

RNA FOR PREVENTING OR TREATING TUBERCULOSIS**Technical Field**

5 The disclosure provides agents and methods for preventing or treating tuberculosis using RNA. The RNA encoding antigens of *Mycobacterium tuberculosis*, immunogenic variants or fragments thereof is formulated and administered in a way that the antigens, variants or fragments are produced by cells of a subject, in particular after intramuscular or intravenous administration of the RNA.

Background

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

Tuberculosis (TB) is caused by the bacterial pathogen *Mycobacterium tuberculosis* (*Mtb*) and is the leading cause of death from a single infectious agent. *Mtb* is a gram-positive, rod-shaped bacterium from the Mycobacteriaceae family. 15 The more than 4,000 genes encoded within an approximately 4 million base pair genome render *Mtb* a complex pathogenic organism. This is further emphasized by the atypical composition of its cell wall, which has a high lipid content.

Despite the observed trend for reduction in TB cases and TB-related deaths for the last 20 years, 1.42 million people died from TB alone in 2019. In addition to the active form of TB, difficulties arise from latent TB infection (LTBI), when the infected patient doesn't present clinical symptoms. The estimated 2 billion latently infected individuals worldwide pose a huge and unpredictable reservoir of *Mtb* (WORLD HEALTH ORGANIZATION. Global tuberculosis report 2019. Geneva, WORLD HEALTH ORGANIZATION; 2019. ISBN: 978-92-4-156571-4). The high prevalence of HIV-1 infections further increases the risk for TB disease acquisition, activation of a latent TB infection, and death from HIV-TB 25 co-infection. In 2009, 0.2 million deaths were related to HIV-TB comorbidity. The complexity of the *Mtb* cell wall makes the bacterium resistant to environmental impact and to therapy with certain antibiotics. The latter further complicates anti-TB treatment especially in low- and middle-income countries (WORLD HEALTH ORGANIZATION. Global tuberculosis report 2020. Geneva, WORLD HEALTH ORGANIZATION; 2020. ISBN: 978-92-4-001313-1).

An attenuated strain of *Mycobacterium bovis*, bacillus Calmette-Guérin (BCG), is the only licensed TB vaccine, introduced in 1921. The use of the live vaccine BCG is not recommended for immunocompromised individuals and the protective efficacy against pulmonary TB conferred by immunization with BCG is highly variable, ranging from 50-80%. Moreover, passaging of BCG over the decades further attenuated the currently used BCG strains, reducing its protective efficacy (Brosch R, et al. Proc. Natl. Acad. Sci. U.S.A., 2007; 104(13):5596-5601). Thus, there is an unmet medical need for a safer and more effective vaccine to prevent TB, especially for a vaccine that can be administered to 35 immunocompromised individuals.

The pipeline of clinical trials for TB vaccine candidates comprises use of live, live-attenuated, and inactivated mycobacteria, and of *Mtb* antigens as recombinant protein (subunit vaccine) (TuBerculosis Vaccine Initiative (TBVI). Available from: <https://www.tbvi.eu/what-we-do/pipeline-of-vaccines/>). The drawbacks from these vaccine platforms are i. their low safety, due to replication-competent live vaccines still being infectious, ii. low immunogenicity of inactivated vaccines, and iii. the need for addition of adjuvants to subunit vaccines to enhance immunogenicity. To 40 date, most vaccine candidates have failed to demonstrate better protection from TB or from the development of TB compared to placebo in clinical trials.

For all these reasons, novel agents for preventing or treating tuberculosis are required.

Summary

The present disclosure provides compositions which are useful as TB vaccines. The compositions provided herein comprise RNA for delivering *Mtb* antigens to a subject. The findings described herein demonstrate that RNA described herein, e.g., non-modified uridine containing mRNA (uRNA) or nucleoside modified mRNA (modRNA), expressing antigens of *Mycobacterium tuberculosis*, immunogenic variants or fragments thereof, is useful for preventing or treating tuberculosis. The RNA encoding antigens of *Mycobacterium tuberculosis*, immunogenic variants or fragments thereof is formulated and administered in a way that the antigens, variants or fragments can be produced and preferably secreted by patient cells to prevent or combat tuberculosis.

In particular, the present disclosure describes in some embodiments RNA components encoding antigen 85 A (Ag85A), antigen *Mtb* 72F (M72; a recombinant protein derived from two *Mtb* proteins, Mtb32A and Mtb39A; hereafter indicated as M72 only), 6 kilodalton early secretory antigenic target (ESAT6), resuscitation-promoting factor A (RpfA), resuscitation-promoting factor D (RpfD), hypoxic response protein 1 (Hrp1), virulence associated protein B47 (VapB47), and heparin-binding hemagglutinin A (HbhA). The present disclosure observes that immunization of mice with these RNA components induced strong antigen-specific T- and B-cell responses, whereby the immune response elicited by two i.m. doses given 21 days apart was higher than that elicited by a single subcutaneous immunization with BCG.

Mtb displays differential gene expression patterns during its active and dormant (non-dividing) phases (Andersen P, et al. Cold Spring Harb Perspect Med, 2014; 4(6):a018523). To prevent development of TB, there should be immunity against antigens specific for each of the various stages of *Mtb* infection. The TB vaccine candidate developed here comprising the RNA components described above is designed to induce protective immune responses against antigens specific for different stages of *Mtb* infection.

Unlike the attenuated vaccine BCG, this TB vaccine candidate does not carry the risks associated with infection and may therefore be given to people who cannot be administered live organism (such as pregnant women and immunocompromised persons).

In one aspect, the disclosure provides a composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes a set of antigenic amino acid sequences, wherein the set of antigenic amino acid sequences comprises (i) at least one *Mtb* antigen from the acute phase of the *Mtb* life cycle, an immunogenic variant thereof, or an immunogenic fragment of the *Mtb* antigen or the immunogenic variant thereof, (ii) at least one *Mtb* antigen from the latent phase of the *Mtb* life cycle, an immunogenic variant thereof, or an immunogenic fragment of the *Mtb* antigen or the immunogenic variant thereof, and (iii) at least one *Mtb* antigen from the resuscitation phase of the *Mtb* life cycle, an immunogenic variant thereof, or an immunogenic fragment of the *Mtb* antigen or the immunogenic variant thereof.

In some embodiments, the at least one *Mtb* antigen from the acute phase of the *Mtb* life cycle, an immunogenic variant thereof, or an immunogenic fragment of the *Mtb* antigen or the immunogenic variant thereof comprises (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof; and/or (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof.

PepA (component of M72 fusion protein) has been reported to be expressed during the acute phase of *Mtb* infection. Accordingly, the at least one *Mtb* antigen from the acute phase of the *Mtb* life cycle, an immunogenic variant thereof, or an immunogenic fragment of the *Mtb* antigen or the immunogenic variant thereof comprises (optionally in addition to one or more of those described above) an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof.

In some embodiments, the at least one Mtb antigen from the latent phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof comprises

- (i) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof; and/or
- 5 (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof.

In some embodiments, the at least one Mtb antigen from the resuscitation phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof comprises

- (i) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof; and/or
- 10 (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof.

In one aspect, the disclosure provides a composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes a set of antigenic amino acid sequences, wherein each antigenic amino acid sequence comprises an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof and each RNA molecule encodes at least two of the antigenic amino acid sequences as fusion molecule.

In some embodiments, each RNA molecule encodes two of the antigenic amino acid sequences as fusion molecule.

In some embodiments, the Mtb antigens, immunogenic variants, or immunogenic fragments in a fusion molecule are not linked by a linker comprising a sequence which is heterologous to the Mtb antigens, immunogenic variants, or immunogenic fragments.

In some embodiments, the set of antigenic amino acid sequences comprises two or more, three or more, four or more, five or more, six or more, seven or more, or eight or more or all of the following:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- 25 (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- 30 (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- 35 (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof; and
- 40 (ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

In one aspect, the disclosure provides a composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes at least one antigenic amino acid sequence comprising an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof,

wherein the RNA encodes an amino acid sequence comprising a secretory signal peptide at the N-terminus of the encoded amino acid sequence and wherein the secretory signal peptide is not endogenous to the Mtb antigen. In some embodiments, the secretory signal peptide is of human origin. In some embodiments, the secretory signal peptide is not of human origin.

5 In some embodiments,

(i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44, or a functional fragment of the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44; and/or

10 (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43.

15 In some embodiments,

(i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, or a functional fragment of the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63; and/or

20 (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, or a fragment of the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62.

In some embodiments,

(i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, or a functional fragment of the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71; and/or

30 (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, or a fragment of the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70.

In one aspect, the disclosure provides a composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes at least one antigenic amino acid sequence comprising an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof, wherein the RNA comprises modified uridines and/or is formulated in lipid nanoparticles.

40 In some embodiments, the at least one RNA molecule encodes two or more, three or more, four or more, five or more, six or more, seven or more, or eight or more of the following:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- 5 (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- 10 (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- 15 (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof; and
- (ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

In some embodiments, the at least one RNA molecule encodes the following amino acid sequences:

- 20 (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- 25 (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- 30 (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof;
- 35 (ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

- In one aspect, the disclosure provides a composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all of the following amino acid sequences:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;

- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- 5 (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- 10 (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof; and
- (ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.
- 15 In some embodiments, the amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof and the amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof are present as a fusion protein.
- In some embodiments, the fusion protein comprises an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof.
- 20 In some embodiments, the at least one RNA molecule encodes two or more, three or more, four or more, five or more, six or more, seven or more, or all of the following amino acid sequences:
 - (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
 - 25 (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
 - (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
 - (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
 - 30 (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
 - (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
 - 35 (vii) an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof; and
 - (viii) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.
- In one aspect, the disclosure provides a composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes two or more, three or more, four or more, five or more, six or more, seven or more, or all of the following amino acid sequences:
 - 40 (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;

- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- 5 (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- 10 (vii) an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof; and
- (viii) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.
- 15 In some embodiments,
- (i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 298 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 positions 2 to 298 of SEQ ID NO: 20, or an immunogenic fragment of the amino acid sequence of positions 2 to 298 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 positions 2 to 298 of SEQ ID NO: 20; and/or
- 20 (ii) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, or a fragment of the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19.
- 25 In some embodiments,
- (i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 323 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 positions 27 to 323 of SEQ ID NO: 44, or an immunogenic fragment of the amino acid sequence of positions 27 to 323 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 positions 27 to 323 of SEQ ID NO: 44; and/or
- 30 (ii) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43.
- 35 In some embodiments,
- (i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 95 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
- 40

acid sequence of positions 2 to 95 of SEQ ID NO: 4, or an immunogenic fragment of the amino acid sequence of positions 2 to 95 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 positions 2 to 95 of SEQ ID NO: 4; and/or

- (ii) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, or a fragment of the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21.

In some embodiments,

- (i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, or an immunogenic fragment of the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46; and/or

- (ii) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, or a fragment of the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45.

In some embodiments,

- (i) the amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6, or an immunogenic fragment of the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6; and/or

- (ii) the RNA sequence encoding the amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, or a fragment of the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22.

In some embodiments,

- (i) the amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, or an immunogenic fragment of the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52; and/or

- (ii) the RNA sequence encoding the amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of

positions 2298 to 2591 of SEQ ID NO: 51, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, or a fragment of the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51.

In some embodiments,

- (i) the amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8, or an immunogenic fragment of the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, or a fragment of the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23.

In some embodiments,

- (i) the amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 324 to 465 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 positions 324 to 465 of SEQ ID NO: 44, or an immunogenic fragment of the amino acid sequence of positions 324 to 465 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 positions 324 to 465 of SEQ ID NO: 44; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43.

In some embodiments,

- (i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 407 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 positions 2 to 407 of SEQ ID NO: 10, or an immunogenic fragment of the amino acid sequence of positions 2 to 407 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 positions 2 to 407 of SEQ ID NO: 10; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, or a fragment of the nucleotide

sequence of positions 4 to 1221 of SEQ ID NO: 24, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24.

In some embodiments,

- (i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, or an immunogenic fragment of the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or a fragment of the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47.

In some embodiments,

- (i) the amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12, or an immunogenic fragment of the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, or a fragment of the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25.

In some embodiments,

- (i) the amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, or an immunogenic fragment of the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, or a fragment of the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45.

In some embodiments,

- (i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 723 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 positions 2 to 723 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence of positions 2 to 723 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 positions 2 to 723 of SEQ ID NO: 29; and/or

- 5 (ii) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, or a fragment of the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, or the nucleotide sequence having at least 99%, 98%, 97%, 96%,
10 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28.

In some embodiments,

- (i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino
15 acid sequence of positions 27 to 748 of SEQ ID NO: 52, or an immunogenic fragment of the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52; and/or

- (ii) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the nucleotide sequence of
20 positions 132 to 2297 of SEQ ID NO: 51, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, or a fragment of the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51.

In some embodiments,

- (i) the amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino
25 acid sequence of positions 2 to 199 of SEQ ID NO: 18, or an immunogenic fragment of the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18; and/or
30

- (ii) the RNA sequence encoding the amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of
35 positions 4 to 597 of SEQ ID NO: 30, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, or a fragment of the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30.

In some embodiments,

- (i) the amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 433 to 630 of SEQ ID
40 NO: 48, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, or an immunogenic fragment of the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, or a fragment of the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47.

In some embodiments, the composition or medical preparation comprises:

(i) an RNA molecule encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;

(ii) an RNA molecule encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;

(iii) an RNA molecule encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof; and

(iv) an RNA molecule encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof.

In some embodiments,

(i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43.

In some embodiments,

(i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45.

In some embodiments,

- (i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47.
- In some embodiments,
- (i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51.
- In some embodiments,
- (i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 44, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 44; and/or
- (ii) (a) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 1448 of SEQ ID NO: 43, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 1448 of SEQ ID NO: 43; or
- (b) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 43, 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: 43.
- In some embodiments,
- (i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 46, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 46; and/or

- (ii) (a) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 872 of SEQ ID NO: 45, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 872 of SEQ ID NO: 45; or
- (b) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 45, 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: 45.

In some embodiments,

- (i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 48, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 48; and/or
- (ii) (a) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 1943 of SEQ ID NO: 47, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 1943 of SEQ ID NO: 47; or
- (b) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 47, 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: 47.

In some embodiments,

- (i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 52, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 52; and/or
- (ii) (a) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 2591 of SEQ ID NO: 51, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 2591 of SEQ ID NO: 51; or
- (b) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 51, 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: 51.

In some embodiments,

(i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 63, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 63; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 62, 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: 62.

In some embodiments,

(i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 65, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 65; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 64, 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: 64.

In some embodiments,

(i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 67, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 67; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 66, 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: 66.

In some embodiments,

(i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 69, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 69; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 68, 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: 68.

In some embodiments,

(i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 71, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 71; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 70, 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: 70.

In some embodiments,

(i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 73, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 73; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 72, 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: 72.

In some embodiments,

(i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 75, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 75; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 74, 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: 74.

In some embodiments,

(i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 77, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 77; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 76, 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: 76.

In some embodiments, the RNA encodes an amino acid sequence comprising a secretory signal peptide.

In some embodiments, the secretory signal peptide is fused, preferably N-terminally, to the amino acid sequence.

In some embodiments, the secretory signal peptide is not endogenous to the Mtb antigen. In some embodiments, the secretory signal peptide is of human origin. In some embodiments, the secretory signal peptide is not of human origin.

