAGCGGCAGCTATGAGCAGTGGGGCCACATCGGCATCGTGTA CGATGGCGGCAATACCAGCACCTTACCATCCTGGAACAGAACTGGAACGGCTACGCCA ACAAGAAACCCACCAAGAGAGTGGACAACCTACTACGGCCTGACACACTTATCGAGATC CCCGTGAAGGCCGGCACCACCGTGAAGAAAGAGACAGCCAAGAAGTCCGCCAGCAAG ACCCCTGCTCCTAAGAAGAAAGCCACACTGAAGGTGTCCAAGAACCACATCAACTACAC CATGGACAAGCGGGGCAAAAAGCCGAAGGCATGGTCATCCACAACGACGCCGGCAG ATCTAGCGGCCAGCAGTACGAAAACAGCCTGGCCAATGCCGGCTATGCCAGATACGCC AATGGAATCGCCACTACTATGGCAGCGAGGGCTATGTGTGGGAAGCCATCGACGCCA AGAATCAGATCGCCTGGCATAAGGCGACGGAACCGGCCCAATTCCGGCAATTTAG ATTCGCCGGCATCGAAGTGTGCCAGAGCATGTCTGCCAGCGACGCCCAGTTCCTGAAGA ATGAGCAGGCCGTGTTCCAGTTTACCGCCGAGAAGTTCAAAGAGTGGGGACTGACCCC TAACAGAAAGACCGTGCGGCTGCACATGGAATTCGTGCCTACCGCATGTCCCCACCGCT CTATGGTGCTGCACACCGGCTTTAACCCGTGACACAGGGCAGACCTAGCCAGGCCATC ATGAACAAGCTGAAGGACTACTTCATTAAAGCAGATCAAGAACTACATGGATAAGGGCA CCAGCAGCAGCACCGTGGTCAAGGATGGCAAGACCAGCAGCGCCTCTACACCAGCCAC AAGACCTGTGACAGGCAGCTGGAAGAAGAACCAGTACGGCACCTGGTACAAGCCCGA GAACGCCACCTTCGTGAACGGCAACCAGCCTATCGTGACCAGAATCGGCAGCCCCCTCC TGAATGCCCCTGTGCGAGGAAATCTTCTGCTGGCGCCACCATCGTGTATGACGAAGTC TGTATTAGGCCGGGCCACATCTGGATCGGCTACAACGCCTACAACGGCAACAGAGTGTA CTGCCCCGTGCGGACTTGTGAGGGCGTTCACCTAATCAGATTCCAGGCGTGGCCTGGG GCGTGTTCAAATGActcagCTGGTACTGCATGCACGCAATGCTAGCTGCCCTTTCCCGT CCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGC CCCACTCACCACCTCTGCTAGTTCCAGACACCTCCCAAGCACGAGCAATGCAGCTCAAA ACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAAC GAAAGTTTAACTAAGCTATACTAACCCAGGGTGGTCAATTTCTGCCAGCCACACCTT GGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAA AA AAAAA
L0482	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGGPREAPKTQIYTVKKGDTLSAIA LKYKTTVSNIQNTNNIANPNLIFIGQKLKVPMTPLVEPKPTVSSNNKSNSNSTLNYLKTLEN RGWDFDGSYGWQCFDLNVVYWNHLYGHGLKGYGAKDIPYANNFNSEAKIYHNTPTFKAE PGDLVVFSGRFGGGYGHTAIVLNGDYDGKLMKFQSLDQNWNNGGWRKAIEVAHKVVHN YENDMIFIRPFKKA*

	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGATTGGACAT GGAGGGTGTTTTGTCTGCTCGCAGTTGCTCCAGGTGCACATAGCATGAAACACCACCAT CACCACCATTCTGCTGGGCTTGAGGTGCTGTTCAAGGACCTAGAGAGGCTCCCAAAC ACAGATTTACACCGTCAAGAAGGGAGATACGCTTAGCGCCATTGCACTGAAATACAAGA CCACCGTATCAAACATCCAGAACCAATAACATAGCCAATCCAATCTCATCTTCATTG GCCAGAAGCTCAAGGTTCCCATGACTCCACTGGTCAACCGAAACCTAAGACAGTGTCC TCCAATAACAAGAGTAACAGCAACTCCTCTACACTGAACTATCTCAAGACGTTGGAGAA TCGCGGTTGGGACTTTGACGGCAGTTACGGCTGGCAATGCTTCGACTTGGTGAATGTGT ACTGGAATCACCTGTATGGGCATGGACTGAAAGGCTATGGGGCCAAAGACATACCTTAT GCTAACAAATTTCAACTCTGAGGCCAAGATCTACCACAACACTCCTACCTTCAAGGCAGAA CCAGGAGATCTGGTTGTGTTTAGCGGACGGTTTGGCGGAGGCTATGGGCATACTGCCA TTGTGCTTAATGGGGATTACGATGGCAAACCTGATGAAATTCCAGTCACTGGACCAGAAC TGAACAATGGTGGTTGGCGAAAAGCGGAAGTAGCCCAAGGTCGTCCACAATTACG AGAACGACATGATCTTCATCCGTCCTTTAAGAAAGCGTGAActcgagCTGGTACTGCATGC ACGCAATGCTAGCTGCCCCCTTCCCGTCTGGGTACCCCGAGTCTCCCCGACCTCGGGT CCCAGGTATGCTCCACCTCCACTGCCCCACTCACCACCTCTGCTAGTTCAGACACCTC CCAAGCACGCAGCAATGCAGCTCAAACGCTTAGCCTAGCCACACCCCCACGGGAAACA GCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTT GGTCAATTTCTGTGCCAGCCACACCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAGCATATGACTAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
L0483	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGPAATHEHSAQWLNNYKKGYG YGPYPLGINGGMHYGVDFFMNIGTPVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQ WYMHLSKYNVKVGDYVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPLKS AGYGKAGGTVTPPTNTGWKTNKYGTLYKSESASFTPNTDIITRTTGPFPSMPQSGVLKAGQ TIHYDEVKQDGHVWVGYTGNNSGQRIYLPVRTWNKSTNTLGLVLWGTIK*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGATTGGACCT GGAGAGTGTTTTGCCTTCTCGCTGTTGCCCTGGTGCCCATTCATGAAGCACCACCACC ACCATCATAGTGCAGGGCTGGAAGTTCTGTTTCAGGGTCCAGCAGCAACGCATGAGCAT TCCGCTCAATGGCTCAACAACTACAAGAAGGGTTACGGCTATGGCCCCTATCCTCTGGG GATAAACGGAGGGATGCACTATGGAGTCGATTTCTTCATGAATATTGGCACTCCAGTGA AAGCGATTTCCAGCGGGAAGATTGTGGAAGCTGGGTGGAGCAACTATGGTGGTGGCA ATCAGATCGGACTGATTGAGAATGACGGCGTGCACAGACAGTGGTATATGCATCTGAG TAAGTACAACGTCAAGGTGGGAGACTACGTCAAAGCCGGGCAGATCATAGGATGGTCC GGCTCTACAGGTTACTCAACAGCGCCACATCTCCACTTTCAGCGCATGGTGAATAGCTTC AGCAATTCAACCGCCCAAGATCCGATGCCCTTCTGAAGTCTGCCGGCTATGGGAAAGC TGGAGGAACAGTGACACCTACTCCCAATACCGGATGGAAAAACGAACAAGTACGGCACT CTGTACAAATCCGAATCTGCCAGTTTACACCCAATACCGACATCATCACCCGAACCACT GGACCTTTTCGGTCAATGCCGAGTCTGGAGTACTGAAAGCAGGGCAAACCTATCCACTA TGACGAGGTGATGAAGCAGGATGGCCACGTATGGGTGGGGTACACAGGCAATAGCGG CCAGAGGATTTACTTGCCAGTCCGTACCTGGAACAAAAGCACGAACACTCTTGCGGTTT TGTGGGGCACCATCAAGTGActcgagCTGGTACTGCATGCACGCAATGCTAGCTGCCCT TTCCCGTCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTC CACCTGCCCCACTCACCACCTCTGCTAGTTCAGACACCTCCAAGCACGCAGCAATGCA GCTCAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGC AATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTTCTGTGCCAGC CACACCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTA

[illegible]

	AGTCAGCTGTCAATGGTCTGAAGAAGTGActcgagCTGGTACTGCATGCACGCAATGCTA GCTGCCCCTTTCCCGTCCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGC TCCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCAGACACCTCCCAAGCACGCA GCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAA CCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCACAGGGTTGGTCAATTTG TGCCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAT ATGACTAAA AAAAAAAAAAAAAAAAAAAA
L0489	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGQPMMLMTKNQAEKWFDNSLGK QFNPDGWYGFQCYDYANMFFMLATGERLQGLYAYNIPFDNKAKIEKYQIIKNYDSFLPQK LDIVVFPKYGGGAGHVEIVESANLNTFTSFGQNWNGKGWNTNGVAQPGWGPETVTRHV HYDNPMPYFIRLNFNNLSVGNKAKGIIKQATTKEAVIKPKIMLVAGHGYNDPGAVGNG TNERDFIRKYITPNIKYL RHAGHEVALYGGSSQSQDMYQDTAYGVNVGNKKDYGLYWVK SQGYDIVLEIHLDAAGESASGGHVISSQFNADTIDKSIQDVIKNNLGQIRGVTPRNDLLNVN VSAEININYRLSELGITNKNMDWIKKNYDLYSKLIAGAIHGKPIGGLVAGNVKTSAKNKKN PPVPAGYTLKNNVPYKKEQGNVTANVKGNVNRDGYSTNSRITGVLNNTTITYDGAYCI NGYRWITYIANSQRRYIATGEVDKAGNRISFSGKFSTI*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGACTGGACTT GGCGGGTCTTCTGTTTGCTTCCGTCGCCCCGGAGCACACTCTATGAAGCACCACCAC CATCACCCTCTGCCGGCTGGAAGTGCTGTTTCAGGGCCCTATGCTGATGACCAAGAA CCAGGCCGAGAAGTGTTTCGACAACAGCCTGGGCAAGCAGTTCAACCCCGATGGATGG TACGGCTTCCAGTGCTACGACTACGCCAACATGTTCTTCATGCTGGCCACCGGCGAGAG ACTGCAGGGCCTGTACGCCTACAACATCCCCCTTCGATAACAAGGCCAAGATCGAGAAGT ACGGCCAGATCATCAAGAACTACGACAGCTTCCTGCCTCAGAACTGGACATCGTGGTG TCCCCAGCAAATATGGCGGCGGAGCCGGACACGTGGAATCGTGGAAGCGCCAACC TGAACACCTTACCAGCTTCGGCCAGAACTGGAACGGCAAAGGCTGGACCAATGGCGT GGCACAACCTGGCTGGGGACCTGAGACAGTGACAAGACACGTGCACTACTACGACAAC CCCATGTACTTCATCCGGCTGAACCTCCCAACAACCTGAGCGTGGGCAACAAAGCCAA GGGCATCATCAAGCAGGCCACCACCAAGAAAGAAGCCGTGATCAAGCCCAAGAAAATC ATGCTGGTGGCTGGCCACGGCTACAACGATCCTGGCGCTGTTGGCAACGGCACCAACG AGAGAGACTTCATCAGAAAGTACATCACGCCCAATATCGCCAAGTACCTGAGGCACGCC GGACATGAGGTTGCCCTGTATGGCGGATCTAGCCAGAGCCAGGACATGTACCAGGATA CCGCCTACGGCGTGAACTGGGAAACAAGAAGGACTACGGCCTGTACTGGGTCAAGAG CCAGGGCTACGACATCGTGCTGGAATCCACCTGGATGCCGCTGGCGAATCTGCCTCTG GCGGACATGTGATCATCAGCAGCCAGTTCAATGCCGACACCATCGACAAGAGCATCCAG GATGTGATCAAGAACAACCTGGGCCAGATCCGGGGCGTGACCCTAGAAACGATCTGC TGAATGTGAACGTGTCCGCCGAGATCAACATCAACTACCGGCTGAGCGAGCTGGGCTTC ATCACCACAAGAACGACATGGACTGGATCAAGAAGAATTACGACCTCTACAGCAAGCT GATCGCTGGCGCCATCCACGGCAAGCCTATTGGAGGACTGGTGGCCGGCAACGTGAAA ACCAGCGCCAAGAACAAGAAAAACCTCCTGTGCCTGCCGGCTACACCCTGGACAAGA ACAATGTGCCCTACAAGAAAGAGCAGGGCAACTACACCGTGGCCAACGTGAAGGGAAA CAACGTGCGCGACGGCTACAGCACCACAGCAGAATTACAGGCGTGCTGCCCAACAAC ACCACAATCACCTACGACGGCGCCTACTGCATCAACGGCTATCGGTGGATCACCTACATT GCCAACTCCGGCCAGCGGCGGTATATTGCCACAGGCGAAGTGGAACAAGGCCGGCAATA GAATCAGCTCCTTCGGCAAGTTACGACCATCTAAActcgagCTGGTACTGCATGCACGCA TGCTAGCTGCCCTTTCCCGTCCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGG TATGCTCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCAGACACCTCCCAAGC

	ACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGT GATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCCAGGGTTGGTCA ATTCGTGCCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAGCATATGACTAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA
L0499	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGPTQNYVVKGDYLGAIKKFN TTVSKIQLNNIKNPFIIGQVLKVPFISSNNSSTNTNRRNNSSTLNYLNSLAYKGWDFDG AYGNQCFDLNVVYWNHLYGHGLRGYGAKDIPFANNFRGEATVYRNTPTFKAELGDVAVFS GRYGGGFGHTAIVLNGNYDGKLMKFQSLDQNWYGGDSRRADIAHKVVHNYDYDMYFIRP YKKA*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGACTGGACTT GGAGGGTGTCTGCCTGTTGGCTGTAGCACCAGGAGCACACTCCATGAAACATCACCAT CATCACCCTCTGCCGACTTGAGGTCTGTTTCAAGGTCCCACACAGAATTACGTGGT GAAGAAGGGCGATTACCTGGGTGCGATCGCCAAGAAGTTCAACACCACAGTGAGCAAG ATCCAGAACTCAACAACATCAAGAATCCCACTTCATCCTGATTGGCCAAGTGCTGAA AGTACCGTTTATAAGCTCTAACAACCAAGCACCAATACGAACCGCCGGAACAACCTCA GTAGCACACTGAACCTGAATTCCTCGCATACAAAGGGTGGGATTCGATGGAGCC TATGGCAATCAGTGTTTGACCTCGTCAACGTGTACTGGAATCACCTCTATGGTCATGGG CTGAGAGGGTACGGTGCTAAGGACATACTTTTGCCAATAATTTTCGCGGCGAAGCTAC TGTGTATCGGAATACCCCAACCTTCAAAGCAGAGCTTGGAGATGTTGCTGTTTTCTCAG GGAGGTATGGAGGAGGCTTTGGCCATACTGCCATTGTCCTGAATGGCAACTATGACGG GAACTGATGAAATCCAGTCTTTGGACCAGAATTGGTATGGCGGGGATAGTAGACGT GCCGATATTGCGCACAAGTTGTCCACAACCTATGACTACGACATGACTTTATCCGACCT TACAAGAAGGCCTGActgagCTGGTACTGCATGCACGCAATGCTAGCTGCCCTTTCCCG TCCTGGGTACCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTG CCCCACTCACCACCTCTGCTAGTTCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAA AACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAA CGAAAGTTTAACTAAGCTATACTAACCCCAGGGTTGGTCAATTCGTGCCAGCCACACCC TGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAA AA AAAAA
L0502	MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGPMNKVDWLKSKHDLKNLTLLI VVGAFILFILLFMLMNHSDQGSNRITVSENAELRDPNAAYPVIYQVEKGDFTFRLSKSGK WIEVESRDGSEKGWIAGWHTSLNIEADSNKKVNPLKNKVIVLDPGHGGNDQGASSNTSKK SLEKNYTLKTAKELRRALEKAGATVKLTRSDDTYVSLDDRKLNGDAYISIHNSLSDSPNANGA TVYWYKNNQEALAE TLNTNIQKKAMLSNRGTRQENFQVLRQTNKPAVLLELGYISNPTDEV MIRNQIHRQVVEEAVVDGLKQYFAS*
	ACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGACTGGACTT GGCGCGTGTTCCTGCTGCTCGCGTAGCTCCTGGAGCCCATAGCATGAAGCACCATCAT CACCACCATTCAGCGGGATTGGAGGTTCTTTCAAGGCCCATGAATAAGGTGGATGC TTGGCTCAGCAAACACGACCTTAAGAACCTCACTACCTTGATCGTTGTAGGAGCCTTCAT TCTGTTCTGATACTGCTGTTGCTGATGAATCACTCTGATCAGGGGTCAAACCGCAT AACTGTTTCCGAGAATGCCGAACCTCGGACAGATCCAAACGCCGCTATCCGGTCATCT ATCAGGTCGAGAAAGGGGACACTTTCACCAGACTGTCTAAAAGCGGCAAATGGATTGA GGTCGAGAGTAGAGATGGGAGTGAGAAGGGGTGGATAGCAGGATGGCACACATCCCT GAACATTGAAGCCGATAGCAACAAGAAAGTGAACCCTCTCAAGAACAAAGTGATCGTG CTGGATCCTGGTCATGGTGGCAACGATCAGGGCGCTTCAAGCAATACCAGCAAGAAGT

	CACTGGAGAAGAATTACACGCTGAAAACAGCCAAGGAACTGCGACGGGCTCTTGAGAA AGCAGGAGCTACGGTGAACTGACACGGTCTGACGACACCTATGTGAGTCTGGACGAC AGGAAGCTGAATGGCGACGCGTATATCTCCATTACAATGACAGCCTGGATTCTCCAA TGCAAATGGGGCTACCGTCTACTGGTACAAGAACAACCAAGAGGCCCTGGCAGAAACC CTCAATACCAACATCCAGAAGAAGGCGATGTTGTCCAATAGGGGTACTAGGCAGGAAA ACTTTCAGGTGCTTAGACAGACAAACAAACCCGAGTTCTGTTGGAACCTCGGCTACATC TCCAACCAACAGACGAGGTGATGATCCGCAATCAGATTCACCGTCAAGTAGTGGAAG AAGCAGTGGTCGATGGCCTGAAACAGTACTTTGCCAGTTGActgagCTGGTACTGCATG CACGCAATGCTAGCTGCCCCCTTCCCGTCTGGGTACCCCGAGTCTCCCCGACCTCGGG TCCCAGGTATGCTCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGACACCT CCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACCCCCACGGGAAAC AGCAGTGATTAACCTTTAGCAATAAACGAAAGTTAACTAAGCTATACTAACCCAGGG TTGGTCAATTTCTGTGCCAGCCACACCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
--	---

Example 4: *In vitro* expression of RNA-encoded lysin variants

In vitro expression and secretion of the generated lysin constructs is analyzed by lipofection of the mRNA into HEK293T/17 cells and subsequent quantification of secreted His-tagged protein by Bio-layer interferometry (BLI). One day prior to lipofection, 1.2×10⁶ HEK293T/17 cells are seeded in 3 mL DMEM (Life Technologies GmbH, cat. no. 31966-021) + 10% fetal bovine serum (FBS, Biochrom GmbH, cat. no. S0115) in 6-well plates. For lipofection, 3 µg mRNA is formulated under sterile and RNase-free conditions using 400 ng mRNA per µL Lipofectamine MessengerMax (Thermo Fisher Scientific, cat. No. LMRNA015) and applied per 10 cm² culture dish to the HEK293T/17 cells at approximately 80 % confluence. After 20 h of expression, supernatants are collected under sterile conditions and stored at -20°C until further use. The levels of secreted lysins in the supernatant are quantified by the label-free optical analytical technique BLI through detection of His-tagged proteins. In short, a single point measurement of supernatant using the pre-hydrated Ni-NTA biosensor FortéBio BLItz® System (FortéBio) is performed. The measured binding rate is used to calculate the lysin concentration based on a calibration curve.

Example 5: Bactericidal activity of lysins encoded in and secreted from HEK293T/17 cells

Supernatants are serially diluted with lysin buffer (20 mM NaPO₄ pH 6.0, 100 mM NaCl) and tested for bactericidal activity. For this purpose, 30 µL of supernatants (and dilutions thereof) are mixed with *S. aureus* cells in 70 µL BHI medium such that the reaction contains 5×10⁵

CFU/mL. The optical density at 620 nm (OD₆₂₀) of these solutions is monitored in a plate reader over the course of 24 hours to quantify bacterial growth. The growth of *S. aureus* or its failure will reveal the levels of active lysin in the (diluted) supernatant. Supernatants from mock-transfected HEK293T/17 cells serve as a control to rule out that conditioned culture medium itself has a lytic activity on *S. aureus* cells. The failure of *S. aureus* cells to grow in the presence of supernatants from cells expressing and secreting lysins indicates that the level of active lysins in the supernatant is sufficiently high to lyse the *S. aureus* cells in the culture. The fold dilutions at which bacterial growth starts to appear further allows to identify the most potent lysin constructs, whose lytic activity depends on expression, secretion and enzymatic activity rates.

Taken together, these experiments reveal mRNA-encoded lysins which can be expressed and secreted in an active form in human cells, resulting in successful lysis of *S. aureus*.

Example 6: Endolysins expressed in HEK cells

15 mRNA encoded endolysin is expressed by HEK cells

To assess expression of phage endolysins in eukaryotic cells, HEK293T cells were transfected with modRNA endolysin constructs. Specifically, constructs 9_LysPA26 and 10_LysPA26 were used which both contain a FLAG tag for detection, and 10_LysPA26 contains an additional signal sequence for protein secretion (husec).

20

Construct 9_LysPA26: pST5-T7-AGA-dEarl-hAg-FLAG-LysPA26-FI-A30LA70

aa sequence:

MDYKDDDDKRTNNIDAIKEHEGLRLVAYLDSVGVTIGYGDTPDVVKGLTITKEEAELRLKRLVEFEGY
VNTYVKVPLKQHQFDALVSLVYNIGPTNFKTSTLLKLNAGDYIGAADQFLVWNKGRVDGKLVVIGLANR
25 RAKERKQFLGE

nt sequence:

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGACTACAAAGAC
GATGACGACAAGAGAACCAACAACATCGATGCCATCAAAGAACATGAAGGACTGAGACTGGTGGCTT
ACCTGGATTCTGTGGGAGTGTGGACAATTGGATATGGAGACACAGGACCTGACGTGGTGAAGGGACT
30 GACAATCACAAAGGAAGAAGCCGAAAAGAGACTGAGAAAGAGACTGGTGAATTTGAAGGATATGT
GAACACCTATGTGAAAGTGCCTCTGAAACAGCACCAGTTTGATGCTCTGGTGAGCCTGGTGTACAACAT
TGGACCCACAACTTCAAGACAAGCAGCTGCTGAAGAAGCTGAATGCTGGAGATTACATTGGAGCTG
CTGATCAGTTCCTGGTGTGGAACAAAGGCAGAGTGGATGGCAAAGCTGGTGGTGTCAAAGGCCTGGC
CAACAGGAGAGCCAAAGAAAGAAAACAGTTCCTGGGAGAATGATGACTCGAGCTGGTACTGCATGCA
35 CGCAATGCTAGCTGCCCCTTCCCGTCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATG

CTCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCAATGC
 AGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACG
 AAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTCGTGCCAGCCACACCCTGGAGCTAGCA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAA
 5 AAAAAAAAAAAAAAAAAAAAAAAAAA

Construct 10_LysPA26: pST5-T7-AGA-dEarI-hAg-husec(opt)-FLAG-LysPA26-FI-A30LA70

aa sequence:

MRVMAPRTLILLSGALALTETWAGSDYKDDDDKRTNNIDAIKEHEGLRLVAYLDSVGVTIGYDGTGPD
 10 VVKGLTITKEEAERLRKRLVEFEGYVNTYVKVPLKQHQFDALVSLVYNIGPTNFKTSTLLKLNAGDYIGAA
 DQFLVWNKGRVDGKLVVIKGLANRRAKERKQFLGE

nt sequence:

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGAGAGTGATGGCC
 CCCAGAACCCTGATCCTGCTGCTGCTGGCGCCCTGGCCCTGACAGAGACATGGGCCGGAAGCGACTA
 15 CAAAGACGATGACGACAAGAGAACCAACAACATCGATGCCATCAAAGAACATGAAGGACTGAGACTG
 GTGGCTTACCTGGATTCTGTGGGAGTGTGGACAATTGGATATGGAGACACAGGACCTGACGTGGTGA
 AGGGACTGACAATCACAAAGGAAGAACCGAAAAGAGACTGAGAAAGAGACTGGTGGAATTTGAAG
 GATATGTGAACACCTATGTGAAAGTGCCTCTGAAACAGCACCAGTTTGATGCTCTGGTGAGCCTGGTGT
 ACAACATTGGACCCACAACTTCAAGACAAGCACACTGCTGAAGAAGCTGAATGCTGGAGATTACATT
 20 GGAGCTGCTGATCAGTTCCTGGTGTGGAACAAAGGCAGAGTGGATGGCAAACCTGGTGGTATCAAAG
 GCCTGGCCAACAGGAGAGCCAAAGAAAGAAAACAGTTCCTGGGAGAATGATGACTCGAGCTGGTACT
 GCATGCACGCAATGCTAGCTGCCCTTTCCCGTCTGGGTACCCGAGTCTCCCCGACCTCGGGTCCC
 AGGTATGCTCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGACACCTCCCAAGCACGCA
 GCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCA
 25 ATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTCGTGCCAGCCACACCCTGGA
 GCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

Transfected cells were assessed for viability, intracellular protein expression and surface
 30 protein expression using flow cytometry (Figure 2). Cells were stained with a viability dye, and
 then to determine intracellular protein expression cells were permeabilized and stained with
 an anti-FLAG antibody conjugate to a fluorescent dye. Surface staining was performed without
 the permeabilization step. Transfection rate indicates the percentage of HEK cells positive for
 presence of the protein and expression rate is representative of the mean fluorescence of the
 35 positive cells. There is clear expression of protein in HEK cells, particularly evident in
 intracellular staining. Since the constructs do not encode for a transmembrane domain,
 expression on the surface was not expected, however interestingly there is some surface
 expression in the construct (10_LysPA26) with a secretion signal.

mRNA encoded endolysin is secreted by HEK cells

Protein secretion of the same constructs was assessed via Western blot (Figure 3). Culture supernatants and cell lysates were collected from two separate transfection experiments, with 350 ng and 1 µg of RNA, done in parallel as a test for optimal RNA concentration for protein expression. Protein was detected using an HRP-conjugated anti-FLAG antibody.