In some embodiments,

- 5 (i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44, or a functional fragment of the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44; and/or

- 10 (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43.

15 In some embodiments,

- (i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, or a functional fragment of the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63; and/or

- 20 (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, or a fragment of the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62.

25 In some embodiments,

- (i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, or a functional fragment of the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71; and/or

- 30 (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, or a fragment of the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70.

In some embodiments, the composition or medical preparation comprises:

- (i) RNA comprising the nucleotide sequence of SEQ ID NO: 43, 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: 43;
- 40 (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 45, 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: 45;
- (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 47, 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: 47; and

(iv) RNA comprising the nucleotide sequence of SEQ ID NO: 51, 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: 51.

In some embodiments, the composition or medical preparation comprises:

- (i) RNA comprising the nucleotide sequence of SEQ ID NO: 43;
- 5 (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 45;
- (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 47; and
- (iv) RNA comprising the nucleotide sequence of SEQ ID NO: 51.

In some embodiments, the composition or medical preparation comprises:

- (i) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 44, or a nucleotide
- 10 sequence encoding 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;
- (ii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 46, or a nucleotide sequence encoding 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: 46;
- 15 (iii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 48, or a nucleotide sequence encoding 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: 48; and
- (iv) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 52, or a nucleotide sequence encoding an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity
- 20 to the amino acid sequence of SEQ ID NO: 52.

In some embodiments, the composition or medical preparation comprises:

- (i) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 44;
- (ii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 46;
- (iii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 48; and
- 25 (iv) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 52.

In some embodiments, the composition or medical preparation comprises:

- (i) RNA comprising the nucleotide sequence of SEQ ID NO: 62, 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: 62;
- (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 64, or a nucleotide sequence having at least 99%, 98%,
- 30 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 64;
- (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 66, 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: 66; and
- (iv) RNA comprising the nucleotide sequence of SEQ ID NO: 68, 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: 68.

35 In some embodiments, the composition or medical preparation comprises:

- (i) RNA comprising the nucleotide sequence of SEQ ID NO: 62;
- (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 64;
- (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 66; and
- (iv) RNA comprising the nucleotide sequence of SEQ ID NO: 68.

40 In some embodiments, the composition or medical preparation comprises:

- (i) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 63, or a nucleotide sequence encoding 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: 63;

- (ii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 65, or a nucleotide sequence encoding 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: 65;
- (iii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 67, or a nucleotide sequence encoding 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: 67; and
- (iv) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 69, or a nucleotide sequence encoding 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: 69.
- 10 In some embodiments, the composition or medical preparation comprises:
- (i) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 63;
- (ii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 65;
- (iii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 67; and
- (iv) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 69.
- 15 In some embodiments, the composition or medical preparation comprises:
- (i) RNA comprising the nucleotide sequence of SEQ ID NO: 70, 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: 70;
- (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 72, 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: 72;
- 20 (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 74, 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: 74; and
- (iv) RNA comprising the nucleotide sequence of SEQ ID NO: 76, 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: 76.
- In some embodiments, the composition or medical preparation comprises:
- 25 (i) RNA comprising the nucleotide sequence of SEQ ID NO: 70;
- (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 72;
- (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 74; and
- (iv) RNA comprising the nucleotide sequence of SEQ ID NO: 76.
- In some embodiments, the composition or medical preparation comprises:
- 30 (i) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 71, or a nucleotide sequence encoding 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: 71;
- (ii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 73, or a nucleotide sequence encoding 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: 73;
- 35 (iii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 75, or a nucleotide sequence encoding 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: 75; and
- (iv) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 77, or a nucleotide sequence encoding 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: 77.
- 40 In some embodiments, the composition or medical preparation comprises:
- (i) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 71;
- (ii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 73;

(iii) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 75; and

(iv) RNA comprising a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the RNA is formulated in lipid nanoparticles.

In some embodiments, the lipid nanoparticles comprise each of:

5 (i) a cationically ionizable lipid;

(ii) a steroid;

(iii) a neutral lipid; and

(iv) a polymer-conjugated lipid.

10 In some embodiments, the cationically ionizable lipid is present in a concentration ranging from about 40 to about 60 mol percent of the total lipids.

In some embodiments, the steroid is present in a concentration ranging from about 30 to about 50 mol percent of the total lipids.

In some embodiments, the neutral lipid is present in a concentration ranging from about 5 to about 15 mol percent of the total lipids.

15 In some embodiments, the polymer-conjugated lipid is present in a concentration ranging from about 1 to about 10 mol percent of the total lipids.

In some embodiments, the cationically ionizable lipid is within a range of about 40 to about 60 mole percent, the steroid is within a range of about 30 to about 50 mole percent, the neutral lipid is within a range of about 5 to about 15 mole percent, and the polymer-conjugated lipid is within a range of about 1 to about 10 mole percent.

20 In some embodiments, the cationically ionizable lipid is or comprises ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate).

In some embodiments, the steroid is or comprises cholesterol.

In some embodiments, the neutral lipid is or comprises a phospholipid.

In some embodiments, the phospholipid is or comprises distearoylphosphatidylcholine (DSPC).

25 In some embodiments, the polymer-conjugated lipid is or comprises a polyethylene glycol (PEG)-lipid.

In some embodiments, the PEG-lipid is or comprises 2-[(polyethylene glycol)-2000]-*N,N*-ditetradecylacetamide.

In some embodiments, the lipid nanoparticles comprise:

(a) ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate);

(b) cholesterol;

30 (c) distearoylphosphatidylcholine (DSPC); and

(d) 2-[(polyethylene glycol)-2000]-*N,N*-ditetradecylacetamide.

In some embodiments, ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate) is within a range of about 40 to about 60 mole percent, cholesterol is within a range of about 30 to about 50 mole percent, distearoylphosphatidylcholine (DSPC) is within a range of about 5 to about 15 mole percent, and 2-[(polyethylene

35 glycol)-2000]-*N,N*-ditetradecylacetamide is within a range of about 1 to about 10 mole percent.

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

In some embodiments, the 5' cap is or comprises a cap1 structure.

In some embodiments, the 5'-cap is or comprises $m_2^{7,3'}\text{-OGppp}(m_1^{2'-O})\text{ApG}$.

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

40 In some embodiments, the 5'-UTR is or comprises a modified human alpha-globin 5'-UTR.

In some embodiments, the 5'-UTR is or comprises the nucleotide sequence of SEQ ID NO: 56, 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: 56.

- In some embodiments, the 5'-UTR is or comprises the nucleotide sequence of SEQ ID NO: 57, 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: 57.
- In some embodiments, the RNA comprises a 3'-UTR.
- 5 In some embodiments, the 3'-UTR is or comprises a first sequence from the amino terminal enhancer of split (AES) messenger RNA and a second sequence from the mitochondrial encoded 12S ribosomal RNA.
- In some embodiments, the 3'-UTR is or comprises the nucleotide sequence of SEQ ID NO: 58, 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: 58.
- 10 In some embodiments, the RNA comprises a polyA sequence.
- In some embodiments, the polyA sequence is an interrupted sequence of A nucleotides.
- In some embodiments, the polyA sequence comprises 30 adenine nucleotides followed by 70 adenine nucleotides, wherein the 30 adenine nucleotides and 70 adenine nucleotides are separated by a linker sequence of 10 nucleotides.
- In some embodiments, the polyA sequence is or comprises the nucleotide sequence of SEQ ID NO: 59, or a nucleotide
- 15 sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 59.
- In some embodiments, the sequence downstream from the open reading frame, i.e., 3'-UTR and polyA sequence, is or comprises the nucleotide sequence of SEQ ID NO: 60, 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: 60.
- 20 In some embodiments, the RNA comprises a 5'-cap, a 5'-UTR, a 3'-UTR, and a polyA sequence.
- In some embodiments, the RNA comprises modified uridines.
- In some embodiments, the RNA comprises modified uridines in place of all uridines.
- In some embodiments, the modified uridines are N1-methyl-pseudouridine.
- In some embodiments, the coding sequence of the RNA is codon-optimized and/or is characterized in that its G/C
- 25 content is increased compared to the parental sequence.
- In some embodiments, the RNA is in a liquid formulation.
- In some embodiments, the RNA is in a frozen formulation.
- In some embodiments, the RNA is in a lyophilized formulation.
- In some embodiments, the RNA is formulated for injection.
- 30 In some embodiments, the RNA is formulated for intramuscular administration.
- In some embodiments, the composition or medical preparation 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 composition or medical preparation is a vaccine.
- 35 In some embodiments, the composition or medical preparation is a kit.
- In some embodiments, different RNA molecules are in separate vials.
- In some embodiments, the composition or medical preparation further comprises instructions for use of the composition or medical preparation for treating or preventing tuberculosis.
- In some embodiments, the composition or medical preparation is for pharmaceutical use.
- 40 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 tuberculosis.
- In some embodiments, the composition or medical preparation is for administration to a human.

In one aspect, the disclosure provides a method of vaccinating a subject comprising administering the composition described herein to the subject.

In some embodiments, the vaccination is for preventing tuberculosis.

In some embodiments, administration is by intramuscular administration.

- 5 In some embodiments, the method comprises administering to the subject at least one dose of the composition.
- In some embodiments, the method comprises administering to the subject at least two doses of the composition.
- In some embodiments, an amount of the RNA of at least 10 µg per dose is administered.
- In some embodiments, the subject is a human.

In further aspects, the disclosure provides the following:

- 10 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, or a fragment of the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, or the nucleotide sequence having at least 99%, 98%, 97%, 96%,
 15 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19.
- A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, or a fragment of the nucleotide
 20 sequence of positions 4 to 285 of SEQ ID NO: 21, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21.
- A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%,
 25 85%, or 80% identity to the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, or a fragment of the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22.
- A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%,
 30 or 80% identity to the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, or a fragment of the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23.
- A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%,
 35 or 80% identity to the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, or a fragment of the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24.
- 40 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, or a fragment of the nucleotide

sequence of positions 4 to 462 of SEQ ID NO: 25, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25.

- 5 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, or a fragment of the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28.

- 10 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof and comprising the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, or a fragment of the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30.

- 15 A polypeptide comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof.

- 20 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof.

- A polypeptide comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof.

- 25 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof.

- A polypeptide comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

- 30 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

- A polypeptide comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof.

- 35 A nucleic acid, e.g., RNA, encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof.

In some embodiments,

- 40 (i) a polypeptide comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 44, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 44; and/or

(ii) (a) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 1448 of SEQ ID NO: 43, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 1448 of SEQ ID NO: 43; or

(b) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 43, 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: 43.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 46, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 46; and/or

(ii) (a) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 872 of SEQ ID NO: 45, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 872 of SEQ ID NO: 45; or

(b) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 45, 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: 45.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 48, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 48; and/or

(ii) (a) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 1943 of SEQ ID NO: 47, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 1943 of SEQ ID NO: 47; or

(b) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 47, 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: 47.

In some embodiments,

- (i) a polypeptide comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 52, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 52; and/or
- 5 (ii) (a) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 2591 of SEQ ID NO: 51, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 2591 of SEQ ID NO: 51;
- 10 or
- (b) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 51, 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: 51.
- 15 In some embodiments,
- (i) a polypeptide comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 63, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 63; and/or
- 20 (ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 62, 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: 62.
- 25 In some embodiments,
- (i) a polypeptide comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 65, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 65; and/or
- 30 (ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 64, 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: 64.
- 35 In some embodiments,
- (i) a polypeptide comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 67, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 67; and/or
- 40

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 66, 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: 66.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 69, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 69; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 68, 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: 68.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 71, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 71; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 70, 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: 70.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 73, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 73; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 72, 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: 72.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 75, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 75; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 74, 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: 74.

In some embodiments,

(i) a polypeptide comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 77, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 77; and/or

(ii) a nucleic acid, e.g., RNA, encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 76, 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: 76.

Brief description of the Figures

Figure 1: RNA mixes of four modRNAs encoding 4 (RNA mix 1), 6 (RNA mix 2), or 8 (RNA mix 3) *Mtb* antigens. *Top* General mRNA construct structure with 5'-cap, 5'- and 3'-untranslated regions (UTR), the open reading frame (ORF), and a poly-adenosine tail. *Bottom*, RNA mixes 1 to 3 comprising 4 to 8 *Mtb* antigens sequences tested. All antigen-encoding sequences were fused on the N-terminus to a major histocompatibility complex (MHC) class I signal peptide fragment (sec) that mediates translocation into the endoplasmic reticulum. **A)** The modRNA constructs in RNA mix 1 encode *Mycobacterium tuberculosis* (*Mtb*) antigens Ag85A(Δ1-41), M72, ESAT6, and HbhA separately. **B)** The modRNA constructs in RNA mix 2 encode antigens Ag85A(Δ1-41), and M72 separately and 2 fusion-antigens: Hrp1-ESAT6 and RpfD-HbhA. **C)** The modRNA constructs in RNA mix 3 encode four fusion-antigens: ESAT6-RpfD, Ag85A(Δ1-41)-Hrp1, RpfA-HbhA, and M72-vapB47. The modRNAs mixtures were formulated with lipid nanoparticles 315 (LNP-315). UTR: untranslated region, poly(A): poly-adenosine tail.

Figure 2: Immunization schedule for *in vivo* mouse immunogenicity studies. Mice received **A)** two intramuscular (i.m.) injections (on days 0 and 21) or **B)** three intravenous (i.v.) injections (on days 0, 7, and 21). Arrowheads indicate injection days. Blood samples (indicated by the arrows) were collected for serum analysis of antigen-specific IgG antibodies on day 14, 28, and 42 after the first injection. On the last experimental day (day 42) mice spleens were dissected for isolation of splenocytes, for subsequent analysis of T-cell responses to antigen-specific peptides.

Figure 3: Scheme of the mRNA constructs used to determine the most immunogenic mRNA platform and *Mtb* antigens. *Top* General mRNA construct structure with 5'-cap, 5'- and 3'-untranslated regions (UTR), the open reading frame (ORF), and a poly-adenosine tail. *Bottom, A-E)* The antigen-encoding mRNA constructs included in the ORF. All antigen-encoding sequences were fused to an N-terminal major histocompatibility complex (MHC) class I signal peptide fragment (sec) that mediates translocation into the endoplasmic reticulum. One construct (B) contained a C-terminal MHC class I transmembrane and cytoplasmic domain (MITD), a cell trafficking-signal for cell membrane anchoring. *Mycobacterium tuberculosis* antigens tested: Ag85A, ESAT6, HbhA, Hrp1, M72, RpfA, RpfD, vapB47.

Figure 4: PPD stimulation of splenocytes showed better induction of immune response in mice immunized with pseudouridine-modified mRNA in comparison to unmodified mRNA and self-amplifying mRNA.