- 5 Protein expression was clearly visible at the expected size of ~17 kDa in cell lysates of cells transfected with 9_LysPA26 at the concentration of 1 µg. Protein secreted into the cell supernatant was detected with a faint band in supernatant of cells transfected with 10_LysPA26, the construct containing the secretion signal, and the band was absent in samples without the secretion signal. The results demonstrate clear protein expression, as
10 well as secretion when a secretion signal is included in the constructs.

Testing secreted endolysin for effectiveness in reducing viable cell count

- To verify the ability of the endolysin to break down bacterial cells extracellularly, log (OD₆₀₀ 0.8) and stationary phase (OD₆₀₀ 2) bacterial cells are incubated with culture supernatant of HEK cells expressing RNA encoded endolysin. XTT assay, a colorimetric assay in which
15 metabolically active cells reduce a tetrazolium salt to a formazan dye, is used to measure bacterial cell viability. Bacterial species (Gram-positive bacteria) not susceptible to LysPA26 endolysin are used as a negative control.

Testing secreted endolysin for effectiveness in reducing bacterial biofilms

- To investigate the potential of the endolysin to clear bacterial biofilms, 48h old biofilms are
20 incubated with culture supernatant of HEK cells expressing endolysin. Surface-attached bacterial biofilms are formed in 96 well plates by seeding log culture bacteria diluted to OD₆₀₀ 0.05 and static incubation at 37°C for 48h in LB media. The media containing planktonic, not attached, cells are carefully removed, and biofilms overlayed with culture supernatant and incubated for 2h at 37°C. Crystal violet stain is subsequently used as a stain for determining
25 biofilm biomass and XTT colorimetric assay to determine cell viability. Bacterial biofilms incubated with culture supernatant not containing secreted endolysin are used as a control.

Example 7: RNA constructs for endolysin expression in human subjects

- Different RNA constructs encoding endolysins as schematically shown in Figure 4 were
30 prepared. The different constructs either do not contain or contain a signal peptide. All constructs contain all the elements necessary for expression of the lysin LysPA26 in human

cells, including a cap, a 5' UTR, a 3' UTR and a polyA sequence. An overview of the constructs is given in the following table:

SEQ ID #	Construct	Signal peptide	Length of RNA [nt]	Mw [Da]
6	pST5-T7-AGA-dEarI-hAg-LysPA26-FI-A30LA70	-	899	16206.5
7	pST5-T7-AGA-dEarI-hAg-husec(opt)-LysPA26-FI-A30LA70	husec	974	
8	pST1-OPAL-LysPA26-A30LA70	-	7972	
9	pST1-OPAL-husec(opt)-LysPA26-A30LA70	husec	8047	

- 5 RNA encoded endolysins are tested against Gram-negative bacteria, given that according to the World Health Organization the top three resistant bacterial pathogens for which novel treatment is urgently needed, carbapenem-resistant *Acinetobacter baumannii*, carbapenem-resistant *Pseudomonas aeruginosa*, and carbapenem-resistant *Enterobacteriaceae*, all belong to the group of Gram-negative bacteria (WHO, 2017). Following the initial testing, this
- 10 treatment option is also used in other types of bacteria – including those belonging to the Gram-positive group, such as methicillin and vancomycin resistant *Staphylococcus aureus* and penicillin non-susceptible *Streptococcus pneumoniae*.

Example 8: Endolysins expressed in HEK cells

- 15 HEK293T cells were transfected with 0.5 and 1 µg of modRNA and saRNA of construct 9_LysPA26 and construct 10_LysPA26 (each containing a HiBiT tag for detection) in an 96 well plate. The nucleotide and amino acid sequences of fusion-protein nucleoside-modified RNA constructs are shown in SEQ ID NOs 111, 112 and SEQ ID NOs 113, 114, respectively. The nucleotide and amino acid sequences of fusion-protein self-amplifying RNA constructs are
- 20 shown in SEQ ID NOs 115, 112 and SEQ ID NOs 116, 114, respectively. Transfection process was as previously described. After 18h, for detection of extracellular protein the Nano-Glo

HiBit extracellular reagent was used. As per manufacturers instructions LgBit protein was diluted 1:100 and the Nano-Glo HiBit extracellular Substrate was diluted 1:50 in the necessary volume of the Nano-Glo HiBit Extracellular buffer. Plates were equilibrated to room temperature for 10 min, and 100 µl of the reagent was added to each well (to have equal
5 volume of culture media and reagent). Plates were placed on an orbital shaker for 3 min at 300 rpm and luminescence measured after 10 mins using the Tecan Spark microplate reader. Data is represented in relative luminescence units (RLU). The same procedure was conducted for detecting protein in the cell lysate, except the Nano-Glo Lytic Reagents were used from the Nano-Glo Lytic Detection System kit, in which the reagents contain a cell lytic compound.
10 Results are shown in Figure 5. Left panels show supernatant and cell lysates for modRNA. The constructs with secretion signals (10_LysPA26) are clearly detectable at both concentrations in the supernatant, while the constructs without secretion signals (9_LysPA26) cannot be detected. The same result is seen for saRNA. This confirms that the expression and secretion of endolysins can be seen for different RNA platforms. All constructs can be detected in the
15 cell lysates showing that the endolysins are expressed in HEK cells irrespective of the RNA platform used.

Example 9: In vitro expression of DNA and RNA encoded lysin variants

In vitro expression and secretion of several lysin constructs is analyzed by lipofection of DNA
20 or mRNA into HEK293T/17 cells and subsequent quantification of secreted His-tagged protein by Western Blot (see table 6). One day prior to lipofection, 0.9×10^6 HEK293T/17 cells are seeded in 3 ml DMEM (Life Technologies GmbH, cat. no. 31966-021) + 10% fetal bovine serum (FBS, Biochrom GmbH, cat. No. S0115) per well in 6-well plates. For lipofection, 3, 5 and 7.5 µg, of mRNA or 2.5 µg of DNA are transfected. mRNA is formulated using the RiboJuice™
25 mRNA Transfection Kit (Sigma Aldrich, cat. no. TR-1013), DNA using Lipofectamine™ LTX Reagent with PLUS™ Reagent (Life Technologies GmbH, cat. no. 15338100) according to the manufacturer's instructions. After 24, 48 or 72 hours of expression, supernatants are collected, subjected to PNGase F digest (New England Biolabs GmbH, cat. no. P0704L) under denaturing conditions to remove potential N-glycosylation and stored at -20°C until further
30 use. The levels of secreted lysins are quantified by Western Blot through detection by an anti-His antibody (Abcam, ab49781 for N-terminal His-tag and Abcam, ab1187 for C-terminal His-

tag) or an anti-Flag antibody (Sigma A8592). Expression levels were quantified by quantification of chemiluminescent signals using Image Lab 6. Mammalian expression of a prototype endolysin was also assessed using an mRNA formulated in lipid nanoparticles (LNP-mRNA). Here 1.5 µg of formulated mRNA were transfected into HEK293T/17 or CHO-K1 cells.

- 5 Growth medium for CHO-K1 is F12-K (21127022, Thermo Fisher Scientific). Medium was changed to Opti-MEM prior to transfection for both cell lines.

Table 6: List of constructs tested for expression in HEK293T/17 cells. Expression from DNA as well as from mRNA transfections was assessed

Construct #	RNA/DNA	Construct	Signal peptide	Length of DNA/RNA [nt]	Predicted Mw [kDa] (without signal peptide)
11	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(IGHV1-2)-His-3C-hL0483-Fl-A30LA70	Ig heavy chain V-I region HG3 (MDWTWRVFCLLAVAPGAHS)	1313 (T7 start to end of Poly-A-tail)	28.87
12	DNA	pcDNA3.1(+)-SS(IGHV1-2)-His-3C-hL0483	Ig heavy chain V-I region HG3 (MDWTWRVFCLLAVAPGAHS)	914 (T7 start to end of lysin)	28.87
13	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hL0473-His-Fl-A30LA70	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	2099 (T7 start to end of Poly-A-tail)	58.05
14	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hL0482-His-Fl-A30LA70	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	1268 (T7 start to end of Poly-A-tail)	27.21
15	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hL0483-His-Fl-A30LA70	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	1361 (T7 start to end of Poly-A-tail)	29.93
16	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	1073 (T7 start to end	19.85

		FLAG-hL0484m-His-Fi-A30LA70		of Poly-A-tail)	
17	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hIL2-His-Fi-A30LA70	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	1022 (T7 start to end of Poly-A-tail)	18.4
18	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(hIL2)-FLAG-hIL2-His-Fi-A30LA70	Human IL2 (MYRMQLLSICALSLVTNS)	1010 (T7 start to end of Poly-A-tail)	18.4
19	DNA	pcDNA3.1(+)-SS(hINS)-FLAG-hL0473-His	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	1700 (T7 start to end of His tag)	58.05
20	DNA	pcDNA3.1(+)-SS(hINS)-FLAG-hL0482-His	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	869 (T7 start to end of His tag)	27.21
21	DNA	pcDNA3.1(+)-SS(hINS)-FLAG-hL0483-His	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	962 (T7 start to end of His tag)	29.93
22	DNA	pcDNA3.1(+)-SS(hINS)-FLAG-hL0484m-His	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	674 (T7 start to end of His tag)	19.85
23	DNA	pcDNA3.1(+)-SS(hINS)-FLAG-hIL2-His	Human insulin (MALWMRLLPLLALLALWGPDPAAA)	623 (T7 start to end of His tag)	18.4
24	DNA	pcDNA3.1(+)-SS(hIL2)-FLAG-hIL2-His	Human IL2 (MYRMQLLSICALSLVTNS)	611 (T7 start to end of His tag)	18.4
25	RNA	pST4-T7-AGA-dEarl-hAg-Kozak-SS(IGkappa)-His-3C-hL0483(PM)-FLAG-Fi-A30LA70	Mouse IGkappa MDMRAPAGIFGFLVLPFGYRS	935 (T7 start to end of Poly-A-tail)	29.1

Construct 11: pST4-T7-AGA-dEarl-hAg-Kozak-SS(IGHV1-2)-His-3C-hL0483-Fi-A30LA70

aa sequence:

MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGQPAATHEHSAQWLNNYKKGYGYPYPLGIN
 GGMHYGVDFMNIPTVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKVGDYVK
 AGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPLKSAGYGKAGGTVTPNTGWKTNKYGTLY
 KSESASFTPNTDIIITRTTGPFRSMPQSGVLKAGQTIHYDEVKQDGHVWVGYTGNSGQRIYLPVRTWNK
 5 STNTLGVLWGNIK*
 nt sequence:
 AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGATTGGACCTG
 GAGAGTGTTCCTTCTCGCTGTTGCCCTGGTGCCCATTCATGAAGCACCACCACCACCATCATA
 GTGCAGGGCTGGAAGTTCTGTTTCAGGGTCCAGCAGCAACGCATGAGCATTCCGCTCAATGGCTCAA
 10 CAACTACAAGAAGGGTTACGGCTATGGCCCTATCCTCTGGGGATAAACGGAGGGATGCACTATGG
 AGTCGATTTCTTCATGAATATTGGCACTCCAGTGAAAGCGATTTCAGCGGGAAGATTGTGGAAGCT
 GGGTGGAGCAACTATGGTGGTGGCAATCAGATCGGACTGATTGAGAATGACGGCGTGACAGACA
 GTGGTATATGCATCTGAGTAAGTACAACGTCAAGGTGGGAGACTACGTCAAAGCCGGGCAGATCAT
 AGGATGGTCCGGCTCTACAGGTTACTCAACAGCGCCACATCTCCACTTTCAGCGCATGGTGAATAGC
 15 TTCAGCAATCAACCGCCCAAGATCCGATGCCCTTCTGAACTCTGCCGGCTATGGGAAAGCTGGAG
 GAACAGTGACACCTACTCCCAATACCGGATGGAAAACGAACAAGTACGGCACTCTGTACAAATCCGA
 ATCTGCCAGTTTCACACCCAATACCGACATCATACCCGAACCACTGGACCTTTTCGGTCAATGCCGC
 AGTCTGGAGTACTGAAAGCAGGGCAAATATCCACTATGACGAGGTGATGAAGCAGGATGGCCACG
 TATGGGTGGGGTACACAGGCAATAGCGGCCAGAGGATTTACTTGCCAGTCCGTACCTGGAACAAAA
 20 GCACGAACACTCTTGCGTTTTGTGGGGCACCATCAAGTGACTCGAGCTGGTACTGCATGCACGCAA
 TGCTAGCTGCCCTTTCCCGTCTGGGTACCCGAGTCTCCCCGACCTCGGGTCCAGGTATGCTCC
 CACCTCCACCTGCCCCACTACCACTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGC
 TCAAAACGCTTAGCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAA
 AGTTTAATAAGCTATACTAATCCAGGGTTGGTCAATTTCTGTCAGCCACACCTGGAGCTAGCAA
 25 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Construct 12: pcDNA3.1(+)-SS(IGHV1-2)-His-3C-hL0483

aa sequence

MDWTWRVFCLLAVAPGAHSMKHHHHHSAGLEVLFGQPAATHEHSAQWLNNYKKGYGYPYPLGIN
 GGMHYGVDFMNIPTVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKVGDYVK
 AGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPLKSAGYGKAGGTVTPNTGWKTNKYGTLY
 YKSESASFTPNTDIIITRTTGPFRSMPQSGVLKAGQTIHYDEVKQDGHVWVGYTGNSGQRIYLPVRTWN
 KSTNTLGVLWGNIK*
 35 nt sequence:
 GAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGATTG
 GACCTGGAGAGTGTTCCTTCTCGCTGTTGCCCTGGTGCCCATTCATGAAGCACCACCACCACC
 40 ATCATAGTGACAGGGCTGGAAGTTCTGTTTCAGGGTCCAGCAGCAACGCATGAGCATTCCGCTCAATG
 GCTCAACAACACTACAAGAAGGGTTACGGCTATGGCCCTATCCTCTGGGGATAAACGGAGGGATGCA
 CTATGGAGTCAATTTCTTCATGAATATTGGCACTCCAGTGAAAGCGATTTCAGCGGGAAGATTGTG
 GAAGCTGGGTGGAGCAACTATGGTGGTGGCAATCAGATCGGACTGATTGAGAATGACGGCGTGCA
 CAGACAGTGGTATATGCATCTGAGTAAGTACAACGTCAAGGTGGGAGACTACGTCAAAGCCGGGCA
 45 GATCATAGGATGGTCCGGCTCTACAGGTTACTCAACAGCGCCACATCTCCACTTTCAGCGCATGGTG

AATAGCTTCAGCAATTCAACCGCCCAAGATCCGATGCCCTTCCTGAAGTCTGCCGGCTATGGGAAAG
 CTGGAGGAACAGTGACACCTACTCCAATACCGGATGGAAAACGAACAAGTACGGCACTCTGTACA
 AATCCGAATCTGCCAGTTTCACACCCAATACCGACATCATCCCCGAACCACTGGACCTTTTCGGTCA
 ATGCCCGAGTCTGGAGTACTGAAAGCAGGGCAAATATCCACTATGACGAGGTGATGAAGCAGGAT
 5 GGCCACGTATGGGTGGGGTACACAGGCAATAGCGGCCAGAGGATTTACTTGCCAGTCCGTACCTGG
 AACAAAAGCACGAACACTCTTGCGTTTTGTGGGGCACCATCAAGTGA

Construct 13: pST4-T7-AGA-dEarI-hAg-Kozak-SS(hINS)-FLAG-hL0473-His-FI-A30LA70

10 aa sequence:
 MALWMRLPLLALLALWGPDPAAAMKDYKDDDDKGGGGSGMAKTKKQIEQRLQAYAKGTVDSPYRIT
 KPTSYDPSFGVMEVGAIADADRYHAQCQDLITDYVLWLTDNKVRPWGNAKDQINQSYGTGFKIHKNLP
 STVAKTGWIAVFTSGTYSQYGHIGIVYSPGDTNKFHIIQNWNGWANKKPLTRWDNYGLTHFIEVPYA
 KETITTKKTPAKAATSKPAKKISVSHNNINLQSNFAKRGYNPKGVIHNDAGRSTAQQYENSLKGANQNR
 15 LNNGIAHAYADSSHIWYAIIEPERIAYHTGDGVKKGTGNYDYFGIEVCQSMSASDKEFMANEQVVLQEVA
 RLLKKWKLPANRNTVRLHLEFSATSCPHRSLTMRTGFNPVTQGRPSDAIHKLKDIWIKIRAYMEGKTPSI
 KTVANKPGSASTPATKSNSDGWKVNAYNTYYKTERASFTVGSTRIAVRSTGPFTTCPFTGWLNPVGTITY
 DEVMVQNGFVWLGYDSYNGRRYVPIRTAQGTPGSSNYVLGPLWGTISAAAGGGGSHHHHHH*

20 Nt sequence:

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGCCCTGTGGAT
 GAGACTGCTGCCTCTGCTTCTGCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATGAAGGAC
 TACAAAGACGATGACGACAAGGGGAGGAGGCGGCTCCGGAATGGCCAAAACCAAGAAACAGATAGA
 25 ACAGAGACTGCAAGCTTACGCAAAGGGGACCGTGGATTCCCTTACCGAATACCAAGCCAACCAAGC
 TACGACCCCTCCTTCGGTGTCTGATGGAGGTTGGCGCTATAGATGCGGATCGCTACTATCATGCCAGT
 GTCAAGATCTGATAACTGACTATGTACTTTGGCTCACTGACAACAAAGTAAGACCTTGGGGTAATGC
 CAAGGACCAGATTAACCAGAGCTATGGGACTGGTTTCAAATCCACAAGAATCTGCCAGTACCGTG
 GCTAAAACAGGATGGATTGCCGTGTTTACATCCGGGACCTACTCACAGTACGGCCACATCGGCATCG
 30 TGTACAGTCCAGGAGATACCAACAAATTCCATATCATCGAGCAGAATTGGAATGGATGGGCAAACAA
 GAAACCTCTGACACGTTGGGACAACCTACTACGGGCTGACCCACTTTATCGAGGTCCCATATGCGAAA
 GAGACAATCACACGAAGAAAACCTGCACCTAAAGCAGCTACCTCAAACCGGCTAAGAAGATCTCA
 GTTTCTCACAACAACATCAACCTTACAGAGCACTTTGCAAAGAGAGGGTACAACCCAAAGGGAGTGG
 TGATTACAATGACGCCGGACGCTCTACTGCGCAACAGTATGAAAACAGCCTGAAGGGCGCTAATCA
 35 GAACAGGCTGAATAACGGCATTGCCATGCATATGCTGACAGCAGCCACATTTGGTATGCTATCGAA
 CCCGAGAGAATCGCCTATCACACAGGCGATGGCGTTAAGAAGGGAACAGGCAATTACGACTACTTC
 GGTATTGAAGTATGCCAGAGCATGAGCGCAAGTGACAAAGAGTTTCATGGCCAACGAACAGGTGGTA
 CTGCAGGAGGTGGCTAGGTTGCTCAAGAAATGGAAGCTTCCCGCCAATCGGAATACCGTGCGGCTC
 CATCTGGAATTCTCTGCCACATCTTGCCCTCATAGGAGCCTCACAATGCGGACCGGTTTCAATCCGGT
 40 GACTCAGGGTAGACCTCAGACGCCATCATCAAGAAGTTGAAGGATATCTGGATAAAAGAGATTGCG
 CGCATATATGGAGGGTAAAACCCCATCCATTAAGACTGTGCGCAATAAGCCCCGGATCAGCCAGTACC
 CCTGCAACCAAAATCCAACCTCCGATGGGTGGAAGGTCAATGCCTACAATACCTACTACAAAACGGAAC
 GGGCCTCCTTACAGTGGGCTCTACTCGGATTGCTGTGCGTAGTACAGGACCTTTACTACATGCCCG
 TTTACAGGGTGGCTGAACCCCGGAGTCACGATAACCTATGACGAGGTTATGGTGCAAGACGGGTTT
 45 GTGTGGTTGGGCTATGATAGCTACAATGGGAGGAGGTATGTCCCCATTGCGACTGCCCAAGGAACCT
 CTGGCAGCTCCAATTATGTGCTTGCCCTCTGTGGGGCACGATTTCTGCGGCCGCGGAGGAGGCG
 GCTCCACCACCACCACCATCATTGACTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCTTT

CCCGTCCTGGGTACCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGCCCC
 ACTCACCACTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCT
 AGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTAT
 ACTAACCCAGGGTTGGTCAATTCGTGCCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAA
 5 AAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Construct 14: pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hL0482-His-FI-A30LA70

10 Aa sequence

MALWMRLLPLLALLALWGPDPAAAMKDYKDDDDKGGGSGREAPKTQIYTVKKGDTLSAIALKYKTTV
 SNIQNTNNIANPNLIFIGQKLKVPMTPLVEPKPKTVSSNNKSNSNSTLNYLKTLENRGWDFDGSYGWQC
 FDLNVVYWNHLYGHGLKGYGAKDIPYANNFNSEAKIYHNTPTFKAEPGDLVVFSGRFGGGYGHTAIVLN
 15 GDYDGKLMKFQSLDQNWNNGGWRKA EVAHKV VHN YENDMIFIRPFKAAAAGGGGSHHHHHH*

Nt sequence

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGCCCTGTGGAT
 GAGACTGCTGCCTCTGCTTGTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTATGAAGGAC
 20 TACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAAGAGAGGGCTCCAAAACACAGATTTA
 CACCGTCAAGAAGGGAGATACGCTTAGCGCCATTGCACTGAAATACAAGACCACCGTATCAAACATC
 CAGAACACCAATAACATAGCCAATCCCAATCTCATCTTCATTGGCCAGAAGCTCAAGGTTCCCATGAC
 TCCACTGGTCGAACCGAAACCTAAGACAGTGTCTCCAATAACAAGAGTAACAGCAACTCCTCTACAC
 TGA ACTATCTCAAGACGTTGGAGAATCGCGGTTGGGACTTTGACGGCAGTTACGGCTGGCAATGCTT
 25 CGACTTGGTGAATGTGTACTGGAATCACCTGTATGGGCATGGACTGAAAGGCTATGGGGCCAAAGA
 CATACCTTATGCTAACAATTTCAACTCTGAGGCCAAGATCTACCACAACACTCCTACCTCAAGGCAG
 AACCAGGAGATCTGGTTGTGTTTAGCGGACGGTTTGGCGGAGGCTATGGGCATACTGCCATTGTGCT
 TAATGGGGATTACGATGGCAAATGATGAAATTCAGTCACTGGACCAGAACTGGAACAATGGTGG
 TTGGCGAAAAGCGGAAGTAGCCACAAGGTCGTCCACAATTACGAGAACGACATGATCTTCATCCGT
 30 CCTTTAAGAAAGCGGCGGCCGCGGAGGAGGCGGCTCCACCACCACCACCATCATTGACTCGAG
 CTGGTACTGCATGCACGAATGCTAGCTGCCCCTTTCCCGTCCTGGGTACCCCGAGTCTCCCCGACC
 TCGGGTCCCAGGTATGCTCCACCTCCACCTGCCCCACTCACCACTCTGCTAGTTCCAGACACCTCCC
 AAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTA
 ACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTTCTGTGCCAGC
 35 CACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAA
 AA

Construct 15: pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hL0483-His-FI-A30LA70