C57BL/6 mice (5 animals per group) were immunized with the indicated mRNA constructs/mix of mRNA constructs (6-antigen cassette, 6-antigen cassette with MITD, 6-antigen mix, 2-antigen mix, 6-antigen cassette + 2 antigens) as unmodified mRNA (uRNA), pseudouridine-modified mRNA (modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg (100 µL dose volume) uRNA or intramuscularly twice (days 0 and 21) with 4 µg (in 20 µL dose volume) modRNA or saRNA. Mice in a reference group were subcutaneously (s.c.) injected with 10⁶ colony forming units (CFU; 100 µL dose volume) of bacillus Calmette-Guerin (BCG) once, on day 0. Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5 × 10⁵ cells) in culture were treated overnight (12-16 h) with 10 µg/mL purified protein-derivative (PPD) and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate, bars represent group mean spot-forming units (SFU) $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, α = 0.05. * = P <0.05, ** = P <0.01, *** = P <0.001, **** = P <0.0001, MITD: MHC class I transmembrane and cytoplasmic domain, modRNA: nucleoside-modified mRNA, saRNA: self-amplifying mRNA, uRNA: unmodified mRNA, SD: standard deviation, SFU: spot-forming units.

Figure 5: Immunogenicity induced by the 6-antigen cassette using three mRNA platforms. C57BL/6 mice (5 animals per group) were immunized with the 6-antigen cassette mRNA construct (encoding *Mycobacterium tuberculosis* antigens Ag85A, ESAT6, vapB47, Hrp1, RpfA, and RpfD) as unmodified mRNA (uRNA), or pseudouridine-modified mRNA (modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg uRNA (100 µL dose volume) or intramuscularly twice (days 0 and 21) with 4 µg modRNA or saRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5 × 10⁵ cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1 and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate, bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4⁺ and CD8⁺ T cells (10⁵ cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on days 14, 28, and 42 after the first immunization to obtain the serum. Antigen-specific IgG was determined in serum samples (1:100 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, α = 0.05. * = P <0.05, ** = P <0.01, *** = P <0.001, **** = P <0.0001, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, MITD: MHC class I transmembrane and cytoplasmic domain, modRNA: nucleoside-modified mRNA, saRNA: self-amplifying mRNA, uRNA: unmodified mRNA, SD: standard deviation, SFU: spot-forming units.

40

Figure 6: Immunogenicity induced by the 6-antigen cassette with MITD using three mRNA platforms.

C57BL/6 mice (5 animals per group) were immunized with the 6-antigen cassette mRNA construct (encoding *Mycobacterium tuberculosis* antigens Ag85A, ESAT6, vapB47, Hrp1, RpfA, RpfD) containing a C-terminal MHC class I transmembrane and cytoplasmic domain (MITD) as unmodified mRNA (uRNA), or pseudouridine-modified mRNA

(modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg uRNA (100 µL dose volume) or intramuscularly twice (days 0 and 21) with 4 µg modRNA or saRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1 and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate, bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4 $^{+}$ and CD8 $^{+}$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on days 14, 28, and 42 after the first immunization to obtain the serum. Antigen-specific IgG was determined in serum samples (1:100 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, MITD: MHC class I transmembrane and cytoplasmic domain, modRNA: nucleoside-modified mRNA, saRNA: self-amplifying mRNA, uRNA: unmodified mRNA, SD: standard deviation, SFU: spot-forming units.

Figure 7: Immunogenicity induced by the 6-antigen mix using three mRNA platforms. C57BL/6 mice (5 animals per group) were immunized with the 6-antigen mix mRNA constructs (separately encoded *Mycobacterium tuberculosis* antigens Ag85A, ESAT6, vapB47, Hrp1, RpfA, RpfD) as unmodified mRNA (uRNA), or pseudouridine-modified mRNA (modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg uRNA (100 µL dose volume) or intramuscularly twice (days 0 and 21) with 4 µg modRNA or saRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1 and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate, bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4 $^{+}$ and CD8 $^{+}$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on days 14, 28, and 42 after the first immunization to obtain the serum. Antigen-specific IgG was determined in serum samples (1:300 dilution for Ag85A and Hrp1, and 1:100 dilution for the other antigens) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, MITD: MHC class I transmembrane and cytoplasmic domain, modRNA: nucleoside-modified mRNA, saRNA: self-amplifying mRNA, uRNA: unmodified mRNA, SD: standard deviation, SFU: spot-forming units.

Figure 8: Immunogenicity induced by the 2-antigen mix using three mRNA platforms. C57BL/6 mice (5 animals per group) were immunized with the 2-antigen mix mRNA constructs (separately encoded *Mycobacterium tuberculosis* antigens HbhA and M72) as unmodified mRNA (uRNA), or pseudouridine-modified mRNA (modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with

C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg uRNA (100 µL dose volume) or intramuscularly twice (days 0 and 21) with 4 µg modRNA or saRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5 × 10⁵ cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1 and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate, bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4⁺ and CD8⁺ T cells (10⁵ cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on days 14, 28, and 42 after the first immunization to obtain the serum. Antigen-specific IgG was determined in serum samples (1:100 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, MITD: MHC class I transmembrane and cytoplasmic domain, modRNA: nucleoside-modified mRNA, saRNA: self-amplifying mRNA, uRNA: unmodified mRNA, SD: standard deviation, SFU: spot-forming units.

Figure 9: Immunogenicity induced by the M72 antigen using modRNA platform. C57BL/6 mice (5 animals per group) were immunized with modRNA encoding M72 formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 µg modRNA or saRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5 × 10⁵ cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1 and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate, bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4⁺ and CD8⁺ T cells (10⁵ cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on days 14, 28, and 42 after the first immunization to obtain the serum. Antigen-specific IgG was determined in serum samples (1:25 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, MITD: MHC class I transmembrane and cytoplasmic domain, modRNA: nucleoside-modified mRNA, saRNA: self-amplifying mRNA, uRNA: unmodified mRNA, SD: standard deviation, SFU: spot-forming units.

Figure 10: Immunogenicity induced by the 6-antigen cassette + 2-antigens using three mRNA platforms. C57BL/6 mice (5 animals per group) were immunized with the 6-antigen cassette + 2-antigens mixture of mRNA constructs (separately encoded *Mycobacterium tuberculosis* antigens M72 and HbhA and the construct encoding a fusion protein of antigens Ag85A, ESAT6, vapB47, Hrp1, RpfA, and RpfD) as unmodified mRNA (uRNA), or pseudouridine-modified mRNA (modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg uRNA (100 µL dose volume) or intramuscularly twice (days 0 and 21) with 4 µg modRNA or saRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were

dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 $\mu\text{g/mL}$ overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1 and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate, bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4 $^+$ and CD8 $^+$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on days 14, 28, and 42 after the first immunization to obtain the serum. Antigen-specific IgG was determined in serum samples (1:100 dilution) by ELISA. Results are shown as ΔOD . Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. ΔOD : 620 nm absorbance subtracted from 450 nm absorbance, MITD: MHC class I transmembrane and cytoplasmic domain, modRNA: nucleoside-modified mRNA, saRNA: self-amplifying mRNA, uRNA: unmodified mRNA, SD: standard deviation, SFU: spot-forming units.

Figure 11: Scheme the mRNA constructs used to test fusion of signal peptides alone or in combination with transmembrane domains, using 2 different codon optimizations. *Top* General mRNA construct structure with 5'-cap, 5'- and 3'-untranslated regions (UTR), the open reading frame (ORF), and a poly-adenosine (A) tail. *Bottom*, the antigen-encoding mRNA constructs included in the ORF. All antigen-encoding sequences were codon optimized (opt1 or opt10). Antigen-encoding sequences were included in the mRNA backbone alone or fused to a secretion signal peptide (sec, SP1 or SP2) to its N-terminus with or without a transmembrane domain (TMD1, TMD2, or TMD3) fused to its C-terminus. *Mycobacterium tuberculosis* antigens tested: Ag85A(Δ 1-41), RpfA, and vapB47.

Figure 12: Immunogenicity induced by Ag85A with or without a signal peptide. C57BL/6 mice (5 animals per group) were immunized with codon optimized (opt1 or opt10) pseudouridine-modified mRNA (modRNA) construct encoding *Mycobacterium tuberculosis* antigen Ag85A(Δ 1-41) alone or with a secretion signal peptide sequence (sec, SP1, or SP2) fused to its N-terminus, as described in Figure 11. The modRNAs were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 μg modRNA (20 μL dose volume). Mice in the control group were injected intramuscularly with 20 μL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 $\mu\text{g/mL}$ overlapping peptide pools covering the Ag85A(Δ 1-41) antigen or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4 $^+$ and CD8 $^+$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on day 42 after the first immunization to obtain the serum. Ag85A-specific IgG was determined in serum samples (1:2700 dilution) by ELISA. Results are shown as ΔOD . Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Dunnett's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, ΔOD : 620 nm absorbance subtracted from 450 nm absorbance, SD: standard deviation, SFU: spot-forming units, SP: secretion signal peptide.

Figure 13: Immunogenicity induced by signal peptide-fused Ag85A with or without a transmembrane domain sequence. C57BL/6 mice (5 animals per group) were immunized with codon optimized (opt1 or opt10) pseudouridine-modified mRNA (modRNA) construct encoding *Mycobacterium tuberculosis* antigen Ag85A(Δ 1-41) with an N-terminally fused secretion signal peptide sequence (sec, SP1, or SP2) and a transmembrane domain sequence

fused to its C-terminus, as described in Figure 11. The modRNAs were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 µg modRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the Ag85A(Δ 1-41) antigen or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4 $^{+}$ and CD8 $^{+}$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on day 42 after the first immunization to obtain the serum. Ag85A-specific IgG was determined in serum samples (1:300 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Dunnett's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, SD: standard deviation, SFU: spot-forming units, SP: secretion signal peptide, TMD: transmembrane domain.

Figure 14: Immunogenicity induced by RpfA with or without a signal peptide. C57BL/6 mice (5 animals per group) were immunized with codon optimized (opt1 or opt10) pseudouridine-modified mRNA (modRNA) construct encoding *Mycobacterium tuberculosis* antigen RpfA alone or with a secretion signal peptide sequence (sec, SP1, or SP2) fused to its N-terminus, as described in Figure 11. The modRNAs were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 µg modRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the RpfA antigen or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4 $^{+}$ and CD8 $^{+}$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C)** Blood samples were collected from mice on day 42 after the first immunization to obtain the serum. RpfA-specific IgG was determined in serum samples (1:300 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Dunnett's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, SD: standard deviation, SFU: spot-forming units, SP: secretion signal peptide.

Figure 15: Immunogenicity induced by signal peptide-fused RpfA with or without a transmembrane domain sequence. C57BL/6 mice (5 animals per group) were immunized with codon optimized (opt1 or opt10) pseudouridine-modified mRNA (modRNA) construct encoding *Mycobacterium tuberculosis* antigen RpfA with an N-terminally fused secretion signal peptide sequence (sec, SP1, or SP2) and a transmembrane domain sequence fused to its C-terminus, as described in Figure 11. The modRNAs were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 µg modRNA (20 µL dose volume). Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the RpfA antigen or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were

measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **B**) CD4⁺ and CD8⁺ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C**) Blood samples were collected from mice on day 42 after the first immunization to obtain the serum. RpfA-specific IgG was determined in serum samples (1:300 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Dunnett's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, SD: standard deviation, SFU: spot-forming units, SP: secretion signal peptide, TMD: transmembrane domain.

Figure 16: Immunogenicity induced by VapB47 with or without a signal peptide. C57BL/6 mice (5 animals per group) were immunized with codon optimized (opt1 or opt10) pseudouridine-modified mRNA (modRNA) construct encoding *Mycobacterium tuberculosis* antigen VapB47 alone or with a secretion signal peptide sequence (sec, SP1, or SP2) fused to its N-terminus, as described in Figure 11. The modRNAs were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 μ g modRNA (20 μ L dose volume). Mice in the control group were injected intramuscularly with 20 μ L phosphate buffer saline (buffer). **A**) On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 μ g/mL overlapping peptide pools covering the VapB47 antigen or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **B**) CD4⁺ and CD8⁺ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C**) Blood samples were collected from mice on day 42 after the first immunization to obtain the serum. VapB47-specific IgG was determined in serum samples (1:300 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Dunnett's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, Δ OD: 620 nm absorbance subtracted from 450 nm absorbance, SD: standard deviation, SFU: spot-forming units, SP: secretion signal peptide.

Figure 17: Immunogenicity induced by signal peptide-fused VapB47 with or without a transmembrane domain sequence. C57BL/6 mice (5 animals per group) were immunized with codon optimized (opt1 or opt10) pseudouridine-modified mRNA (modRNA) construct encoding *Mycobacterium tuberculosis* antigen VapB47 with an N-terminally fused secretion signal peptide sequence (sec, SP1, or SP2) and a transmembrane domain sequence fused to its C-terminus, as described in Figure 11. The modRNAs were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 μ g modRNA (20 μ L dose volume). Mice in the control group were injected intramuscularly with 20 μ L phosphate buffer saline (buffer). **A**) On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 μ g/mL overlapping peptide pools covering the VapB47 antigen or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **B**) CD4⁺ and CD8⁺ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. **C**) Blood samples were collected from mice on day 42 after the first immunization to obtain the serum. VapB47-specific IgG was determined in serum samples (1:300 dilution) by ELISA. Results are shown as Δ OD. Samples were measured in

duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Dunnett's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, ΔOD : 620 nm absorbance subtracted from 450 nm absorbance, SD: standard deviation, SFU: spot-forming units, SP: secretion signal peptide, TMD: transmembrane domain.

5

Figure 18: PPD-induced cellular immune response in mice immunized with three different modRNA mixes. C57BL/6 mice (5 animals per group) were immunized with mixtures of codon optimized (opt10) pseudouridine-modified mRNA (modRNA) constructs as described in Figure 1. The modRNAs mixtures were formulated with Acuitas ALC-315 lipid nanoparticles (LNP-315). Mice were injected intramuscularly twice (days 0 and 21) with 4 μ g modRNA (20 μ L dose volume). Mice in a reference group were subcutaneously (s.c.) injected with 10^6 colony forming units (CFU; 100 μ L dose volume) of bacillus Calmette-Guerin (BCG) once, on day 0. Mice in the control group were injected intramuscularly with 20 μ L phosphate buffer saline (buffer). **A)** On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (1.25×10^5 cells) in culture were treated overnight (12-16 h) with 10 μ g/mL purified protein-derivative (PPD), and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Samples were measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **B)** CD4 $^+$ and CD8 $^+$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, SD: standard deviation.

10

15

20

Figure 19: Antigen-specific cellular immune responses in mice immunized with three different modRNA mixes. C57BL/6 mice (5 animals per group) were immunized with mixtures of codon optimized (opt1) pseudouridine-modified mRNA (modRNA) constructs as described in Figure 1. The modRNAs mixtures were formulated with Acuitas ALC-315 lipid nanoparticles (LNP-315). Mice were injected intramuscularly twice (days 0 and 21) with 4 μ g modRNA (20 μ L dose volume). Mice in the control group were injected intramuscularly with 20 μ L phosphate buffer saline (buffer). On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. **A)** Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 μ g/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. **B)** Because spot counts from treatment of splenocytes with Ag85A- and M72-specific peptide pools were too high when 5×10^5 splenocytes were used, response to these antigens was analyzed using 1.25×10^5 splenocytes, treated as in A. Samples were measured in duplicate; bars represent group mean spot-forming units (SFU) $\pm$ SD. **C)** CD4 $^+$ and **D)** CD8 $^+$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as in A. Pooled T cell samples were measured in triplicates; bars represent group replicates mean spot-forming units (SFU) $\pm$ SD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, SD: standard deviation.