40 Aa sequence

MALWMRLLPLLALLALWGPDPAAAMKDYKDDDDKGGGSGAATHEHSAQWLNNYKKGYGYPYPL
 GINGGMHYGVDFFMNIGTPVKAISSKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKVGD
 YVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPLKSAGYGKAGGTVTPTNTGWKTNKY
 45 GTLYKSESASFTPTNDIITRTTGPFRSMPQSGVLKAGQTIHYDEVMMKQDGHVWVGYTGNSSGQRIYLPVRT
 WNKSTNTLGLVLWGTTIKAAAGGGGSHHHHHH*

Nt sequence

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGCCCTGTGGAT
 GAGACTGCTGCCTCTGCTTGTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATGAAGGAC
 5 TACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAGCAGCAACGCATGAGCATTCCGCTCA
 ATGGCTCAACAACACTACAAGAAGGGTTACGGCTATGGCCCCCTATCCTCTGGGGATAAACGGAGGGAT
 GCACTATGGAGTCGATTTCTTCATGAATATTGGCACTCCAGTGAAAGCGATTTCCAGCGGGAAGATT
 GTGGAAGCTGGGTGGAGCAACTATGGTGGTGGCAATCAGATCGGACTGATTGAGAATGACGGCGT
 GCACAGACAGTGGTATATGCATCTGAGTAAGTACAACGTCAAGGTGGGAGACTACGTCAAAGCCGG
 10 GCAGATCATAGGATGGTCCGGCTCTACAGGTTACTCAACAGCGCCACATCTCCACTTTAGCGCATG
 GTGAATAGCTTCAGCAATTCAACCGCCCAAGATCCGATGCCCTTCTGAAGTCTGCCGGCTATGGGA
 AAGCTGGAGGAACAGTGACACCTACTCCCAATACCGGATGGAAAACGAACAAGTACGGCACTCTGT
 ACAAATCCGAATCTGCCAGTTTACACCCAATACCGACATCATCACCGAACCCTGGACCTTTTCGG
 TCAATGCCGCACTCTGGAGTACTGAAAGCAGGGCAAATATCCACTATGACGAGGTGATGAAGCAG
 15 GATGGCCACGTATGGGTGGGGTACACAGGCAATAGCGGCCAGAGGATTTACTTGCCAGTCCGTACC
 TGGAACAAAAGCAGCAACACTCTTGGCGTTTTGTGGGGCACCATCAAGGCGGCCGCCGAGGAGGC
 GGCTCCCAACCACCACCACCATCATTGACTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCTT
 TCCCGTCTGGGTACCCCGAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGCCC
 CACTCACCACCTCTGCTAGTTCCAGACACCTCCCAAGCAGCAGCAATGCAGCTCAAAACGCTTAGCC
 20 TAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTA
 TACTAACCCAGGGTTGGTCAATTTCTGTCAGCCACACCTGGAGCTAGCAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

25 Construct 16: pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hL0484m-His-FI-A30LA70

MALWMRLPLLALLALWGPDPAAAMKDYKDDDDKGGGSGDQFLDEKEGNHTTAYRDGSGIWTICR
 GATMVDGKPIVPGMRLSKEKCDQVNAIERDKALAWVERNIKVPLTEPQKAGIASFCPYNIGPGKCFPSTF
 YKRLNAGDRKGACEAIRWWIKDGGRDCCRISNNCYGQVIRRDQESALACWGIDQAAAGGGGSHHHHH
 30 H*

Nt sequence

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGCCCTGTGGAT
 35 GAGACTGCTGCCTCTGCTTGTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATGAAGGAC
 TACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAGATCAGTTTCTGACGAGAAGGAAGG
 CAATCACACCACTGCGTATAGGGATGGCTCTGGGATATGGACCATCTGTAGAGGCGCCACAATGGTC
 GATGGCAAACCCGTGATACCCGGAATGCGTCTGTCAAAGGAGAAGTGCATCAGGTGAATGCCATC
 GAACGCGACAAAGCTCTGGCCTGGGTTGAGCGGAACATCAAGGTTCCACTGACGGAACCACAGAAA
 40 GCCGGAATTGCGTCCTTCTGTCCGTACAACATTGGCCCTGGAAAGTGCTTTCCCTCCACTTTCTACAA
 GCGCCTGAATGCAGGGGACAGAAAAGGAGCCTGTGAGGCTATCAGGTGGTGGATCAAAGATGGAG
 GTAGGGACTGTCGAATCAGGAGCAACAACCTGCTATGGGCAAGTGATTGAGAGATCAGGAGAGT
 GCTCTGGCTTGTGGGGTATTGACCAGGCGGCCCGGAGGAGGCGGCTCCCAACCACCACCACCAT
 CATTGACTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCTTTCCCGTCTGGGTACCCCGAG
 45 TCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTC
 CAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAA
 CAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAA

TTTCGTGCCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATG
 ACTAAA
 AAAAAAA

5 Construct 17: pST4-T7-AGA-dEarl-hAg-Kozak-SS(hINS)-FLAG-hIL2-His-FI-A30LA70

Aa sequence

MALWMRLLPLLALLALWGPDPAAAMKDYKDDDDKGGGGSGAPTSSSTKKTQLQLEHLLDLQMILNGI
 10 NNYKNPKLTRMLTFKYMPKKATELKHLCLEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSET
 TFMCEYADETATIVEFLNRWITFCQSIISTLTAAAGGGGSHHHHHH*

Nt sequence

15 AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGCCCTGTGGAT
 GAGACTGCTGCCTCTGCTTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATGAAGGAC
 TACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAGCTCCAACATCTTCTTCAACAAAGAAA
 ACACAGCTTCAGCTTGAACACCTTCTTCTTGATCTTCAGATGATTCTGAATGGAATCAACAATTACAAA
 AATCCAAAAGTACAGAAGTCTGACATTTAAATTTTACATGCCAAAGAAAGCAACAGAAGTAAAC
 20 ACCTTCAGTGCCTTGAAGAAGAACTGAAACCTCTGGAAGAAGTGTGAATCTGGCTCAGAGCAAAAA
 TTTTCACCTGAGACCAAGAGATCTGATCAGCAACATCAATGTGATTGTGCTGGAAGTAAAGGATCT
 GAAACAACATTCATGTGTGAATATGCTGATGAAACAGCAACAATTGTGGAATTTCTGAACAGATGGA
 TCACATTTTGCCAGTCAATCATTTCAACACTGACAGCGGCCGCCGAGGAGGCGGCTCCCACCACCA
 CCACCATCATTGACTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCCTTTCCCGTCTGGGTA
 25 CCCCAGTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGCCCCACTCACCACCTCTG
 CTAGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAACGCTTAGCCTAGCCACACCCCCAC
 GGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGT
 TGGTCAATTTCTGTCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 GCATATGACTAAA
 30 AAAAAAAAAAAAAA

Construct 18: pST4-T7-AGA-dEarl-hAg-Kozak-SS(hIL2)-FLAG-hIL2-His-FI-A30LA70

Aa sequence

35 MYRMQLLSIALSLALVTNSMKDYKDDDDKGGGGSGAPTSSSTKKTQLQLEHLLDLQMILNGINNYKN
 PKLTRMLTFKYMPKKATELKHLCLEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCE
 YADETATIVEFLNRWITFCQSIISTLTAAAGGGGSHHHHHH*

40 Nt sequence

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGTACCGGATGCA
 GCTGCTGAGCTGTATCGCCCTGTCTGCGCCCTGGTCACCAATTCTATGAAGGACTACAAAGACGAT
 GACGACAAGGGAGGAGGCGGCTCCGGAGCTCCAACATCTTCTTCAACAAAGAAAACACAGCTTCAG
 45 CTTGAACACCTTCTTCTTGATCTTCAGATGATTCTGAATGGAATCAACAATTACAAAAATCCAAAAGT
 ACAAGAATGCTGACATTTAAATTTTACATGCCAAAGAAAGCAACAGAAGTAAACACCTTCAGTGCC
 TTGAAGAAGAACTGAAACCTCTGGAAGAAGTGTGAATCTGGCTCAGAGCAAAAATTTTACCTGAG

ACCAAGAGATCTGATCAGCAACATCAATGTGATTGTGCTGGAAGTAAAGGATCTGAAACAACATTC
 ATGTGTGAATATGCTGATGAAACAGCAACAATTGTGGAATTTCTGAACAGATGGATCACATTTTGCC
 AGTCAATCATTTCAACACTGACAGCGGCCGCCGGAGGAGGCGGCTCCACCACCACCACCATCATTG
 ACTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCCCTTCCCGTCTGGGTACCCCGAGTCTCC
 5 CCGGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGAC
 ACCTCCCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCA
 GTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTTG
 TGCCAGCCACACCCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAA
 AA
 10 AAA

Construct 19: pcDNA3.1(+)-SS(hINS)-FLAG-hL0473-His

Aa sequence

15 MALWMRLLPLLALLLWGPDPAAAMKDYKDDDDKGGGSGMAKTKKQIEQRLQAYAKGTVDSPIRIT
 KPTSYDPSFGVMEVGAIADRYHAQCQLITDYVLWLTDNKVRPWGNAKDQINQSYGTGFKIHKNLP
 STVAKTGWIAVFTSGTYSQYGHIGIVSPGDTNKFHIEQNWNGWANKPLTRWDNYGLTHFIEVPYA
 KETITTKKTAAPKAATSKPAKKISVSHNNINLQSNFAKRGYNPKGVVIHNDAGRSTAAQYENSLKGANQNR
 20 LNNGIAHAYADSSHIWYAIERIAHYHTGDGVKKGTGNYDYFGIEVCQSMSASDKEFMANEQVVLQEVA
 RLLKKWKLPANRNTVRLHLEFSATSCPHRSLTMRTGFNPVTQGRPSDAIHKLDIWIKEIRAYMEGKTPSI
 KTVANKPGSASTPATKSNSDGWKVNAYNTYYKTERASFTVGSTRIAVRSTGPFTTCPFTGWLNPVGTITY
 DEVMVQNGFVWLGYDSYNGRRYVPIRTAQGTGSSNYVLGPLWGTISAAAGGGGSHHHHHH

25 Nt sequence

GAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGCCCT
 GTGGATGAGACTGCTGCTCTGCTGCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATG
 AAGGACTACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAATGGCCAAAACCAAGAAACA
 30 GATAGAACAGAGACTGCAAGCTTACGCAAAGGGGACCGTGGATTCCCTTACCGAATACCAAGCC
 AACCAGCTACGACCCCTCCTTCGGTGTCATGGAGGTTGGCGCTATAGATGCGGATCGCTACTATCAT
 GCCCAGTGTCAGATCTGATAACTGACTATGTACTTTGGCTCACTGACAACAAAGTAAGACCTTGGG
 GTAATGCCAAGGACCAGATTAACCAGAGCTATGGGACTGGTTTCAAAATCCACAAGAATCTGCCCAG
 TACCGTGGCTAAAACAGGATGGATTGCCGTGTTACATCCGGGACCTACTCACAGTACGGCCACATC
 35 GGCATCGTGACAGTCCAGGAGATACCAACAAATTCCATATCATCGAGCAGAATTGGAATGGATGG
 GCAAACAAGAAACCTCTGACACGTTGGGACAATACTACGGGCTGACCCACTTTATCGAGGTCCCAT
 ATGCGAAAGAGACAATCACAACGAAGAAAAGTGCACCTAAAGCAGCTACCTCAAAACCGGCTAAGA
 AGATCTCAGTTTCTACAACAACATCAACCTTCAGAGCAACTTTGCAAAGAGAGGGTACAACCCAAA
 GGGAGTGGTGATTACAATGACGCCGACGCTCTACTGCGCAACAGTATGAAAACAGCCTGAAGGG
 40 CGCTAATCAGAACAGGCTGAATAACGGCATTGCCCATGCATATGCTGACAGCAGCCACATTTGGTAT
 GCTATCGAACCCGAGAGAATCGCCTATCACACAGGCGATGGCGTTAAGAAGGGAACAGGCAATTAC
 GACTACTTCGGTATTGAAGTATGCCAGAGCATGAGCGCAAGTGACAAAGAGTTCATGGCCAACGAA
 CAGGTGGTACTGCAGGAGGTGGCTAGGTTGCTCAAGAAATGGAAGCTTCCCGCAATCGGAATACC
 GTGCGGCTCCATCTGGAATTCTCTGCCACATCTTGCCTCATAGGAGCCTACAATGCGGACCGGTTT
 45 CAATCCGGTGACTCAGGGTAGACCCTCAGACGCCATCATCAAGAAGTTGAAGGATATCTGGATAAAA
 GAGATTGCGCATATATGGAGGGTAAACCCCATCCATTAAGACTGTCGCCAATAAGCCCGGATCAG
 CCAGTACCCCTGCAACCAATCCAACCTCCGATGGGTGGAAGGTCAATGCCTACAATACCTACTACAA