25

30

35

Figure 20: Antigen-specific humoral immune responses in mice immunized with three different modRNA mixes. C57BL/6 mice (5 animals per group) were immunized with mixtures of codon optimized (opt1) pseudouridine-modified mRNA (modRNA) constructs as described in Figure 1. The modRNAs mixtures were formulated with Acuitas ALC-315 lipid nanoparticles (LNP-315). Mice were injected intramuscularly twice (days 0 and 21) with 4 μ g modRNA (20 μ L dose volume). Mice in the control group were injected intramuscularly with 20 μ L phosphate buffer saline (buffer). Blood samples were collected from mice on day 42 after the first immunization to obtain the serum. Antigen-

40

specific IgG was determined in serum samples (serum dilution indicated on each graph title) by ELISA. Results are shown as Δ OD. Samples were measured in duplicates; bars represent group mean $\pm$ SD. Statistical analysis were performed with one-way ANOVA, with Tukey's multiple-comparison test, $\alpha = 0.05$. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$, SD: standard deviation.

5

Figure 21: *In vitro* expression of uRNA Mix 3 and modRNA Mix 3

HEK293T/17 cells were transfected with unmodified mRNA or nucleoside-modified mRNA as single mRNA (single, 0.25 μ g mRNA), comprised within a mixture of equal amounts of each of the four fusion-mRNAs (mix 1; 0.25 μ g of each mRNA, total of 1 μ g) or with drug substance containing the four fusion-mRNAs (mix 2; 1 μ g total mRNA). Non-transfected HEK293T/17 cells served as control (NT). Whole cell lysates of non-treated and transfected cells were generated and separated by SDS-PAGE. Proteins were blotted on a nitrocellulose-membrane and expression of fusion-proteins was assessed using antigen-specific antibodies. Expected molecular weights: A: Ag85A-Hrp1, 50 kDa; B: ESAT-6-RpfD, 28 kDa; C: RpfA-HbhA, 120 kDa; D: M72-VapB47, 86 kDa.

Figure 22: T-cell responses induced by immunization with uRNA Mix 3 and modRNA Mix 3

Splenocytes were isolated on study Day 42 from C57BL/6 mice injected with uRNA Mix 3, modRNA Mix 3, or a control (saline). Cells were stimulated for 18 hours with antigen-specific overlapping peptide pools at 2 μ g/mL per peptide and responses were assessed by an IFN- γ ELISpot assay. Bars represent treatment group average. One-way analysis of variance (ANOVA), $\alpha = 0.05$. *** = $p < 0.001$, **** = $p < 0.0001$. Significance values between treatment group and the control group are not depicted.

Figure 23: Antibody responses induced by immunization with uRNA Mix 3 and modRNA Mix 3

Antigen-specific immunoglobulin G (IgG) antibodies in sera from C57BL/6 mice immunized with uRNA Mix 3, modRNA Mix 3, or a saline control were assessed by ELISA. The represented data is from Day 42 after the first immunization. No IgG was detected in the sera from mice in the control group. Bars represent treatment group average; symbols represent values for each mouse sample. One-way analysis of variance (ANOVA), $\alpha = 0.05$. ** = $p < 0.01$. Significance values between treatment group and the control group are not depicted.

Figure 24: CD4⁺ and CD8⁺ T-cell specific responses assessed by intracellular cytokine staining of cells from uRNA- and modRNA Mix 3-injected mice

Splenocytes were isolated on study Day 42 from C57BL/6 mice injected with uRNA Mix 3, modRNA Mix 3, or a control (saline). (A) and (B) Cells were stimulated for 5 to 6 hours with a mix of overlapping peptide pools covering all the antigens at 1 μ g/mL per peptide in the presence of co-stimulatory antibodies (CD28 and CD49d), Golgi Stop and Golgi Plug (Protein transport inhibitors). Cells were stained for intracellular and extracellular markers including viability, CD3, CD4, CD8, IFN- γ , IL-2, and TNF α . Cells were analyzed using a BD-Celesta flow cytometer to identify specific cell types. (A) CD4⁺ and (B) CD8⁺ cells. (C) Splenocytes were stained with a different panel of surface markers in the presence of Mtb32A tetramer (PepA, a component of M72) and cells were acquired in BD-Celesta and analyzed. Data represents cells positive for single cytokines and polyfunctional T cells (IFN- γ , IL-2, and TNF α secreting cells). Bars represent treatment group average; symbols represent values for each mouse sample. One-way analysis of variance (ANOVA), $\alpha = 0.05$. Significance values between treatment group and the control group are not depicted.

Figure 25: T-cell responses induced by immunization with uRNA Mix 3 and modRNA Mix 3

Splenocytes were isolated on study Day 42 from BALB/c mice injected with uRNA Mix 3, modRNA Mix 3, or a control (saline). Total splenocytes were stimulated for 18 hours with antigen-specific overlapping peptide pools at 2 μ g/mL per

peptide and responses were assessed by an IFN- γ ELISpot assay. Bars represent treatment group average ($\pm$ standard deviation). One-way analysis of variance (ANOVA), $\alpha = 0.05$. Significance values between treatment group and the control group are not depicted.

5 **Figure 26: Antibody responses induced by immunization with uRNA Mix 3 and modRNA Mix 3**

Antigen-specific immunoglobulin G (IgG) antibodies in sera from BALB/c mice immunized with uRNA Mix 3, modRNA Mix 3, or a saline control were assessed by ELISA. The represented data is from Day 42 after the first immunization. No IgG was detected in the sera from mice in the control group. Bars represent treatment group average; symbols represent values for each mouse sample. One-way analysis of variance (ANOVA), $\alpha = 0.05$. * = $p < 0.05$. Significance values between treatment group and the control group are not depicted.

10 **Figure 27: Cellular immune responses induced by immunization with uRNA Mix 3 or modRNA Mix 3 in a humanized mouse model**

Splenocytes were isolated on Day 42 from humanized mice (transgenic for HLA alleles A2.1/DR1) injected with 4 μ g uRNA Mix 3, 4 μ g modRNA Mix 3, or a saline control (Buffer). Splenic CD4⁺ and CD8⁺ T cells were magnetically isolated and stimulated with antigen-specific peptide pools in presence of autologous bone marrow-derived dendritic cells. Cellular responses were assessed by an IFN- γ ELISpot assay after $\sim$ 18 hours incubation. Mean spot counts per group using T cells (measured in triplicate wells) are indicated by bars ($\pm$ standard deviation). Counts above 1,500 SFU are too numerous to count (TNTC).

20 **Figure 28: Cellular responses induced by immunization with modRNA Mix 3 or with the individual RNAs comprising modRNA Mix 3**

Splenocytes were isolated on study Day 42 post prime from C57BL/6 mice injected with 4 μ g modRNA Mix 3 or with 1 μ g RNA-LNP (Ag85A-Hrp1, ESAT6-RpfD, RpfA-HbhA, or M72-VapB47). Control group received saline (Buffer). Cells were stimulated with antigen-specific peptide pools and responses were assessed by an IFN- γ ELISpot assay after $\sim$ 18 h incubation. Group mean spot counts are indicated by bars ($\pm$ standard deviation). One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed, P-values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Not indicated in figure: IFN- γ secretion from both the test groups were significant compared to saline group except for HbhA and VapB47. SFUs with 5×10^5 cells per well reached above detectable range of the assay for Ag85A and M72, hence a lower cell number was used for stimulation.

30 **Figure 29: Humoral responses induced by immunization with modRNA Mix 3 or with the individual RNAs comprising modRNA Mix 3**

Antigen-specific IgG antibodies in sera from C57BL/6 mice immunized with 4 μ g modRNA Mix 3 or 1 μ g of RNA-LNP (Ag85A-Hrp1, ESAT6-RpfD, RpfA-HbhA, or M72-VapB47) were assessed by ELISA. Control group received saline (Buffer). The data shown are from Day 42 post prime. Group mean values are indicated by horizontal bars ($\pm$ standard deviation), mean from an individual mouse (measured in duplicates) are depicted as circles. One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed, P-values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. Not indicated in figure: significance values of the test groups compared to saline group. Numbers in brackets above the graph indicate the serum dilution.

40 **Figure 30: Cellular and humoral responses against Ag85A and Hrp1 induced by immunization with modRNA Mix 3 or with the individual RNAs contained in modRNA Mix 3 or with single RNA encoding single antigens**

Splenocytes and blood were isolated on Day 42 post prime from C57BL/6 mice injected with 4 µg modRNA Mix 3 or 1 µg RNA-LNPs encoding Ag85A–Hrp1 or Ag85A or Hrp1. Control group received saline (Buffer). (A) Splenocytes were stimulated with antigen-specific peptide pools and responses were assessed by an IFN-γ ELISpot assay. Group mean values are indicated by bars (± standard deviation). (B) Antigen-specific immunoglobulin G (IgG) antibodies in sera were assessed by ELISA. Group mean values are indicated by horizontal bars (± standard deviation), mean from individual mouse (measured in duplicates) are depicted as circles. One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed, P-values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. Not indicated in figure: significance values of the test groups compared to saline group. Numbers in brackets of B above the graph indicate the serum dilution.

Figure 31: Humoral responses induced by immunization with uRNA Mix 3 and modRNA Mix 3 in Wistar Han rats

(A) Immunization schedule. (B) Antigen-specific IgG antibodies in sera from Wistar Han rats immunized with 30 µg uRNA Mix 3, 30 µg modRNA Mix 3, or a saline control (Buffer) were assessed by ELISA 28 days post first immunization. Group mean values are indicated by horizontal bars (± standard deviation), mean from individual mouse (measured in duplicates) are depicted as circles. One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed, P-values: * = $p < 0.05$, ** = $p < 0.01$. Not indicated in figure: significance values of the test groups compared to saline group. Numbers in brackets above the graph indicate the serum dilution.

Figure 32: Cellular responses induced by injection of uRNA Mix 3 and modRNA Mix 3 in Wistar Han rats

Wistar Han Rats were administered intramuscular injections on Days 0, 7, 14, and 21 with 30 µg uRNA Mix 3 or modRNA Mix 3. Control group received saline. IFN-γ ELISpot assay was performed using splenocytes isolated on Day 28 and stimulated for ~36 h with each of eight Mtb antigen-specific overlapping peptide pools or the respective recombinant Mtb proteins. Group mean values are indicated by bars (± standard deviation). One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed, P-values: * = $p < 0.05$, ** = $p < 0.01$.

Figure 33: Study design and read outs of the *in vivo* Mtb challenge study in C57BL/6 mice

The image shows the immunization schedule for different groups and two termination time points. Arrows below bleeding and termination time points indicate the respective assays performed.

Figure 34: Humoral responses induced by immunization with uRNA Mix 3 and modRNA Mix 3 in C57BL/6 mice after boost and challenge with M. tuberculosis H37Rv

Antigen-specific IgG antibodies in sera from C57BL/6 mice were assessed by ELISA for the indicated groups at various time points. The data shown are from Days (d) 43, 87, and 117. "Prime" is one dose and "prime and boost" is two doses. Group mean values are indicated by horizontal bars (± standard deviations), means from individual mice (measured in duplicates) are depicted as circles. One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed, P-values: ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. Not indicated in figure: significance values of the test groups compared to saline group or NTL group. Numbers in brackets above the graph indicate the serum dilution.

Figure 35: Enumeration of Mtb H37Rv from lung and spleen of infected C57BL/6 mice following immunization

Saline-injected or vaccinated C57BL/6 mice were aerosol infected with ~100 colony forming units (CFU) of Mycobacterium tuberculosis strain H37Rv. "Prime" is one dose and "prime and boost" is two doses. The bacterial burden

in lung right lobe and whole spleen was determined on Day 87, 30 days post-infection (A; n = 10) or Day 117, 60 days post-infection (B; n = 5). Group mean values are indicated by horizontal bars ($\pm$ standard deviations), means from individual mice (measured in duplicates) are depicted as circles. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test between the saline-treated group and all other groups was performed, P-values: * = $p < 0.05$; ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$.

Figure 36: Effect of a third immunization with uRNA Mix 3 and modRNA Mix 3 on humoral responses.

(A) Immunization schedule and timepoints of serum collection. (B) Antigen-specific IgG antibodies in sera from C57BL/6 mice immunized with uRNA Mix 3 or modRNA Mix 3 were assessed by ELISA at the indicated time points. The data shown are from days (d) 42 to 155. Group mean values are indicated by symbols ($\pm$ standard deviations). Numbers in brackets above the graph indicate the serum dilution. The dotted line indicates the third immunization at d134.

Figure 37: N-terminal fusion of signal peptide alternatives to uRNA Mix 3 and modRNA Mix 3

Top General mRNA construct structure with 5'-cap, 5'- and 3'-untranslated regions (UTR), the open reading frame (ORF), and a poly-adenosine tail. *Bottom*, RNA Mix 3 comprising *Mtb* antigen sequences. All antigen-encoding sequences were fused on the N-terminus to a major histocompatibility complex (MHC) class I signal peptide fragment (sec) that mediates translocation into the endoplasmic reticulum or an alternative signal peptide fragment. The RNA mixtures were formulated with lipid nanoparticles 315 (LNP-315). UTR: untranslated region, poly(A): poly-adenosine tail.

Figure 38: In vitro expression of uRNA Mix 3 fused to alternative signal peptides

HEK293T cells were transfected with 1 μ g of uRNA Mix 3 comprising the four fusion-antigens with either SP1, SP2, or sec. At 18h post transfection cells were stained for viability and *Mtb* antigens comprised by uRNA Mix 3 (Ag85A, Hrp1, ESAT-6, RpfD, RpfA, HbhA, M72 and vapB47) with specific antibodies. Data shows mean fluorescence intensities of the antigen-specific staining within the viable cell population. Bars represent mean values, symbols represent technical replicates of transfection and staining. Circles (first bars) show non-transfected (negative control), up-pointing triangles (second bars) show sec, squares (third bars) show SP1, and diamonds (fourth bar) show SP2. One-way analysis of variance (ANOVA) with Dunnett's multiple comparison test, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Not indicated in figure: significance values of the test items compared to non-transfected controls.

Figure 39: In vitro expression of modRNA Mix 3 fused to alternative signal peptides HEK293T cells were transfected with 1 μ g of modRNA Mix 3 comprising the four fusion-antigens with either SP1, SP2, or sec. At 18h post transfection cells were stained for viability and *Mtb* antigens comprised by modRNA Mix 3 (Ag85A, Hrp1, ESAT-6, RpfD, RpfA, HbhA, M72 and vapB47) with specific antibodies. Data shows mean fluorescence intensities of the antigen-specific staining within the viable cell population. Bars represent mean values, symbols represent technical replicates of transfection and staining. Circles (first bars) show non-transfected (negative control), up-pointing triangles (second bars) show sec, squares (third bars) show SP1, and diamonds (fourth bar) show SP2. One-way analysis of variance (ANOVA) with Dunnett's multiple comparison test, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. Not indicated in figure: significance values of the test items compared to non-transfected controls.

Figure 40: Humoral responses induced by immunization with uRNA Mix 3 or modRNA Mix 3 fused to alternative signal peptides

Antigen-specific IgG antibodies in sera from C57BL/6 mice were assessed by ELISA for the indicated groups at d42 post first immunization. Group mean values are indicated by horizontal bars ($\pm$ standard deviations), means from

individual mice (measured in duplicates) are depicted as symbols. Circles show buffer (negative control), up-pointing triangles show sec, squares show SP1, and diamonds show SP2. Dark symbols represent uRNA Mix 3 and white symbols represent modRNA Mix 3. One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed, P-values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Not indicated in figure: significance values of the test groups compared to buffer group. Numbers in brackets above the graph indicate the serum dilution.

Figure 41: Cellular responses from total splenocytes induced by immunization with uRNA Mix 3 or modRNA Mix 3 fused to alternative signal peptides

Splenocytes were isolated on study day 42 post first immunization from C57BL/6 mice injected with 4 μ g uRNA Mix 3 or modRNA Mix 3 (comprising fusion antigens Ag85A-Hrp1, ESAT6-Rpfd, RpfA-HbhA, or M72-VapB47 with SP1, SP2 or sec). Control group received saline (Buffer). Cells were stimulated with antigen-specific peptide pools and responses were assessed by an IFN- γ ELISpot assay after ~18 h incubation. Group mean spot counts ($\pm$ standard deviation) are indicated by bars. Symbols depict mean responses from individual mice measured in duplicates. Circles show buffer (negative control), up-pointing triangles show sec, squares show SP1 and diamonds show SP2. Dark symbols represent uRNA Mix 3 and white symbols represent modRNA Mix 3. Y-axis shows IFN- γ -secreting cells / 5×10^5 splenocytes for Trp1, Hrp1, ESAT-6, RpfA, HbhA and VapB47, and IFN- γ -secreting cells / 1.25×10^5 splenocytes for Ag85A, Rpfd and M72. One-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed between test groups immunized with the uRNA encoded mixes or the modRNA encoded mixes, respectively. P-values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. Not indicated in figure: Significant IFN- γ secretion of test groups versus buffer. Spot counts with 5×10^5 cells per well reached above detectable range of the assay for Ag85A, Rpfd and M72, hence a lower cell number was used for stimulation.

Figure 42: Cellular responses from CD4⁺ T cells induced by immunization with uRNA Mix 3 or modRNA Mix 3 fused to alternative signal peptides

CD4⁺ T cells were isolated from mouse splenocytes pooled by treatment group and stimulated with antigen-specific peptide pools in presence of autologous bone-marrow derived dendritic cells. Pooled T cell samples were measured in duplicates or triplicates where possible; bars represent group replicates mean spot-forming units (SFU) $\pm$ standard deviation. Bars represent group mean values. Bars from left to right are buffer, sec uRNA Mix 3, SP1 uRNA Mix 3, SP2 uRNA Mix 3, sec modRNA Mix 3, SP1 modRNA Mix 3 and SP2 modRNA Mix 3. Y-axis shows IFN- γ -secreting cells / 1×10^5 CD4⁺ cells for all panels.

Figure 43: Cellular responses from CD8⁺ T cells induced by immunization with uRNA Mix 3 or modRNA Mix 3 fused to alternative signal peptides

CD8⁺ T cells were isolated from mouse splenocytes pooled by treatment group and stimulated with antigen-specific peptide pools in presence of autologous bone-marrow derived dendritic cells. Pooled T cell samples were measured in duplicates or triplicates where possible; bars represent group replicates mean spot-forming units (SFU) $\pm$ standard deviation. Bars represent group mean values. Bars from left to right are buffer, sec uRNA Mix 3, SP1 uRNA Mix 3, SP2 uRNA Mix 3, sec modRNA Mix 3, SP1 modRNA Mix 3 and SP2 modRNA Mix 3. Y-axis shows IFN- γ -secreting cells / 1×10^5 CD8⁺ cells for all panels.

Detailed Description

Although the present disclosure is further described in more detail below, it is to be understood that this disclosure is not limited to the particular methodologies, protocols and reagents described herein as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not

intended to limit the scope of the present disclosure which will be limited only by the appended claims. Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art.

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

For example, the present disclosure describes combinations of sequence molecules which may have different levels of sequence identity to a specified sequence, e.g., (i) sequence molecule A comprising the sequence of SEQ ID NO: a, or a sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the sequence of SEQ ID NO: a, (ii) sequence molecule B comprising the sequence of SEQ ID NO: b, or a sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the sequence of SEQ ID NO: b etc. It should be understood that the sequence molecules may be combined in any of the identity levels specified. In some embodiments, the sequence molecules are combined such that the identity levels are identical; e.g., (i) sequence molecule A comprising the sequence of SEQ ID NO: a, (ii) sequence molecule B comprising the sequence of SEQ ID NO: b etc., or (i) sequence molecule A comprising a sequence having at least 90% identity to the sequence of SEQ ID NO: a, (ii) sequence molecule B comprising a sequence having at least 90% identity to the sequence of SEQ ID NO: b etc. In some embodiments, the identity levels are independently selected and are partially or entirely different from each other, i.e., the sequence molecules are combined such that the identity levels are not identical; e.g., (i) sequence molecule A comprising the sequence of SEQ ID NO: a, (ii) sequence molecule B comprising a sequence having at least 90% identity to the sequence of SEQ ID NO: b etc., or (i) sequence molecule A comprising a sequence having at least 90% identity to the sequence of SEQ ID NO: a, (ii) sequence molecule B comprising a sequence having at least 85% identity to the sequence of SEQ ID NO: b etc.

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

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated feature, element, member, integer or step or group of features, elements, members, integers or steps but not the exclusion of any other feature, element, member, integer or step or group of features, elements, members, integers or steps. The term "consisting essentially of" limits the scope of a claim or disclosure to the specified features, elements, members, integers, or steps and those that do not materially affect the basic and novel characteristic(s) of the claim or disclosure. The term "consisting of" limits the scope of a claim or disclosure to the specified features, elements, members, integers, or steps. The term "comprising" encompasses the term "consisting essentially of" which, in turn, encompasses the term "consisting of". Thus, at each occurrence in the present application, the term "comprising" may be replaced with the term "consisting essentially of" or "consisting of". Likewise, at each occurrence in the present application, the term "consisting essentially of" may be replaced with the term "consisting of".

The terms "a", "an" and "the" and similar references used in the context of describing the present disclosure (especially in the context of the claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by the context.

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

The use of any and all examples, or exemplary language (*e.g.*, "such as"), provided herein is intended merely to better illustrate the present disclosure and does not pose a limitation on the scope of the present disclosure otherwise claimed.

- 5 No language in the specification should be construed as indicating any non-claimed element essential to the practice of the present disclosure.

The term "optional" or "optionally" as used herein means that the subsequently described event, circumstance or condition may or may not occur, and that the description includes instances where said event, circumstance, or condition occurs and instances in which it does not occur.

- 10 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.

It should be noted for unambiguousness that whenever a sequence is referred to as being the sequence between the nucleotide at position x and the nucleotide at position y, the resulting sequence includes both the nucleotide at position x and the nucleotide at position y. Similarly, whenever a sequence is referred to as being the sequence between the amino acid at position x and the amino acid at position y, the resulting sequence includes both the amino acid at

position x and the amino acid at position y. Moreover, while the sequences described herein, in particular in the sequence listing, refer to DNA molecules, it is clear that when it is stated in the description or the claims that an RNA comprises a nucleotide sequence as described herein, in particular in the sequence listing, the nucleotide sequence referred to is actually identical to the base-sequence of the DNA molecule described herein, in particular in the sequence listing, e.g., represented in a SEQ ID NO referred to, except that thymine is replaced by uracil.

In the following, definitions and embodiments will be provided which apply to all aspects of the present disclosure. Terms which are defined in the following have the meanings as defined, unless otherwise indicated. Any undefined terms have their art recognized meanings.

- 10 *Mycobacterium tuberculosis* (*M. tuberculosis*) is a non-motile, slowly growing and rod shaped (2-4 µm in length and 0.2-0.5 µm in width) bacterium. *M. tuberculosis* is gram-positive, obligate aerobe, requires a host for growth and reproduction, and does not form spores.

- The term "tuberculosis" or "TB" is used to describe the infection caused by the infective agent "*Mycobacterium tuberculosis*" or "*M. tuberculosis*". Tuberculosis is a potentially fatal contagious disease that can affect almost any part of the body but is most frequently an infection of the lungs. *Mycobacterium tuberculosis*, the causative agent of tuberculosis, is transmitted by airborne droplet nuclei produced when an individual with active disease coughs, speaks, or sneezes. When inhaled, the droplet nuclei reach the alveoli of the lung. In susceptible individuals the organisms may then multiply and spread through lymphatics to the lymph nodes, and through the bloodstream to other sites such as the lung apices, bone marrow, kidneys, and meninges. The development of acquired immunity in 2 to 10 weeks results in a halt to bacterial multiplication. Lesions heal and the individual remains asymptomatic. *M. tuberculosis* bacteria can remain dormant (latent TB) in the body after infection for years, concealed in the phagocytosed cells, and never develop into the disease. Such an individual is said to have tuberculous infection without disease, and will show a positive tuberculin test. The clinical status of latent TB is traditionally associated with the transition of *M. tuberculosis* to a dormant state in response to non-optimal growth conditions *in vivo* due to activation of the host immune response.
- 25 Dormancy is a specific physiological state characterized by significant cessation of metabolic activity and growth, whereas resuscitation from dormancy is a process of restoring cell activity followed by bacterial multiplication, which in case of *M. tuberculosis* can lead to disease progression. The risk of developing active disease with clinical symptoms diminishes with time and may never occur, but is a lifelong risk. Approximately 5% of individuals with tuberculous infection progress to active disease.

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

- 35 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.

- "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 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 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:

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

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 μ M 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.

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.

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.

- 5 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.

- 15 The term "biological activity" means the response of a biological system to a molecule. Such biological systems may be, for example, a cell or an organism. In some embodiments, such response is therapeutically or pharmaceutically useful.

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.

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

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

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

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 by man.

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.

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

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

In some embodiments the degree of similarity, such as identity between a given amino acid sequence and an amino acid sequence which is a variant of said given amino acid sequence, will be at least about 60%, 70%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. In some embodiments, the degree of similarity or identity is given for an amino acid region which is at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90% or about 100% of the entire length of the reference amino acid sequence. For example, if the reference amino acid sequence consists of 200 amino acids, the degree of similarity or identity is given, e.g., for at least about 20, at least about 40, at least about 60, at least about 80, at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, or about 200 amino acids, in some embodiments continuous amino acids. In some embodiments, the degree of similarity or identity is given for the entire length of the

reference amino acid sequence. The alignment for determining sequence similarity, such as sequence identity, can be done with art known tools, such as using the best sequence alignment, for example, using Align, using standard settings, preferably EMBOSS::needle, Matrix: Blosum62, Gap Open 10.0, Gap Extend 0.5.

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

The terms "% identical" 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 antigens or antigenic sequences, one particular function is one or more immunogenic activities displayed by the amino acid sequence from which the fragment or variant is derived. The term "functional fragment" or "functional variant", as used herein, in particular refers to a variant molecule or sequence that comprises an amino acid sequence that is altered by one or more amino acids compared to the amino acid sequence of the parent molecule or sequence and that is still capable of fulfilling one or more of the functions of the parent molecule or sequence, *e.g.*, inducing an immune response. In some embodiments, the modifications in the amino acid sequence of the parent molecule or sequence do not significantly affect or alter the characteristics of the molecule or sequence. In different embodiments, the function of the functional fragment or functional variant may be reduced but still significantly present, *e.g.*, function of the functional fragment or functional variant may be at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the parent molecule or sequence. However, in other embodiments, function of the functional fragment or functional variant may be enhanced compared to the parent molecule or sequence.

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

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

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

encoding antigen is transiently transfected into cells. RNA can be transfected into cells to transiently express its coded protein.

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.

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

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.

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

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.

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 present disclosure. The instructional material of the kit of the present disclosure may, for example, be affixed to a container which contains the compositions/formulations of the present disclosure or be shipped together with a container which contains the compositions/formulations. 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. The term "set", e.g., as used herein in the context of "set of antigenic amino acid sequences", means more than 1, e.g., 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, or 8 or more.

The term "at least one" as used herein in the context of "at least one RNA molecule" means 1 or more, e.g., 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, or 8 or more. In some embodiments, the term "at least one" refers to 1, 2, 3, 4, 5, 6, 7, or 8. In some embodiments, "at least one RNA molecule" refers to a set of RNA molecules, e.g., 4 RNA molecules, wherein each RNA molecule encodes an amino acid sequence comprising at least two different Mtb antigens, immunogenic variants or fragments thereof, e.g., an amino acid sequence comprising two different Mtb antigens, immunogenic variants or fragments thereof. In some embodiments, such at least one RNA molecule or set of RNA molecules comprises the RNA molecules in a mixture, which mixture may be obtainable by transcribing in a common reaction a mixture of DNA templates encoding said RNA molecules.

Prodrugs of a particular compound described herein are those compounds that upon administration to an individual undergo chemical conversion under physiological conditions to provide the particular compound. Additionally, prodrugs can be converted to the particular compound by chemical or biochemical methods in an *ex vivo* environment. For example, prodrugs can be slowly converted to the particular compound when, for example, placed in a transdermal patch reservoir with a suitable enzyme or chemical reagent. Exemplary prodrugs are esters (using an alcohol or a carboxy group contained in the particular compound) or amides (using an amino or a carboxy group contained in the particular compound) which are hydrolyzable *in vivo*. Specifically, any amino group which is contained in the particular compound and which bears at least one hydrogen atom can be converted into a prodrug form. Typical N-prodrug forms include carbamates, Mannich bases, enamines, and enamines.

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 disclosure 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. Furthermore, in the present specification, a structural formula of a compound may represent a specific salt and/or solvate of said compound. It is to be understood, however, that the present disclosure includes all salts (e.g., pharmaceutically acceptable salts) and solvates (e.g., hydrates) and is not limited to the description of the specific salt and/or solvate.

"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 stereoisomers which are non-superimposable and which are not mirror-images of each other. "Tautomers" are structural isomers of the same chemical substance that spontaneously and reversibly interconvert into each other, even when pure, due to the migration of individual atoms or groups of atoms; *i.e.*, the tautomers are in a dynamic chemical equilibrium with each other. An example of tautomers are the isomers of the keto-enol-tautomerism. "Conformers" are stereoisomers that can be interconverted just by rotations about formally single bonds, and include - in particular - those leading to different 3-dimensional forms of (hetero)cyclic rings, such as chair, half-chair, boat, and twist-boat forms of cyclohexane.

The term "solvate" as used herein refers to an addition complex of a dissolved material in a solvent (such as an organic solvent (e.g., an aliphatic alcohol (such as methanol, ethanol, n-propanol, isopropanol), acetone, acetonitrile, ether, and the like), water or a mixture of two or more of these liquids), wherein the addition complex exists in the form of a crystal or mixed crystal. The amount of solvent contained in the addition complex may be stoichiometric or non-

5 stoichiometric. A "hydrate" is a solvate wherein the solvent is water.

In isotopically labeled compounds one or more atoms are replaced by a corresponding atom having the same number of protons but differing in the number of neutrons. For example, a hydrogen atom may be replaced by a deuterium or tritium atom. Exemplary isotopes which can be used in the present disclosure include deuterium, tritium, ^{11}C , ^{13}C , ^{14}C , ^{15}N , ^{18}F , ^{32}P , ^{32}S , ^{35}S , ^{36}Cl , and ^{125}I .