AACGGAACGGGCTCCTTTACAGTGGGCTCTACTCGGATTGCTGTGCGTAGTACAGGACCCTTTACT
 ACATGCCCCTTTACAGGGTGGCTGAACCCCGGAGTCACGATAACCTATGACGAGGTTATGGTGCAG
 AACGGGTTTGTGTGGTTGGGCTATGATAGCTACAATGGGAGGAGGTATGTCCCATTCGCACTGCCC
 AAGGAACTCCTGGCAGCTCCAATTATGTGCTTGCCCTCTGTGGGGCACGATTTCTGCGGCCGCCG
 5 AGGAGGCGGCTCCACCACCACCACCATCATTGA

Construct 20: pcDNA3.1(+)-SS(hINS)-FLAG-hL0482-His

Aa sequence

10 MALWMRLLPLLALLALWGPDPAAAMKDYKDDDDKGGGSGREAPKTQIYTVKKGDTLSAIALKYKTTV
 SNIQNTNNIANPNLIFIGQKLKVPMTPLVEPKPTVSSNNKSNSNSTLNYLKTLENRGWDFDGSYGWQC
 FDLNVVYWNHLYGHGLKGYGAKDIPYANNFNSEAKIYHNTPTFKAEPGDLVVFSGRFGGGYGHTAIVLN
 GDYDGKLMKFQSLDQNWNGGWRKA EVAHKV VHN YENDMIFIRPFKAAAAGGGGSHHHHHH

Nt sequence

GAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGCCCT
 GTGGATGAGACTGCTGCCTCTGCTTGTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATG
 20 AAGGACTACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAAGAGAGGCTCCCAAAACACA
 GATTACACCGTCAAGAAGGGAGATACGCTTAGCGCCATTGCACTGAAATACAAGACCACCGTATCA
 AACATCCAGAACCAATAACATAGCCAATCCAATCTCATCTTCATTGGCCAGAAGCTCAAGGTTCC
 CATGACTCCACTGGTCAACCGAAACCTAAGACAGTGTCTCCAATAACAAGAGTAACAGCAACTCC
 TCTACACTGAACTATCTCAAGACGTTGGAGAATCGCGGTTGGGACTTTGACGGCAGTTACGGCTGGC
 25 AATGCTTCGACTTGGTGAATGTGTACTGGAATCACCTGTATGGGCATGGACTGAAAGGCTATGGGGC
 CAAAGACATACCTTATGCTAACAATTTCAACTCTGAGGCCAAGATCTACCACAACACTCCTACCTCAA
 GGCAGAACCAGGAGATCTGGTTGTGTTAGCGGACGGTTTGGCGGAGGCTATGGGCATACTGCCAT
 TGTGCTTAATGGGGATTACGATGGCAAACCTGATGAAATTCAGTCACTGGACCAGAACTGGAACAAT
 GGTGGTTGGCGAAAAGCGGAAGTAGCCACAAAGGTCGTCACAATTACGAGAACGACATGATCTTC
 30 ATCCGTCCTTTAAGAAAGCGGCGGCCGCCGGAGGAGGCGGCTCCACCACCACCACCATCATTGA

Construct 21: pcDNA3.1(+)-SS(hINS)-FLAG-hL0483-His

Aa sequence:

35 MALWMRLLPLLALLALWGPDPAAAMKDYKDDDDKGGGSGAATHEHSAQWLNNYKKGYGYPYPL
 GINGGMHYGVDFFMNIGTPVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVVKVGD
 YVKAGQIIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPLKSAGYGKAGGTVPPTPNTGWKTNKY
 GTLYKSESASFTPNTDIITRTTGPFPSMPQSGVLKAGQTIHYDEVMMKQDGHVWVGTYGNSGQRIYLPVRT
 WNKSTNTLGVLWGTIKAAAAGGGGSHHHHHH

Nt sequence

GAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGCCCT
 GTGGATGAGACTGCTGCCTCTGCTTGTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATG
 45 AAGGACTACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAGCAGCAACGCATGAGCATT
 CGCTCAATGGCTCAACAACATAAGAAGGGTTACGGCTATGGCCCTATCCTCTGGGGATAAACGGA
 GGGATGCACTATGGAGTCGATTTCTTCATGAATATTGGCACTCCAGTGAAAGCGATTTCCAGCGGGA

AGATTGTGGAAGCTGGGTGGAGCAACTATGGTGGTGGCAATCAGATCGGACTGATTGAGAATGACG
GCGTGACAGACAGTGGTATATGCATCTGAGTAAGTACAACGTCAAGGTGGGAGACTACGTCAAAG
CCGGGCAGATCATAGGATGGTCCGGCTCTACAGTTACTCAACAGCGCCACATCTCCACTTTCAGCG
CATGGTGAATAGCTTCAGCAATTCAACCGCCCAAGATCCGATGCCCTTCTGAAGTCTGCCGGCTATG
5 GGAAAGCTGGAGGAACAGTGACACCTACTCCAATACCGGATGGAAAACGAACAAGTACGGCACTC
TGTACAAATCCGAATCTGCCAGTTTCACACCCAATACCGACATCATCACCCGAACCACTGGACCTTTTC
GGTCAATGCCGCAGTCTGGAGTACTGAAAGCAGGGCAAACCTATCCACTATGACGAGGTGATGAAGC
AGGATGGCCACGTATGGGTGGGGTACACAGGCAATAGCGGCCAGAGGATTTACTTGCCAGTCCGTA
CCTGGAACAAAAGCACGAACACTCTTGCGTTTTGTGGGGCACCATCAAGGCGGCCGCCGGAGGAG
10 GCGGCTCCCAACCACCACCACCATCATTGA

Construct 22: pcDNA3.1(+)-SS(hINS)-FLAG-hL0484m-His

Aa sequence

15 MALWMRLLPLLALLALWGPDPAAAMKDYKDDDDKGGGGSGDQFLDEKEGNHTTAYRDGSGIWTICR
GATMVDGKPIPGMRLSKEKCDQVNAIERDKALAWVERNIVPLTEPQKAGIASFCPYNIGPGKCFPSTF
YKRLNAGDRKGACEAIRWWIKDGGDRCRIRSNNCYGQVIRRDQESALACWGIDQAAAGGGGSHHHHH
H

20

Nt sequence

GAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGCCCT
GTGGATGAGACTGCTGCCTCTGCTTGTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATG
25 AAGGACTACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAGATCAGTTTCTCGACGAGAA
GGAAGGCAATCACACCACTGCGTATAGGGATGGCTCTGGGATATGGACCATCTGTAGAGGCGCCAC
AATGGTCGATGGCAAACCCGTGATACCCGGAATGCGTCTGTCAAAGGAGAAGTGGGATCAGGTGAA
TGCCATCGAACGCGACAAAGCTCTGGCCTGGGTTGAGCGGAACATCAAGGTTCCACTGACGGAACC
ACAGAAAGCCGGAATTGCGTCTTCTGTCCGTACAACATTGGCCCTGGAAAGTGCTTTCCCTCCACTT
30 TCTACAAGCGCCTGAATGCAGGGGACAGAAAAGGAGCCTGTGAGGCTATCAGGTGGTGGATCAA
GATGGAGGTAGGGACTGTGAATCAGGAGCAACAACCTGCTATGGGCAAGTGATTCCGAGAGATCA
GGAGAGTGCTCTGGCTTGTGGGTATTGACCAGGCGGCCGCCGGAGGAGGCGGCTCCCAACCACCA
CCACCATCATTGA

35 Construct 23: pcDNA3.1(+)-SS(hINS)-FLAG-hIL2-His

Aa sequence

40 MALWMRLLPLLALLALWGPDPAAAMKDYKDDDDKGGGGSGAPTSSSTKKTQLQLEHLLDLQMILNGI
NNYKNPKLTRMLTFKFYMPKKATELKHLCLEELKPLEEVLNLAQSKNFHLRPRDLISNINIVIVLELKGSET
TFMCEYADETATIVEFLNRWITFCQSIISTLTAAAGGGGSHHHHHH*

Nt sequence:

45 GAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGCCCT
GTGGATGAGACTGCTGCCTCTGCTTGTCTGCTGGCACTGTGGGGACCTGATCCTGCTGCTGCTATG
AAGGACTACAAAGACGATGACGACAAGGGAGGAGGCGGCTCCGGAGCTCCAACATCTTCTCAACA

AAGAAAACACAGCTTCAGCTTGAACACCTTCTTCTTGATCTTCAGATGATTCTGAATGGAATCAACAA
 TTACAAAAATCCAAACTGACAAGAATGCTGACATTTAAATTTTACATGCCAAAGAAAGCAACAGAA
 CTGAAACACCTTCAGTGCCTTGAAGAAGAACTGAAACCTCTGGAAGAAGTCTGAATCTGGCTCAGA
 GCAAAAATTTTACCTGAGACCAAGAGATCTGATCAGCAACATCAATGTGATTGTGCTGGAAGTGA
 5 AGGATCTGAAACAACATTCATGTGTGAATATGCTGATGAAACAGCAACAATTGTGGAATTTCTGAAC
 AGATGGATCACATTTTGCCAGTCAATCATTTCAACACTGACAGCGGCCGCCGAGGAGGCGGCTCCC
 ACCACCACCACCATCATTGA

Construct 24: pcDNA3.1(+)-SS(hIL2)-FLAG-hIL2-His

10

Aa sequence

MYRMQLLSIALSLALVTNSMKDYKDDDDKGGGGSGAPTSSSTKKTQLQLEHLLLDLQMLNGINNYKN
 PKLTRMLTFKFYMPKKATELKHLCLEELKPLEEVLNLAQSKNFHLRPRDLISINIVIVLELKGSETTFMCE
 15 YADETATIVEFLNRWITFCQSIISTLTAAGGGGSHHHHHH

Nt sequence

GAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGTACCG
 20 GATGCAGCTGCTGAGCTGTATCGCCCTGTCTCTGGCCCTGGTCACCAATTCTATGAAGGACTACAAA
 GACGATGACGACAAGGGAGGAGGCGGCTCCGGAGCTCCAACATCTTCTTCAACAAAGAAAACACAG
 CTTGAGCTTGAACACCTTCTTCTTGATCTTCAGATGATTCTGAATGGAATCAACAATTACAAAAATCCA
 AACTGACAAGAATGCTGACATTTAAATTTTACATGCCAAAGAAAGCAACAGAACTGAAACACCTTC
 AGTGCCTTGAAGAAGAACTGAAACCTCTGGAAGAAGTCTGAATCTGGCTCAGAGCAAAAATTTTCA
 25 CCTGAGACCAAGAGATCTGATCAGCAACATCAATGTGATTGTGCTGGAAGTGAAGGATCTGAAACA
 ACATTCATGTGTGAATATGCTGATGAAACAGCAACAATTGTGGAATTTCTGAACAGATGGATCACAT
 TTTGCCAGTCAATCATTTCAACACTGACAGCGGCCGCCGAGGAGGCGGCTCCCAACCACCACCACCA
 TCATTGA

30 Construct 25: PL084_pST4-T7-AGA-dEarl-hAg-Kozak-SS(IGkappa)-His-3C-hL0483(PM)-FLAG-FI-A30LA70

Aa sequence:

35 MDMRAPAGIFGFLVLFPGYRSHHHHHHLEVLFQGPAA THEHSAQWLNNYKKG YGYGPYPLGINGGM
 HYGVDFFMNIGTPVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVVKVGDYVKAGQI
 IGWSGSTGYSTAPHLHFQRMVNSFSQSTAQDPMFLKSAGYKAGGTVPTPNTGWKTNKYGTLYKSE
 SASFTPNTDIITRTTGPFRSMPQSGVLKAGQTIHYDEVMMKQDGHVWVG YTGNSGQRIYLPVRTWQKST
 NTLGVLVWGTIKSGGSDYKDDDDK*

40

Nt sequence:

AGAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACCATGGACATGCGGGC
 CCCTGCCGGAATCTTCGGCTTTCTGCTGGTGTGTTCCCGGCTACAGAAGCCATCATCATCACCACC
 45 ACCTCGAAGTATTGTTTCAGGGACCCGCCGCCACCCACGAGCACAGCGCTCAATGGCTGAACAAC
 TCAAAAAGGGCTACGGCTATGGTCCTTATCCTCTGGGCATCAACGGCGGGATGCACTACGGCGTGG
 CTTCTTCATGAACATCGGCACCCCGTGAAGGCCATCAGCAGCGGAAAGATCGTGAAGCCGGCTG

GTCCAACACGGCGGGCAACAGATCGGACTGATCGAGAACGACGGCGTGACAGACAGTGGTA
 CATGCATCTGTCTAAGTACAATGTGAAGGTGGGCGACTACGTGAAAGCCGGCCAAATCATTGGATG
 GAGCGGCTCTACCGGTACAGCACAGCCCCACACCTGCACTTCCAGAGAATGGTGAACAGCTTCAGC
 CAGTCCACAGCCCAGGACCCCATGCCTTCTGAAGAGCGCCGGCTACGGCAAGGCCGGAGGCACC
 5 GTGACCCCAACACCTAATACCGGCTGGAAGACCAACAAGTACGGCACCTGTACAAGTCTGAAAGCG
 CTTCTTTTACCCTAACACTGATATCATCACAAGAACCACCGGCCCTTTAGAAGCATGCCCCAGAGC
 GGCGTCTCAAGCTGGACAGACCATCCACTACGACGAGGTGATGAAACAGGATGGCCACGTGTGG
 GTCGGCTACACCGGAAATAGCGGCCAGCGGATCTACCTGCCTGTGCGGACATGGCAGAAGTCCACC
 AACACACTGGGAGTTCTGTGGGGCACAATCAAGTCAGGAGGTTAGATTACAAGGATGACGACGAT
 10 AAGTGAICTGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCCCTTCCCGTCTGGGTACCCGA
 GTCTCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTT
 CCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACCCCCACGGGAA
 ACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCCAGGGTTGGTCA
 ATTCGTGCCAGCCACACCTGGAGCTAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATAT
 15 GACTAAA
 AAAAAAAAAA

Endolysin expression was seen for DNA as well as for mRNA transfected HEK293T/17 cells. A
 dosing and time course experiment of a prototype endolysin expressed from an mRNA
 20 construct revealed that expression was dose, as well as time dependent (Figure 6A and 6B).
 The highest expression levels from mRNA were in the range of 2 µg/ml.

For the determination of pharmacokinetic parameters in mice, an LNP-formulated mRNA was
 generated and tested for expression in mammalian cells. Transfection of 1.5 µg of LNP-mRNA
 led to detectable lysis levels in the cell supernatants. The levels were determined to be in the
 25 ng/ml range (Figure 6C+D).

Expression of lysis variants from mRNA or DNA transfections is shown in Figure 7. 5 µg of
 mRNA or 2.5 µg of DNA were transfected into HEK293T/17 cells and supernatants were
 analyzed by Western Blot, demonstrating expression from DNA as well as from mRNA for
 different lysis variants.

30

Example 10: Bactericidal activity of lysins encoded in and secreted from HEK293T/17 cells

Supernatants from DNA or RNA transfected cells are concentrated (5 fold) using VivaSpin 500
 concentrators (Sartorius VS0102) and tested for bactericidal activity in a 96 well plate. For this
 purpose, 60 µl of supernatants (concentrated or not concentrated) are mixed with *S. aureus*
 35 ATCC® 43300™ in 40 µl cation-adjusted Müller Hinton broth (Merck 90922), supplemented
 with 10% horse serum (Thermo Scientific 26050070), such that the reaction contains 5x10⁵
 cfu/ml. The optical density at 600 nm (OD_{600nm}) of these solutions is monitored in a plate

reader (TECAN Infinite 200) over the course of 24 hours to quantify bacterial growth. Bacterial growth, or the inhibition thereof, reveals the levels of active lysin in the supernatant tested. Supernatants of mock transfected HEK293T/17 cells serve as a control to rule out an effect of the culture medium on lytic activity.

- 5 When testing the non-concentrated supernatants, a clear delay in outgrowth of staphylococci can be observed, meaning that bacterial killing occurs but due to the low concentration of lysin complete eradication is not achieved. The complete inhibition of staphylococcal outgrowth indicates that the level of expressed and active lysins is sufficiently high in the concentrated supernatant to lyse all *S. aureus* bacteria in the culture (Figure 8 A-D).
- 10 The enzymatic activity of RNA expressed lysins secreted in HEK293 supernatants was also monitored using a dye release assay. Crude peptidoglycan extract, obtained by lysis of staphylococci through heat shock, was labelled with remazol brilliant blue (RBB, Sigma, R8001-25G) (Farris et al, J Vis Exp, 2016). After 5 hours of incubation of this substrate with the lysin to be tested, the reaction is stopped by the addition of ethanol, centrifuged down and
- 15 enzymatic activity is assessed by measuring the release of the RBB dye at OD₅₉₅ nm. As a positive control, recombinant lysin was spiked into non lipofected supernatant at 3 µg/ml. Non-lipofected supernatant and supernatant from a luciferase mRNA LNP transfection served as negative controls. In Figure 8E, enzymatic activity is seen for L0483 mRNA-LNP supernatants. The lower levels of activity compared to the spiked recombinant protein, are
- 20 due to the lower concentration in the supernatant.

These experiments reveal DNA and mRNA-encoded lysins which can be expressed and secreted in an active form in human cells resulting in successful lysis of *S. aureus* (Figure 8).

25 **Example 11: Pharmacokinetics and efficacy of mRNA expressed lysin and recombinant lysin in vivo**

To compare the *in vivo* pharmacokinetics of an mRNA-delivered lysin and a recombinantly expressed and administered lysin, mice were injected with lipid-nanoparticle (LNP) formulated mRNA and 2 different doses of recombinant lysin. Expression levels were determined in mouse plasma by quantitation in a sandwich ELISA.

- 30 Specific-pathogen-free male CD1 mice (Charles River) of 11-15g at arrival, were left to acclimatize for at least 7 days prior to a single bolus application of different doses of LNP-

mRNA or recombinant protein. 5 mice were administered with a dose of 30 µg LNP-mRNA and two groups of 3 mice each with recombinant lysin (535µg and 107 µg per mouse, respectively). A blood microsample (~20 µl) was collected from each mouse by tail vein puncture into a capillary and separated for plasma collection at timepoints 1, 2, 4, 8, 12, 24, 48 and 72 hours post application for the LNP-mRNA groups and 0.08, 0.25, 0.5, 1, 2, 4, 8, 24 hours post application for the recombinant lysin. The plasma samples were then analyzed for their lysin content using a sandwich ELISA. Briefly, a rabbit antibody against the HRV3C cleavage site (Abcam PA1-118) was used as capturing antibody and a HRP coupled anti-flag antibody (Thermo Fisher A8592-.2MG) was used as detection antibody.

The analyses of the plasma samples after administration of recombinant protein revealed a poor PK profile for both doses tested (Figure 9A). Lysin in plasma could be detected for only 2 hours post treatment in the high dose; for the low dose, no lysin could be detected 30 min after administration. An important parameter for efficacy of treatment is the time the plasma concentration of lysin above MIC (minimum inhibitory concentration determined to be 250 ng/ml for this lysin). In the high dose group, time above MIC was approximately 100 min, in the low dose group time above MIC was only 30 min.

The profile was very different for the LNP-mRNA administered lysin (Figure 9B). C_{max}, ranged between 787 and 219 ng/ml depending on the mouse, hence lower than determined for the recombinant protein. T_{max} ranged between 4 and 12 hours depending on the mouse (Table 7). The AUC values were comparable in the mRNA groups and in the recombinant protein groups. For maximal efficacy, an endolysin plasma level above MIC is desired, which was seen for 4 of the 5 mice and time over MIC was over 24 hours. Although C_{max} was lower when mRNA was applied, time over MIC was much longer, which we postulate will lead to better efficacy *in vivo*.

Table 7: PK parameters determined for recombinant lysin application and mRNA-LNP application

	30 µg mRNA dose	535 µg protein dose	107 µg protein dose
C_{max} (ng/ml)	219-787	123294	19672
T_{max} (h)	4-12	0.08	0.08
AUC (ng*h*ml⁻¹)	1247-13166	19141	2085

Time over MIC (h)	24-48	1.5	0.5
-------------------	-------	-----	-----

The PK data generated with LNP-mRNA clearly shows good bioavailability of the lysin when administered as mRNA. In previous experiments, these PK parameters were determined for a lysin administered i.v. as recombinant protein. In parallel to assessing the bioavailability of the recombinant protein, an *in vivo* efficacy study was performed, with bacterial burden in different tissues serving as readout.

Specific-pathogen-free male CD1 mice (Charles River) of 11-15g at arrival, were left to acclimatize for at least 7 days prior to IV administration of 5×10^7 cfu of bacteria (*S. aureus* NCTC8189). One hour post-infection, non-treated control mice were euthanized and the dose groups (5 mice per group) were as follows: 535 μ g per mouse, 107 μ g per mouse, Vancomycin at 25 mg/kg and a mock-treatment control (vehicle only), all administered i.v. q12h. Mice were euthanized 24 hours post treatment and bacterial burden in blood, heart and kidney was determined by quantitative culture on plates.

The PK data did not look promising, especially the rapid decline in plasma concentration and the time over MIC being very short, nevertheless efficacy was tested using a doses of 535 μ g or 107 μ g per mouse. Interestingly, despite the suboptimal bioavailability and rapid elimination of the lysin from plasma, efficacy could be seen. Bacterial burden was decreased compared to vehicle only in blood as well as in heart and kidneys and in the range of vancomycin efficacy (Figure 10). No dose dependent effect of the lysin was observed. These results clearly show that the lysin is efficacious in this model, even at very low plasma concentrations.

The PK data generated for the LNP-mRNA administered lysin demonstrates longer time over MIC, indicating longer bioavailability than the recombinant version of the lysin. Rotolo et al (poster ASM microbe 2016) observed that a ratio of AUC/MIC above 1 was required for maximal efficacy (>3-log reduction in bacterial burden in a mouse thigh infection model). This ratio was also taken as basis for dose selection in a phase 2 clinical study of an anti-staphylococcal lysin (Fowler et al, J. Clin Invest, 2020), where positive responses were seen. Taking these results together, we postulate that the LNP-mRNA administration of the lysin should lead to an even higher efficacy than seen for the recombinant protein.

30

Claims

1. A composition or medical preparation comprising RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase.
2. The composition or medical preparation of claim 1, wherein the peptidoglycan hydrolase breaks down peptidoglycan in bacterial cell wall.
3. The composition or medical preparation of claim 1 or 2, wherein the peptidoglycan hydrolase is derived from a bacteriophage.
4. The composition or medical preparation of any one of claims 1 to 3, wherein the peptidoglycan hydrolase is or is derived from an endolysin.
5. The composition or medical preparation of any one of claims 1 to 4, wherein the peptidoglycan hydrolase is modified so as to reduce glycosylation.
6. The composition or medical preparation of any one of claims 1 to 5, wherein the peptidoglycan hydrolase is modified by removing glycosylation sites.
7. The composition or medical preparation of any one of claims 1 to 6, wherein the peptidoglycan hydrolase is modified so as to reduce immunogenicity.
8. The composition or medical preparation of any one of claims 1 to 7, wherein the peptidoglycan hydrolase is modified by removing T cell epitopes.
9. The composition or medical preparation of any one of claims 1 to 8, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1.

10. The composition or medical preparation of any one of claims 1 to 9, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 2, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2, or a fragment of the nucleotide sequence of SEQ ID NO: 2 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2.

11. The composition or medical preparation of any one of claims 1 to 9, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 3, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, or a fragment of the nucleotide sequence of SEQ ID NO: 3 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3.

12. The composition or medical preparation of any one of claims 1 to 11, wherein the peptidoglycan hydrolase is secreted peptidoglycan hydrolase.

13. The composition or medical preparation of any one of claims 1 to 12, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises a signal peptide.

14. The composition or medical preparation of any one of claims 1 to 13, wherein the peptidoglycan hydrolase is fused to a signal peptide.

15. The composition or medical preparation of claim 13 or 14, wherein the signal peptide is cleaved off during secretion or export.

16. The composition or medical preparation of any one of claims 13 to 15, wherein the signal peptide comprises the amino acid sequence of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of

SEQ ID NO: 10, or a fragment of the amino acid sequence of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10.

17. The composition or medical preparation of any one of claims 1 to 16, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4.

18. The composition or medical preparation of any one of claims 1 to 17, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 5, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, or a fragment of the nucleotide sequence of SEQ ID NO: 5 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5.

19. The composition or medical preparation of any one of claims 1 to 18, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises an extended pharmacokinetic (PK) polypeptide.

20. The composition or medical preparation of any one of claims 1 to 19, wherein the peptidoglycan hydrolase is fused to a pharmacokinetic modifying group.

21. The composition or medical preparation of claim 20, wherein the pharmacokinetic modifying group comprises a moiety which is heterologous to the peptidoglycan hydrolase.

22. The composition or medical preparation of claim 20 or 21, wherein the pharmacokinetic modifying group comprises a moiety selected from the group consisting of albumin, an immunoglobulin fragment, transferrin, Fn3, a functional variant thereof, or a functional

fragment of the albumin, immunoglobulin fragment, transferrin, Fn3 or the functional variant thereof.