10 The term "average diameter" refers to the mean hydrodynamic diameter of particles as measured by dynamic light scattering (DLS) with data analysis using the so-called cumulant algorithm, which provides as results the so-called Z_{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} .

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

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:

$$25 \quad R_g = \left(\sum_{i=1}^n m_i \cdot s_i^2 / \sum_{i=1}^n m_i \right)^{1/2}$$

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).

30 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).

40 Thus, one procedure to determine the hydrodynamic radius of a particle or a population of particles (such as the hydrodynamic radius of particles contained in a sample or control composition as disclosed herein or the hydrodynamic

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

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

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.

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

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

- 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 $4\pi^2 n_0 (dn/dc)^2 \lambda_0^{-4} N_A^{-1}$, where n_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 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 using the following term or the reciprocal thereof:

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

wherein c , R_θ and K^* are as defined above. Preferably, the Debye plot is calculated using the following term or the reciprocal thereof:

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

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 hydrodynamic radius of the particles. A monochromatic light source, usually a laser, is shot through a polarizer and into a sample. The scattered light then goes through a second polarizer where it is detected and the resulting image is projected onto a screen. The particles in the solution are being hit with the light and diffract the light in all directions. The diffracted light from the particles can either interfere constructively (light regions) or destructively (dark regions). This process is repeated at short time intervals and the resulting set of speckle patterns are analyzed by an autocorrelator that compares the intensity of light at each spot over time.

- 40 The expression "static light scattering" or "SLS" as used herein refers to a technique to determine the size and size distribution profile of particles, in particular with respect to the radius of gyration of the particles, and/or the molar

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

Nucleic Acids

The term "nucleic acid" comprises deoxyribonucleic acid (DNA), ribonucleic acid (RNA), combinations thereof, and modified forms thereof. The term comprises genomic DNA, cDNA, mRNA, recombinantly produced and chemically synthesized molecules. In some embodiments, a nucleic acid is DNA. In some embodiments, a nucleic acid is RNA. In some embodiments, a nucleic acid is a mixture of DNA and RNA. A nucleic acid may be present as a single-stranded or double-stranded and linear or covalently circularly closed molecule. A nucleic acid can be isolated. The term "isolated nucleic acid" means, according to the present disclosure, that the nucleic acid (i) was amplified *in vitro*, for example via polymerase chain reaction (PCR) for DNA or *in vitro* transcription (using, *e.g.*, an RNA polymerase) for RNA, (ii) was produced recombinantly by cloning, (iii) was purified, for example, by cleavage and separation by gel electrophoresis, or (iv) was synthesized, for example, by chemical synthesis.

The term "nucleoside" (abbreviated herein as "N") relates to compounds which can be thought of as nucleotides without a phosphate group. While a nucleoside is a nucleobase linked to a sugar (*e.g.*, ribose or deoxyribose), a nucleotide is composed of a nucleoside and one or more phosphate groups. Examples of nucleosides include cytidine, uridine, pseudouridine, adenosine, and guanosine.

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

A modified purine (A or G) or pyrimidine (C, T, or U) base moiety is, 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 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.

DNA

Herein, the term "DNA" relates to a nucleic acid molecule which is entirely or at least substantially composed of 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.

RNA

The term "RNA" relates to a nucleic acid molecule which includes ribonucleotide residues. In preferred embodiments, the RNA contains all or a majority of ribonucleotide residues. As used herein, "ribonucleotide" refers to a nucleotide with a hydroxyl group at the 2'-position of a β -D-ribofuranosyl group. RNA encompasses without limitation, double stranded RNA, single stranded RNA, isolated RNA such as partially purified RNA, essentially pure RNA, synthetic RNA, recombinantly produced RNA, as well as modified RNA that differs from naturally occurring RNA by the addition, deletion, substitution and/or alteration of one or more nucleotides. Such alterations may refer to addition of non-nucleotide material to internal RNA nucleotides or to the end(s) of RNA. It is also contemplated herein that nucleotides in RNA may be non-standard nucleotides, such as chemically synthesized nucleotides or deoxynucleotides. For the present disclosure, these altered/modified nucleotides can be referred to as analogs of naturally occurring nucleotides, and the corresponding RNAs containing such altered/modified nucleotides (*i.e.*, altered/modified RNAs) can be referred to as analogs of naturally occurring RNAs. A molecule contains "a majority of ribonucleotide residues" if the content of ribonucleotide residues in the molecule is more than 50% (such as at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%), based on the total number of nucleotide residues in the molecule. The total number of nucleotide residues in a molecule is the sum of all nucleotide residues (irrespective of whether the nucleotide residues are standard (*i.e.*, naturally occurring) nucleotide residues or analogs thereof).

"RNA" includes mRNA, tRNA, ribosomal RNA (rRNA), small nuclear RNA (snRNA), self-amplifying RNA (saRNA), trans-amplifying RNA (taRNA), single-stranded RNA (ssRNA), dsRNA, inhibitory RNA (such as antisense ssRNA, small interfering RNA (siRNA), or microRNA (miRNA)), activating RNA (such as small activating RNA) and immunostimulatory RNA (isRNA). In some embodiments, "RNA" refers to mRNA.

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

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 the mRNA to fold and pair with itself to form double helices.

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

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

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

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

In some embodiments, RNA is *in vitro* transcribed RNA (IVT-RNA) and may be obtained by *in vitro* transcription of an appropriate DNA template. The promoter for controlling transcription can be any promoter for any RNA polymerase. 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 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 see José *et al.*, Future Microbiol., 2009, vol. 4, pp. 837–856). The total genome length of many alphaviruses typically ranges between 11,000 and 12,000 nucleotides, and the genomic RNA typically has a 5'-cap, and a 3' poly(A) tail. The genome of alphaviruses encodes non-structural proteins (involved in transcription, modification and replication of viral RNA and in protein modification) and structural proteins (forming the virus particle). There are typically two open reading frames (ORFs) in the genome. The four non-structural proteins (nsP1–nsP4) are typically encoded together by a first ORF beginning near the 5' terminus of the genome, while alphavirus structural proteins are encoded together by a second ORF which is found downstream of the first ORF and extends near the 3' terminus of the genome. Typically, the first ORF is larger than the second ORF, the ratio being roughly 2:1. In cells infected by an alphavirus, only the nucleic acid sequence encoding non-structural proteins is translated from the genomic RNA, while the genetic information encoding structural proteins is translatable from a subgenomic transcript, which is an RNA molecule that resembles eukaryotic messenger RNA (mRNA; Gould *et al.*, 2010, Antiviral Res., vol. 87 pp. 111–124). Following infection, *i.e.* at early stages of the viral life cycle, the (+) stranded genomic RNA directly acts like a messenger RNA for the translation of the open reading frame encoding the non-structural poly-protein (nsP1234).

Alphavirus-derived vectors have been proposed for delivery of foreign genetic information into target cells or target organisms. In simple approaches, the open reading frame encoding alphaviral structural proteins is replaced by an open reading frame encoding a protein of interest. Alphavirus-based trans-replication (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 certain alphaviral sequence elements to allow recognition and RNA synthesis by the alphaviral replicase.

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

5'-Cap

In some embodiments, the RNA (in particular, mRNA) described herein comprises a 5'-cap structure. In some embodiments, the RNA does not have uncapped 5'-triphosphates. In some embodiments, the RNA (in particular, mRNA) may comprise a conventional 5'-cap and/or a 5'-cap analog. The term "conventional 5'-cap" refers to a cap structure found on the 5'-end of an 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 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 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 m₂^{7,20}G(5')ppSp(5')G (referred to as beta-S-ARCA or β -S-ARCA)), as described in PCT/EP2019/056502. Providing an

RNA (in particular, mRNA) with a 5'-cap structure as described herein may be achieved by *in vitro* transcription of a DNA template in presence of a corresponding 5'-cap compound, wherein said 5'-cap structure is co-transcriptionally incorporated into the generated RNA (in particular, mRNA) strand, or the RNA (in particular, mRNA) may be generated, for example, by *in vitro* transcription, and the 5'-cap structure may be attached to the RNA post-transcriptionally using

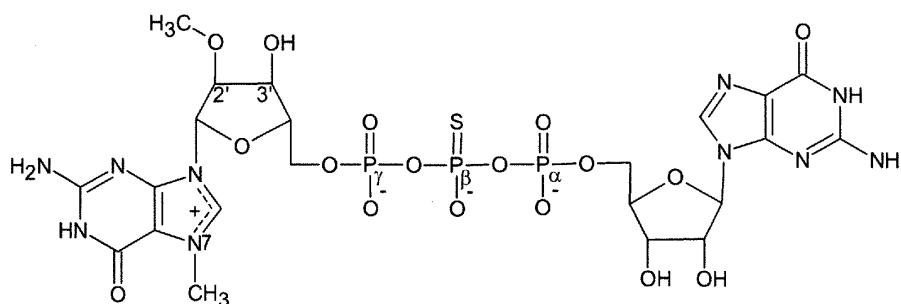
5 capping enzymes, for example, capping enzymes of vaccinia virus.

In some embodiments, the RNA (in particular, mRNA) 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, RNA comprises $m_2^{7,2'-O}G(5')ppSp(5')G$ (in particular its D1 diastereomer) as 5'-cap structure. In some embodiments, RNA comprises $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ as 5'-cap structure.

10 In some embodiments, the RNA (in particular, mRNA) 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

15 OCH_3 moiety at position 2'.

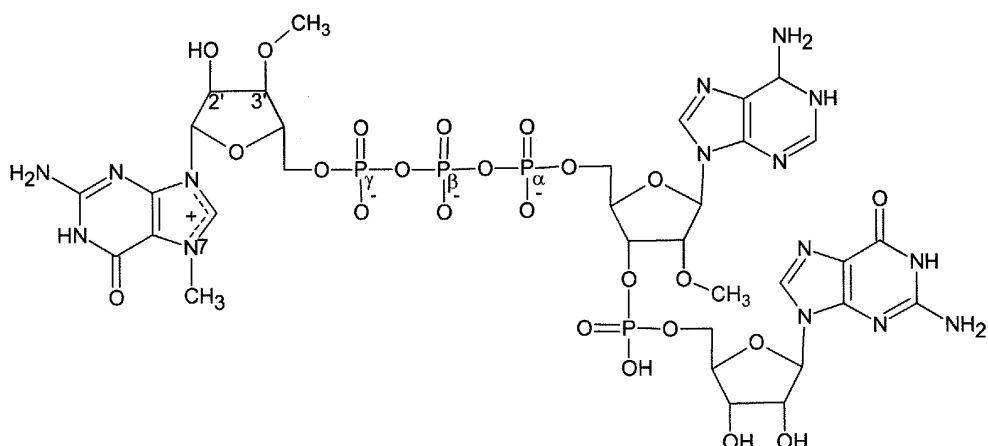
The 5'-cap analog beta-S-ARCA (β -S-ARCA) has the following structure:



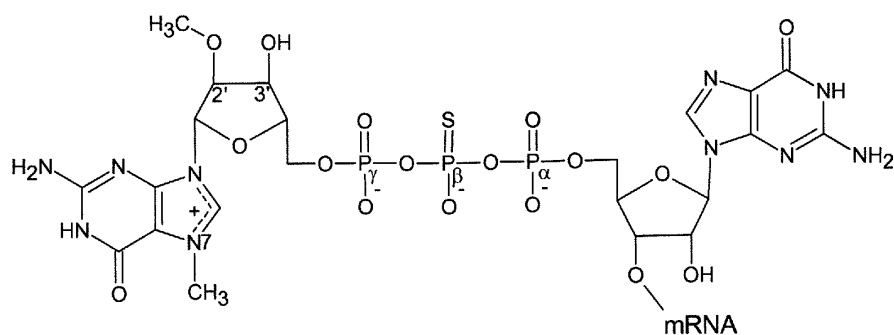
20 The "D1 diastereomer of beta-S-ARCA" or "beta-S-ARCA(D1)" is the diastereomer of beta-S-ARCA which elutes first on an HPLC column compared to the D2 diastereomer of beta-S-ARCA (beta-S-ARCA(D2)) and thus exhibits a shorter retention time. The HPLC preferably is an analytical HPLC. In some embodiments, 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 applied. In some embodiments, a gradient of methanol in ammonium acetate, for example, a 0-25% linear gradient of methanol

25 in 0.05 M ammonium acetate, pH = 5.9, within 15 min is used. UV-detection (VWD) can be performed at 260 nm and fluorescence detection (FLD) can be performed with excitation at 280 nm and detection at 337 nm.

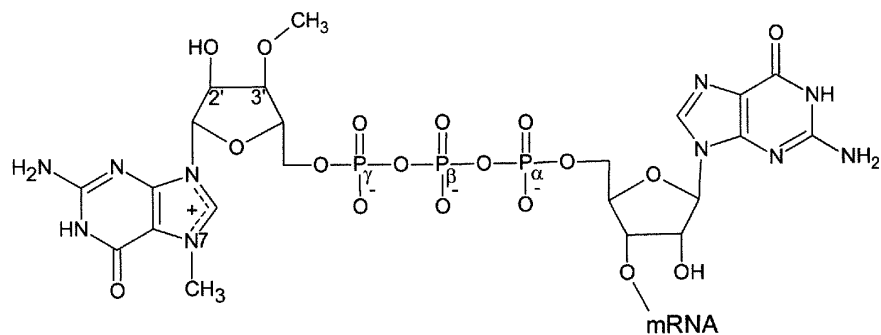
The 5'-cap analog $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ (also referred to as $m_2^{7,3'-O}G(5')ppp(5')m_1^{2'-O}ApG$) which is a building block of a cap1 has the following structure:



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

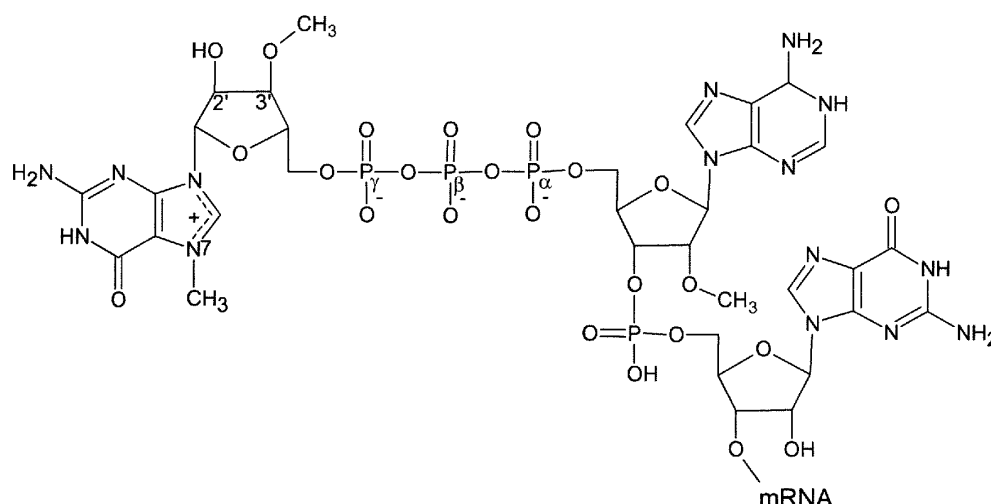


An exemplary cap0 mRNA comprising $m_2^{7,3'0}G(5')ppp(5')G$ and mRNA has the following structure:



5

An exemplary cap1 mRNA comprising $m_2^{7,3'-0}Gppp(m_1^{2'-0})ApG$ and mRNA has the following structure:



Poly-A tail

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 (in particular, mRNA) molecule. Poly-A tails or poly-A sequences are known to those of skill in the art and may follow the 3'-UTR in the RNAs (in particular, mRNAs) described herein. An uninterrupted poly-A tail is characterized by consecutive adenylate residues. In nature, an uninterrupted poly-A tail is typical. RNAs (in particular, mRNAs) 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 (uridylyte), G nucleotides (guanylyte), or C nucleotides (cytidylate). In this context, "consists of" means that all nucleotides in the poly-A tail, *i.e.*, 100% by number of nucleotides in the poly-A tail, are A nucleotides. The term "A nucleotide" or "A" refers to adenylate.