23. The composition or medical preparation of any one of claims 20 to 22, wherein the pharmacokinetic modifying group comprises human albumin, a functional variant thereof, or a functional fragment of the human albumin or the functional variant thereof.

24. The composition or medical preparation of any one of claims 1 to 23, wherein the amino acid sequence comprising a peptidoglycan hydrolase is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence, wherein the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

25. The composition or medical preparation of any one of claims 1 to 24, wherein the RNA comprises a modified nucleoside in place of uridine.

26. The composition or medical preparation of any one of claims 1 to 25, wherein the RNA comprises a modified nucleoside in place of each uridine.

27. The composition or medical preparation of claim 25 or 26, wherein the modified nucleoside is selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1 ψ), and 5-methyl-uridine (m5U).

28. The composition or medical preparation of any one of claims 25 to 27, wherein the modified nucleoside is N1-methyl-pseudouridine (m1 ψ).

29. The composition or medical preparation of any one of claims 1 to 28, wherein the RNA comprises a cap0 5' cap.

30. The composition or medical preparation of any one of claims 1 to 29, wherein the RNA comprises the 5' cap m₂^{7,2'}O^{3'}G(5')ppSp(5')G (in particular its D1 diastereomer).

31. The composition or medical preparation of any one of claims 1 to 28, wherein the RNA comprises a cap1 5' cap.
32. The composition or medical preparation of any one of claims 1 to 28 and 31, wherein the RNA comprises the 5' cap $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$.
33. The composition or medical preparation of any one of claims 1 to 32, wherein the RNA comprises a 5' UTR.
34. The composition or medical preparation of claim 33, wherein the 5' UTR comprises the nucleotide sequence of SEQ ID NO: 11, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11.
35. The composition or medical preparation of any one of claims 1 to 34, wherein the RNA comprises a 3' UTR.
36. The composition or medical preparation of claim 35, wherein the 3' UTR comprises the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12.
37. The composition or medical preparation of any one of claims 1 to 36, wherein the RNA comprises a poly-A sequence.
38. The composition or medical preparation of claim 37, wherein the poly-A sequence comprises at least 100 nucleotides.
39. The composition or medical preparation of claim 37 or 38, wherein the poly-A sequence comprises or consists of the nucleotide sequence of SEQ ID NO: 13, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 13.

40. The composition or medical preparation of any one of claims 1 to 39, wherein the RNA is mRNA.

41. The composition or medical preparation of any one of claims 1 to 40, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 6, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence of SEQ ID NO: 6 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6.

42. The composition or medical preparation of any one of claims 1 to 40, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 7, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7, or a fragment of the nucleotide sequence of SEQ ID NO: 7 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7.

43. The composition or medical preparation of any one of claims 1 to 39, wherein the RNA is self-amplifying RNA.

44. The composition or medical preparation of any one of claims 1 to 39 and 43, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 8, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8, or a fragment of the nucleotide sequence of SEQ ID NO: 8 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8.

45. The composition or medical preparation of any one of claims 1 to 39 and 43, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the

nucleotide sequence of SEQ ID NO: 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9, or a fragment of the nucleotide sequence of SEQ ID NO: 9 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9.

46. The composition or medical preparation of any one of claims 1 to 45, wherein the RNA is formulated as a liquid, formulated as a solid, or a combination thereof.

47. The composition or medical preparation of any one of claims 1 to 46, wherein the RNA is formulated for injection.

48. The composition or medical preparation of any one of claims 1 to 47, wherein the RNA is formulated for intravenous administration.

49. The composition or medical preparation of any one of claims 1 to 48, wherein the RNA is formulated or is to be formulated as lipid particles.

50. The composition or medical preparation of claim 49, wherein the RNA lipid particles are lipid nanoparticles (LNP).

51. The composition or medical preparation of any one of claims 1 to 50, which is a pharmaceutical composition.

52. The composition or medical preparation of claim 51, wherein the pharmaceutical composition further comprises one or more pharmaceutically acceptable carriers, diluents and/or excipients.

53. The composition or medical preparation of any one of claims 1 to 50, wherein the medical preparation is a kit.

54. The composition or medical preparation of claim 53, further comprising instructions for use of the RNA for treating or preventing bacterial infection.

55. The composition or medical preparation of any one of claims 1 to 54, which is for administration to a human.

56. The composition or medical preparation of any one of claims 1 to 55 for pharmaceutical use.

57. The composition or medical preparation of claim 56, wherein the pharmaceutical use comprises a therapeutic or prophylactic treatment of a disease or disorder.

58. The composition or medical preparation of claim 57, wherein the therapeutic or prophylactic treatment of a disease or disorder comprises treating or preventing bacterial infection.

59. A method of treating bacterial infection in a subject comprising administering RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase to the subject.

60. The method of claim 59, wherein the peptidoglycan hydrolase breaks down peptidoglycan in bacterial cell wall.

61. The method of claim 59 or 60, wherein the peptidoglycan hydrolase is derived from a bacteriophage.

62. The method of any one of claims 59 to 61, wherein the peptidoglycan hydrolase is or is derived from an endolysin.

63. The method of any one of claims 59 to 62, wherein the peptidoglycan hydrolase is modified so as to reduce glycosylation.

64. The method of any one of claims 59 to 63, wherein the peptidoglycan hydrolase is modified by removing glycosylation sites.

65. The method of any one of claims 59 to 64, wherein the peptidoglycan hydrolase is modified so as to reduce immunogenicity.

66. The method of any one of claims 59 to 65, wherein the peptidoglycan hydrolase is modified by removing T cell epitopes.

67. The method of any one of claims 59 to 66, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1, or a fragment of the amino acid sequence of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1.

68. The method of any one of claims 59 to 67, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 2, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2, or a fragment of the nucleotide sequence of SEQ ID NO: 2 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2.

69. The method of any one of claims 59 to 67, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 3, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, or a fragment of the nucleotide sequence of SEQ ID NO: 3 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3.

70. The method of any one of claims 59 to 69, wherein the peptidoglycan hydrolase is secreted peptidoglycan hydrolase.

71. The method of any one of claims 59 to 70, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises a signal peptide.

72. The method of any one of claims 59 to 71, wherein the peptidoglycan hydrolase is fused to a signal peptide.

73. The method of claim 71 or 72, wherein the signal peptide is cleaved off during secretion or export.

74. The method of any one of claims 71 to 73, wherein the signal peptide comprises the amino acid sequence of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10, or a fragment of the amino acid sequence of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10.

75. The method of any one of claims 59 to 74, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises the amino acid sequence of SEQ ID NO: 4, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4, or a fragment of the amino acid sequence of SEQ ID NO: 4, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 4.

76. The method of any one of claims 59 to 74, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 5, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, or a fragment of the nucleotide sequence of SEQ ID NO: 5 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5.

77. The method of any one of claims 59 to 76, wherein the amino acid sequence comprising a peptidoglycan hydrolase comprises an extended pharmacokinetic (PK) polypeptide.

78. The method of any one of claims 59 to 77, wherein the peptidoglycan hydrolase is fused to a pharmacokinetic modifying group.

79. The method of claim 78, wherein the pharmacokinetic modifying group comprises a moiety which is heterologous to the peptidoglycan hydrolase.

80. The method of claim 78 or 79, wherein the pharmacokinetic modifying group comprises a moiety selected from the group consisting of albumin, an immunoglobulin fragment, transferrin, Fn3, a functional variant thereof, or a functional fragment of the albumin, immunoglobulin fragment, transferrin, Fn3 or the functional variant thereof.

81. The method of any one of claims 78 to 80, wherein the pharmacokinetic modifying group comprises human albumin, a functional variant thereof, or a functional fragment of the human albumin or the functional variant thereof.

82. The method of any one of claims 59 to 81, wherein the amino acid sequence comprising a peptidoglycan hydrolase is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence, wherein the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

83. The method of any one of claims 59 to 82, wherein the RNA comprises a modified nucleoside in place of uridine.

84. The method of any one of claims 59 to 83, wherein the RNA comprises a modified nucleoside in place of each uridine.

85. The method of claim 83 or 84, wherein the modified nucleoside is selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1 ψ), and 5-methyl-uridine (m5U).

86. The method of any one of claims 83 to 85, wherein the modified nucleoside is N1-methyl-pseudouridine (m1ψ).

87. The method of any one of claims 59 to 86, wherein the RNA comprises a cap0 5' cap.

88. The method of any one of claims 59 to 87, wherein the RNA comprises the 5' cap m₂^{7,2'-O}G(5')ppSp(5')G (in particular its D1 diastereomer).

89. The method of any one of claims 59 to 86, wherein the RNA comprises a cap1 5' cap.

90. The method of any one of claims 59 to 86 and 89, wherein the RNA comprises the 5' cap m₂^{7,3'-O}Gppp(m₁^{2'-O})ApG.

91. The method of any one of claims 59 to 90, wherein the RNA comprises a 5' UTR.

92. The method of claim 91, wherein the 5' UTR comprises the nucleotide sequence of SEQ ID NO: 11, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11.

93. The method of any one of claims 59 to 92, wherein the RNA comprises a 3' UTR.

94. The method of claim 93, wherein the 3' UTR comprises the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12.

95. The method of any one of claims 59 to 94, wherein the RNA comprises a poly-A sequence.

96. The method of claim 95, wherein the poly-A sequence comprises at least 100 nucleotides.

97. The method of claim 95 or 96, wherein the poly-A sequence comprises or consists of the nucleotide sequence of SEQ ID NO: 13, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 13.

98. The method of any one of claims 59 to 97, wherein the RNA is mRNA.

99. The method of any one of claims 59 to 98, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 6, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence of SEQ ID NO: 6 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6.

100. The method of any one of claims 59 to 98, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 7, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7, or a fragment of the nucleotide sequence of SEQ ID NO: 7 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7.

101. The method of any one of claims 59 to 97, wherein the RNA is self-amplifying RNA.

102. The method of any one of claims 59 to 97 and 101, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 8, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8, or a fragment of the nucleotide sequence of SEQ ID NO: 8 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 8.

103. The method of any one of claims 59 to 97 and 101, wherein the RNA encoding an amino acid sequence comprising a peptidoglycan hydrolase comprises the nucleotide sequence of SEQ ID NO: 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9, or a fragment of the nucleotide sequence of SEQ ID NO: 9 or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9.

104. The method of any one of claims 59 to 103, wherein the RNA is formulated as a liquid, formulated as a solid, or a combination thereof.

105. The method of any one of claims 59 to 104, wherein the RNA is administered by injection.

106. The method of any one of claims 59 to 105, wherein the RNA is administered by intravenous administration.

107. The method of any one of claims 59 to 106, wherein the RNA is formulated as lipid particles.

108. The method of claim 107, wherein the RNA lipid particles are lipid nanoparticles (LNP).

109. The method of any one of claims 59 to 108, wherein the RNA is formulated as a pharmaceutical composition.

110. The method of claim 109, wherein the pharmaceutical composition further comprises one or more pharmaceutically acceptable carriers, diluents and/or excipients.

111. The method of any one of claims 59 to 110, wherein the subject is a human.

112. The method of any one of claims 59 to 111, wherein the infection is a chronic infection.

113. The method of any one of claims 59 to 112, wherein the bacteria causing infection are multi-resistant bacteria.

114. The method of any one of claims 59 to 113, wherein the bacteria causing infection are Gram-negative bacteria, Gram-positive bacteria, or both.

115. The method of any one of claims 59 to 114, wherein the bacteria are selected from *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacteriaceae*, *Porphyromonas*

gingivalis, *Helicobacter pylori*, *Chlamydia trachomatis*, *Enterobacteriaceae*, *E. coli*, *Klebsiella pneumoniae*, *Salmonella*, *Shigella*, *Borellia*, *Campylobacter jejuni*, *Neisseria gonorrhoeae*, *Chlamydia trachomatis*, *Vibrio cholerae*, and *Fusobacterium nucleatum*.

116. The method of any one of claims 59 to 115, wherein the bacteria are resistant to carbapenem.

117. The method of any one of claims 59 to 114, wherein the bacteria are selected from *Staphylococci*, *Streptococci*, *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Clostridioides difficile*, *Cutibacterium acnes*, *Mycobacterium tuberculosis*, *Gardnerella*, *Enterococcus faecalis*, *Enterococcus faecium*, *Lactobacillus iners*, *Bacillus subtilis* and *Bacillus anthracis*.

118. The method of any one of claims 59 to 117, wherein the bacteria are resistant to methicillin and vancomycin or non-susceptible to penicillin.

119. The method of any one of claims 59 to 114, 117 and 118, wherein the bacteria are methicillin-resistant *Staphylococcus aureus* (MRSA).