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

In some embodiments, the poly(A) cassette present in the coding strand of DNA essentially consists of dA nucleotides, but is interrupted by a random sequence of the four nucleotides (dA, dC, dG, and dT). Such random sequence may be 5 to 50, 10 to 30, or 10 to 20 nucleotides in length. Such a cassette is disclosed in WO 2016/005324 A1, hereby incorporated by reference. Any poly(A) cassette disclosed in WO 2016/005324 A1 may be used in the present disclosure. A poly(A) cassette that essentially consists of dA nucleotides, but is interrupted by a random sequence having an equal distribution of the four nucleotides (dA, dC, dG, dT) and having a length of *e.g.*, 5 to 50 nucleotides

shows, on DNA level, constant propagation of plasmid DNA in *E. coli* and is still associated, on RNA level, with the beneficial properties with respect to supporting RNA stability and translational efficiency is encompassed. Consequently, in some embodiments, the poly-A tail contained in an RNA (in particular, mRNA) molecule described herein essentially consists of A nucleotides, but is interrupted by a random sequence of the four nucleotides (A, C, G, U). Such random

5 sequence may be 5 to 50, 10 to 30, or 10 to 20 nucleotides in length.

In some embodiments, the poly(A) tail comprises 30 adenine nucleotides followed by 70 adenine nucleotides, wherein the 30 adenine nucleotides and 70 adenine nucleotides are separated by a linker sequence of 10 nucleotides.

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.

10 In some embodiments, a poly-A tail may comprise at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly-A tail may essentially consist of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly-A tail may consist of at least 20, at least 30,

15 some embodiments, the poly-A tail comprises the poly-A tail shown in SEQ ID NO: 59. 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.

Untranslated regions (UTR)

20 In some embodiments, RNA (in particular, mRNA) described in present disclosure comprises a 5'-UTR and/or a 3'-UTR. The term "untranslated region" or "UTR" relates to a region in a DNA molecule which is transcribed but is not translated into an amino acid sequence, or to the corresponding region in an RNA molecule, such as an mRNA molecule. An untranslated region (UTR) can be present 5' (upstream) of an open reading frame (5'-UTR) and/or 3' (downstream) of an open reading frame (3'-UTR). A 5'-UTR, if present, is located at the 5'-end, upstream of the start codon of a protein-

25 encoding region. A 5'-UTR is downstream of the 5'-cap (if present), *e.g.*, directly adjacent to the 5'-cap. A 3'-UTR, if present, is located at the 3'-end, downstream of the termination codon of a protein-encoding region, but the term "3'-UTR" does generally not include the poly-A sequence. Thus, the 3'-UTR is upstream of the poly-A sequence (if present), *e.g.*, directly adjacent to the poly-A sequence. Incorporation of a 3'-UTR into the 3'-non translated region of an RNA (preferably mRNA) molecule can result in an enhancement in translation efficiency. A synergistic effect may be achieved

30 by incorporating two or more of such 3'-UTRs (which are preferably arranged in a head-to-tail orientation; *cf.*, *e.g.*, Holtkamp *et al.*, Blood 108, 4009-4017 (2006)). The 3'-UTRs may be autologous or heterologous to the RNA (*e.g.*, mRNA) into which they are introduced. In certain embodiments, the 3'-UTR is derived from a globin gene or mRNA, such as a gene or mRNA of alpha2-globin, alpha1-globin, or beta-globin, *e.g.*, beta-globin, *e.g.*, human beta-globin. For example, the RNA (*e.g.*, mRNA) may be modified by the replacement of the existing 3'-UTR with or the insertion

35 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.

In some embodiments, a 5'-UTR is or comprises a modified human alpha-globin 5'-UTR. A particularly preferred 5'-UTR comprises the nucleotide sequence of SEQ ID NO: 56. In some embodiments, a 3'-UTR comprises a first sequence from the amino terminal enhancer of split (AES) messenger RNA and a second sequence from the mitochondrial

40 encoded 12S ribosomal RNA. A particularly preferred 3'-UTR comprises the nucleotide sequence of SEQ ID NO: 58.

In some embodiments, RNA comprises a 5'-UTR comprising the nucleotide sequence of SEQ ID NO: 56, 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: 56.

In some embodiments, RNA comprises a 3'-UTR comprising the nucleotide sequence of SEQ ID NO: 58, 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: 58.

5 Chemical modification

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

10 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), 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-taurinomethyl-2-thio-uridine (tm5s2U), 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-2-thio-uridine (m5s2U), 1-methyl-4-thio-pseudouridine (m1s4 ψ), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine (m3 ψ), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp3 ψ), 5-(isopentenylaminomethyl)uridine (inm5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm5s2U), α -thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m5Um), 2'-O-methyl-pseudouridine (ψ m), 2-thio-2'-O-methyl-uridine (s2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm5Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm5Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine (cmnm5Um), 3,2'-O-dimethyl-uridine (m3Um), 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm5Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, 5-[3-(1-E-propenylamino)uridine, or any other modified uridine known in the art.

35 An RNA (preferably mRNA) which is modified by pseudouridine (replacing partially or completely, preferably completely, uridine) is referred to herein as " ψ -modified", whereas the term "m1 ψ -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 m1 ψ - or m5U-modified RNAs usually exhibit decreased immunogenicity compared to their unmodified forms and, thus, are preferred in applications where the induction of an immune response is to be avoided or minimized. In some embodiments, the RNA (preferably mRNA) contains N(1)-methylpseudouridine replacing completely uridine.

Codon optimization and GC enrichment

The codons of the RNA (in particular, mRNA) described in the present disclosure may further be optimized, *e.g.*, to increase the GC content of the RNA and/or to replace codons which are rare in the cell (or subject) in which the peptide or polypeptide of interest is to be expressed by codons which are synonymous frequent codons in said cell (or subject).

- 5 In some embodiments, the amino acid sequence encoded by the RNA (in particular, mRNA) described in the present disclosure is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence. This also includes embodiments, wherein one or more sequence regions of the coding sequence are codon-optimized and/or increased in the G/C content compared to the corresponding sequence regions of the wild type coding sequence. In some embodiments, the codon-optimization and/or the increase
- 10 in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

- The term "codon-optimized" refers to the alteration of codons in the coding region of a nucleic acid molecule to reflect the typical codon usage of a host organism without preferably altering the amino acid sequence encoded by the nucleic acid molecule. Within the context of the present disclosure, coding regions may be codon-optimized for optimal expression in a subject to be treated using the RNA (in particular, mRNA) described herein. Codon-optimization is
- 15 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 (in particular, mRNA) may be modified such that codons for which frequently occurring tRNAs are available are inserted in place of "rare codons".

- In some embodiments, the guanosine/cytosine (G/C) content of the coding region of the RNA (in particular, mRNA) described herein is increased compared to the G/C content of the corresponding coding sequence of the wild type RNA, wherein the amino acid sequence encoded by the RNA is preferably not modified compared to the amino acid sequence
- 20 encoded by the wild type RNA. This modification of the RNA sequence is based on the fact that the sequence of any RNA region to be translated is important for efficient translation of that 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 codons for the stability can be determined (so-called alternative codon usage). Depending on the amino acid to be encoded by the RNA, there are various possibilities for modification of the RNA sequence, compared to its wild type sequence. In particular, codons which contain A and/or U nucleotides can be modified by substituting these codons by other codons, which code for the same amino acids but contain no A and/or U or contain a lower content of A and/or U nucleotides.

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

Non-immunogenic RNA

- 35 The term "non-immunogenic RNA" (such as "non-immunogenic mRNA") as used herein refers to RNA that does not induce a response by the immune system upon administration, *e.g.*, to a mammal, or induces a weaker response than would have been induced by the same RNA that differs only in that it has not been subjected to the modifications and treatments that render the non-immunogenic RNA non-immunogenic, *i.e.*, than would have been induced by standard RNA (stdRNA). In certain embodiments, non-immunogenic RNA is rendered non-immunogenic by incorporating
- 40 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 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 incorporation of modified nucleosides, any modified nucleoside may be used as long as it lowers or suppresses immunogenicity of the RNA.

- 5 Particularly preferred are modified nucleosides that suppress RNA-mediated activation of innate immune receptors. In some embodiments, the modified nucleosides comprise a replacement of one or more uridines with a nucleoside comprising a modified nucleobase. In some embodiments, the modified nucleobase is a modified uracil. In some
10 embodiments, the nucleoside comprising a modified nucleobase is selected from the group consisting of 3-methyl-uridine (m³U), 5-methoxy-uridine (mo⁵U), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s²U), 4-thio-uridine (s⁴U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho⁵U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-oxyacetic acid (cmo⁵U), uridine 5-oxyacetic acid methyl ester (mcmo⁵U), 5-carboxymethyl-uridine (cm⁵U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm⁵U), 5-carboxyhydroxymethyl-uridine methyl ester (mchm⁵U), 5-methoxycarbonylmethyl-uridine (mcm⁵U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm⁵s²U), 5-aminomethyl-2-thio-uridine (nm⁵s²U), 5-methylaminomethyl-uridine (nmn⁵U), 1-ethyl-pseudouridine, 5-methylaminomethyl-2-thio-uridine (nmn⁵s²U), 5-methylaminomethyl-2-seleno-uridine (nmn⁵se²U), 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-taurinomethyl-2-thio-uridine (tm⁵s²U), 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-2-thio-uridine (m⁵s²U), 1-methyl-4-thio-pseudouridine (m¹s⁴ψ), 4-thio-
20 1-methyl-pseudouridine, 3-methyl-pseudouridine (m³ψ), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m⁵D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp³U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp³ ψ), 5-(isopentenylaminomethyl)uridine (inm⁵U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm⁵s²U), α-thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m⁵Um), 2'-O-methyl-pseudouridine (ψm), 2-thio-2'-O-methyl-uridine (s²Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm⁵Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm⁵Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine (cmnm⁵Um), 3,2'-O-dimethyl-uridine (m³Um), 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm⁵Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-
30 uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, and 5-[3-(1-E-propenylamino)uridine. In certain embodiments, the nucleoside comprising a modified nucleobase is pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ) or 5-methyl-uridine (m5U), in particular N1-methyl-pseudouridine.

- 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
35 25%, at least 50%, at least 75%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% of the uridines.

- During synthesis of mRNA by *in vitro* transcription (IVT) using T7 RNA polymerase significant amounts of aberrant products, including double-stranded RNA (dsRNA) are produced due to unconventional activity of the enzyme. dsRNA induces inflammatory cytokines and activates effector enzymes leading to protein synthesis inhibition. Formation of
40 dsRNA can be limited during synthesis of mRNA 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 mRNA. 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.

5 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
10 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 dsRNA) 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
15 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
20 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.).

25 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.

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) TWEEN-20) containing 5% (w/v) skim milk
30 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 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
35 standard RNA with the same sequence. In some embodiments, translation is enhanced by a factor of 2-fold relative to its unmodified counterpart. In some embodiments, translation is enhanced by a 3-fold factor. In some embodiments, translation is enhanced by a 4-fold factor. In some embodiments, translation is enhanced by a 5-fold factor. In some embodiments, translation is enhanced by a 6-fold factor. In some embodiments, translation is enhanced by a 7-fold factor. In some embodiments, translation is enhanced by an 8-fold factor. In some embodiments, translation is
40 enhanced by a 9-fold factor. In some embodiments, translation is enhanced by a 10-fold factor. In some embodiments, translation is enhanced by a 15-fold factor. In some embodiments, translation is enhanced by a 20-fold factor. In some embodiments, translation is enhanced by a 50-fold factor. In some embodiments, translation is enhanced by a 100-fold factor. In some embodiments, translation is enhanced by a 200-fold factor. In some embodiments, translation is enhanced by a 500-fold factor. In some embodiments, translation is enhanced by a 1000-fold factor. In some

embodiments, translation is enhanced by a 2000-fold factor. In some embodiments, the factor is 10-1000-fold. In some embodiments, the factor is 10-100-fold. In some embodiments, the factor is 10-200-fold. In some embodiments, the factor is 10-300-fold. In some embodiments, the factor is 10-500-fold. In some embodiments, the factor is 20-1000-fold. In some embodiments, the factor is 30-1000-fold. In some embodiments, the factor is 50-1000-fold. In some
 5 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 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
 10 immunogenicity is reduced by a 3-fold factor. In some embodiments, innate immunogenicity is reduced by a 4-fold factor. In some embodiments, innate immunogenicity is reduced by a 5-fold factor. In some embodiments, innate immunogenicity is reduced by a 6-fold factor. In some embodiments, innate immunogenicity is reduced by a 7-fold factor. In some embodiments, innate immunogenicity is reduced by an 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
 15 factor. In some embodiments, innate immunogenicity is reduced by a 15-fold factor. In some embodiments, innate immunogenicity is reduced by a 20-fold factor. In some embodiments, innate immunogenicity is reduced by a 50-fold factor. In some embodiments, innate immunogenicity is reduced by a 100-fold factor. In some embodiments, innate immunogenicity is reduced by a 200-fold factor. In some embodiments, innate immunogenicity is reduced by a 500-fold factor. In some embodiments, innate immunogenicity is reduced by a 1000-fold factor. In some embodiments,
 20 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 without triggering a detectable innate immune response. In some
 25 embodiments, the term refers to a decrease such that the non-immunogenic RNA (especially mRNA) can be repeatedly administered without eliciting an innate immune response sufficient to detectably reduce production of the protein encoded by the non-immunogenic RNA. In some embodiments, the decrease is such that the non-immunogenic RNA (especially mRNA) can be repeatedly administered without eliciting an innate immune response sufficient to eliminate detectable production of the protein encoded by the non-immunogenic RNA.

"Immunogenicity" is the ability of a foreign substance, such as RNA, to provoke an immune response in the body of a
 30 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.

Antigen-coding RNA and use thereof for inducing an immune response

35 Generally, RNA (in particular, mRNA) described in the present disclosure comprises a nucleic acid sequence encoding a peptide or polypeptide comprising one or more *Mycobacterium tuberculosis* antigens, immunogenic variants or fragments thereof, for inducing an immune response against *Mycobacterium tuberculosis* in a subject. The peptide or polypeptide for inducing an immune response is also designated herein as "vaccine antigen" or simply "antigen".

In some embodiments, the RNA (in particular, mRNA) is translated into the respective protein upon entering cells of a
 40 subject being administered the RNA, e.g., muscle cells or antigen-presenting cells (APCs).

In some embodiments, the RNA encoding the vaccine antigen is expressed in cells of the subject to provide the vaccine antigen. In some embodiments, the RNA encoding the vaccine antigen is transiently expressed in cells of the subject.

In some embodiments, the vaccine antigen is presented in the context of MHC. In some embodiments, the vaccine antigen is secreted by cells of the subject.

In some embodiments, the RNA encoding the vaccine antigen is administered intramuscularly.

In some embodiments, the RNA encoding the vaccine antigen is administered systemically, e.g., intravenously. In some
5 embodiments, after systemic administration of the RNA encoding the vaccine antigen, expression of the RNA encoding the vaccine antigen in spleen occurs. In some embodiments, after systemic administration of the RNA encoding the vaccine antigen, expression of the RNA encoding the vaccine antigen in antigen presenting cells, preferably professional antigen presenting cells occurs. In some embodiments, the antigen presenting cells are selected from the group consisting of dendritic cells, macrophages and B cells. In some embodiments, after systemic administration of the RNA
10 encoding the vaccine antigen, no or essentially no expression of the RNA encoding the vaccine antigen in lung and/or liver occurs. In some embodiments, after systemic administration of the RNA encoding the vaccine antigen, expression of the RNA encoding the vaccine antigen in spleen is at least 5-fold the amount of expression in lung.

A vaccine antigen comprises an epitope for inducing an immune response against a disease-associated antigen, e.g., a protein of an infectious agent (e.g., Mtb antigen), in a subject. Accordingly, the vaccine antigen comprises an
15 antigenic sequence for inducing an immune response against a disease-associated antigen in a subject. Such antigenic sequence may correspond to a target antigen or disease-associated antigen, an immunogenic variant thereof, or an immunogenic fragment of the target antigen or disease-associated antigen or the immunogenic variant thereof. Thus, the antigenic sequence may comprise at least an epitope of a target antigen or disease-associated antigen or an immunogenic variant thereof.

The antigenic sequences, e.g., epitopes, suitable for use according to the disclosure typically may be derived from a target antigen, i.e. the antigen against which an immune response is to be elicited. For example, the antigenic sequences contained within the vaccine antigen may be a target antigen or a fragment or variant of a target antigen. The antigenic sequence or a procession product thereof, e.g., a fragment thereof, may bind to an antigen receptor such as TCR carried by immune effector cells. In some embodiments, the antigenic sequence is selected from the
20 group consisting of the antigen expressed by a target cell to which the immune effector cells are targeted or a fragment thereof, or a variant of the antigenic sequence or the fragment.

A vaccine antigen which may be provided to a subject according to the present disclosure by administering RNA encoding the vaccine antigen, preferably results in the induction of an immune response, e.g., in the stimulation, priming and/or expansion of immune effector cells, in the subject being provided the vaccine antigen. Said immune
30 response, e.g., stimulated, primed and/or expanded immune effector cells, is preferably directed against a target antigen, in particular a target antigen expressed in diseased cells, tissues and/or organs, i.e., a disease-associated antigen. Thus, a vaccine antigen may comprise the disease-associated antigen, or a fragment or variant thereof. In some embodiments, such fragment or variant is immunologically equivalent to the disease-associated antigen.

The term "immunologically equivalent" means that the immunologically equivalent molecule such as the
35 immunologically equivalent amino acid sequence exhibits the same or essentially the same immunological properties and/or exerts the same or essentially the same immunological effects, e.g., with respect to the type of the immunological effect. In the context of the present disclosure, the term "immunologically equivalent" is preferably used with respect to the immunological effects or properties of antigens or antigen variants used for immunization. For example, an amino acid sequence is immunologically equivalent to a reference amino acid sequence if said amino acid
40 sequence when exposed to the immune system of a subject induces an immune reaction having a specificity of reacting with the reference amino acid sequence. Thus, in some embodiments, a molecule which is immunologically equivalent to an antigen exhibits the same or essentially the same properties and/or exerts the same or essentially the same effects regarding the stimulation, priming and/or expansion of T cells as the antigen to which the T cells are targeted. In the context of the present disclosure, the term "fragment of an antigen" or "variant of an antigen" means an agent

which results in the induction of an immune response, e.g., in the stimulation, priming and/or expansion of immune effector cells, which immune response, e.g., stimulated, primed and/or expanded immune effector cells, targets the antigen, i.e. a disease-associated antigen, in particular when presented by diseased cells, tissues and/or organs. Thus, the vaccine antigen may correspond to or may comprise the disease-associated antigen, may correspond to or may

5 comprise a fragment of the disease-associated antigen or may correspond to or may comprise an antigen which is homologous to the disease-associated antigen or a fragment thereof. If the vaccine antigen comprises a fragment of the disease-associated antigen or an amino acid sequence which is homologous to a fragment of the disease-associated antigen said fragment or amino acid sequence may comprise an epitope of the disease-associated antigen to which the antigen receptor of the immune effector cells is targeted or a sequence which is homologous to an epitope of the

10 disease-associated antigen. Thus, according to the disclosure, a vaccine antigen may comprise an immunogenic fragment of a disease-associated antigen or an amino acid sequence being homologous to an immunogenic fragment of a disease-associated antigen. An "immunogenic fragment of an antigen" according to the disclosure preferably relates to a fragment of an antigen which is capable of inducing an immune response against, e.g., stimulating, priming and/or expanding immune effector cells carrying an antigen receptor binding to, the antigen or cells expressing the

15 antigen. It is preferred that the vaccine antigen (similar to the disease-associated antigen) provides the relevant epitope for binding by the antigen receptor present on the immune effector cells. In some embodiments, the vaccine antigen or a fragment thereof (similar to the disease-associated antigen) is expressed on the surface of a cell such as an antigen-presenting cell (optionally in the context of MHC) so as to provide the relevant epitope for binding by immune effector cells. The vaccine antigen may be a recombinant antigen.

20 In some embodiments of all aspects described herein, the RNA encoding the vaccine antigen is expressed in cells of a subject to provide the antigen or a procession product thereof for binding by the antigen receptor expressed by immune effector cells, said binding resulting in stimulation, priming and/or expansion of the immune effector cells.

An "antigen" according to the present disclosure covers any substance that will elicit an immune response and/or any substance against which an immune response or an immune mechanism such as a cellular response and/or humoral

25 response is directed. This also includes situations wherein the antigen is processed into antigen peptides and an immune response or an immune mechanism is directed against one or more antigen peptides, in particular if presented in the context of MHC molecules. In particular, an "antigen" relates to any substance, such as a peptide or polypeptide, that reacts specifically with antibodies or T-lymphocytes (T-cells). The term "antigen" may comprise a molecule that comprises at least one epitope, such as a T cell epitope. In some embodiments, an antigen is a molecule which,

30 optionally after processing, induces an immune reaction, which may be specific for the antigen (including cells expressing the antigen). In some embodiments, an antigen is a disease-associated antigen, such as an Mtb antigen.

In some embodiments, an antigen is presented or present on the surface of cells of the immune system such as antigen presenting cells like dendritic cells or macrophages. An antigen or a procession product thereof such as a T cell epitope is in some embodiments bound by an antigen receptor. Accordingly, an antigen or a procession product thereof may

35 react specifically with immune effector cells such as T-lymphocytes (T cells).

According to the present disclosure, an antigen or a combination of antigens described herein may induce an immune response, wherein the immune response may comprise a humoral or cellular immune response, or both. In the context of some embodiments of the present disclosure, the antigen is presented by a cell, such as by an antigen presenting cell, in the context of MHC molecules, which results in an immune response against the antigen. An antigen may be a

40 product which corresponds to or is derived from a naturally occurring antigen. According to the present disclosure, an antigen may correspond to a naturally occurring product.

The term "disease-associated antigen" is used in its broadest sense to refer to any antigen associated with a disease. In some embodiments, a disease-associated antigen is a molecule which contains epitopes that will stimulate a host's immune system to make a cellular antigen-specific immune response and/or a humoral antibody response against the

disease. Disease-associated antigens include pathogen-associated antigens, *i.e.*, antigens which are associated with infection by microbes, typically microbial antigens (such as bacterial or viral antigens, *e.g.*, Mtb antigens), or antigens associated with cancer, typically tumors, such as tumor antigens.

The term "bacterial antigen" refers to any bacterial component having antigenic properties, *i.e.* being able to provoke an immune response in an individual. The bacterial antigen may be derived from the cell wall or cytoplasm membrane of the bacterium. The term "bacterial antigen" includes Mtb antigens, *e.g.*, Mtb antigens as described herein.

The term "epitope" refers to an antigenic determinant in a molecule such as an antigen, *i.e.*, to a part in or fragment of the molecule that is recognized by the immune system, for example, that is recognized by antibodies, T cells or B cells, in particular when presented in the context of MHC molecules. An epitope of a protein may comprises a continuous or discontinuous portion of said protein and, *e.g.*, may be between about 5 and about 100, between about 5 and about 50, between about 8 and about 30, or about 10 and about 25 amino acids in length, for example, the epitope may be preferably 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids in length. In some embodiments, the epitope in the context of the present disclosure is a T cell epitope.

Terms such as "epitope", "fragment of an antigen", "immunogenic peptide" and "antigen peptide" are used interchangeably herein and, *e.g.*, may relate to an incomplete representation of an antigen which is, *e.g.*, capable of eliciting an immune response against the antigen or a cell expressing or comprising and presenting the antigen. In some embodiments, the terms relate to an immunogenic portion of an antigen. In some embodiments, it is a portion of an antigen that is recognized (*i.e.*, specifically bound) by a T cell receptor, in particular if presented in the context of MHC molecules. Certain preferred immunogenic portions bind to an MHC class I or class II molecule. The term "epitope" refers to a part or fragment of a molecule such as an antigen that is recognized by the immune system. For example, the epitope may be recognized by T cells, B cells or antibodies. An epitope of an antigen may include a continuous or discontinuous portion of the antigen and may be between about 5 and about 100, such as between about 5 and about 50, between about 8 and about 30, or between about 8 and about 25 amino acids in length, for example, the epitope may be 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids in length. In some embodiments, an epitope is between about 10 and about 25 amino acids in length. The term "epitope" includes T cell epitopes.

The term "T cell epitope" refers to a part or fragment of a protein that is recognized by a T cell when presented in the context of MHC molecules. The term "major histocompatibility complex" and the abbreviation "MHC" includes MHC class I and MHC class II molecules and relates to a complex of genes which is present in all vertebrates. MHC proteins or molecules are important for signaling between lymphocytes and antigen presenting cells or diseased cells in immune reactions, wherein the MHC proteins or molecules bind peptide epitopes and present them for recognition by T cell receptors on T cells. The proteins encoded by the MHC are expressed on the surface of cells, and display both self-antigens (peptide fragments from the cell itself) and non-self-antigens (*e.g.*, fragments of invading microorganisms) to a T cell. In the case of class I MHC/peptide complexes, the binding peptides are typically about 8 to about 10 amino acids long although longer or shorter peptides may be effective. In the case of class II MHC/peptide complexes, the binding peptides are typically about 10 to about 25 amino acids long and are in particular about 13 to about 18 amino acids long, whereas longer and shorter peptides may be effective.

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

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

In some embodiments, vaccine antigen, *i.e.*, an antigen whose inoculation into a subject induces an immune response, is recognized by an immune effector cell. In some embodiments, the vaccine antigen if recognized by an immune effector cell is able to induce in the presence of appropriate co-stimulatory signals, stimulation, priming and/or expansion of the immune effector cell carrying an antigen receptor recognizing the vaccine antigen. In the context of the embodiments of the present disclosure, the vaccine antigen may be, *e.g.*, presented or present on the surface of a cell, such as an antigen presenting cell.

In some embodiments, an antigen is expressed in a diseased cell (such as an infected cell).

In some embodiments, an antigen is presented by a diseased cell (such as an infected cell). In some embodiments, an antigen receptor is a TCR which binds to an epitope of an antigen presented in the context of MHC. In some embodiments, binding of a TCR when expressed by T cells and/or present on T cells to an antigen presented by cells such as antigen presenting cells results in stimulation, priming and/or expansion of said T cells. In some embodiments, binding of a TCR when expressed by T cells and/or present on T cells to an antigen presented on diseased cells results in cytolysis and/or apoptosis of the diseased cells, wherein said T cells release cytotoxic factors, *e.g.*, perforins and granzymes.

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

The terms "T cell" and "T lymphocyte" are used interchangeably herein and include T helper cells (CD4+ T cells) and cytotoxic T cells (CTLs, CD8+ T cells) which comprise cytolytic T cells. The term "antigen-specific T cell" or similar terms relate to a T cell which recognizes the antigen to which the T cell is targeted, in particular when presented on the surface of antigen presenting cells or diseased cells in the context of MHC molecules and preferably exerts effector functions of T cells. T cells are considered to be specific for antigen if the cells kill target cells expressing an antigen. T cell specificity may be evaluated using any of a variety of standard techniques, for example, within a chromium release assay or proliferation assay. Alternatively, synthesis of lymphokines (such as interferon- γ) can be measured.

In some embodiments, the term "target" shall mean an agent such as a cell or tissue which is a target for an immune response such as a cellular immune response. Targets include cells that present an antigen or an antigen epitope, *i.e.*, a peptide fragment derived from an antigen. In some embodiments, the target cell is a cell expressing an antigen and presenting said antigen with class I MHC.

"Antigen processing" refers to the degradation of an antigen into processing products which are fragments of said antigen (*e.g.*, the degradation of a polypeptide into peptides) and the association of one or more of these fragments (*e.g.*, via binding) with MHC molecules for presentation by cells, such as antigen-presenting cells to specific T-cells. Antigen-presenting cells can be distinguished in professional antigen presenting cells and non-professional antigen presenting cells.

The term "professional antigen presenting cells" relates to antigen presenting cells which constitutively express the Major Histocompatibility Complex class II (MHC class II) molecules required for interaction with naive T cells. If a T cell interacts with the MHC class II molecule complex on the membrane of the antigen presenting cell, the antigen presenting cell produces a co-stimulatory molecule inducing activation of the T cell. Professional antigen presenting cells comprise dendritic cells and macrophages.

The term "non-professional antigen presenting cells" relates to antigen presenting cells which do not constitutively express MHC class II molecules, but upon stimulation by certain cytokines such as interferon-gamma. Exemplary, non-professional antigen presenting cells include fibroblasts, thymic epithelial cells, thyroid epithelial cells, glial cells, pancreatic beta cells or vascular endothelial cells.

The term "dendritic cell" (DC) refers to a subtype of phagocytic cells belonging to the class of antigen presenting cells. In some embodiments, dendritic cells are derived from hematopoietic bone marrow progenitor cells. These progenitor cells initially transform into immature dendritic cells. These immature cells are characterized by high phagocytic activity

and low T cell activation potential. Immature dendritic cells constantly sample the surrounding environment for pathogens such as viruses and bacteria. Once they have come into contact with a presentable antigen, they become activated into mature dendritic cells and begin to migrate to the spleen or to the lymph node. Immature dendritic cells phagocytose pathogens and degrade their proteins into small pieces and upon maturation present those fragments at their cell surface using MHC molecules. Simultaneously, they upregulate cell-surface receptors that act as co-receptors in T cell activation such as CD80, CD86, and CD40 greatly enhancing their ability to activate T cells. They also upregulate CCR7, a chemotactic receptor that induces the dendritic cell to travel through the blood stream to the spleen or through the lymphatic system to a lymph node. Here they act as antigen-presenting cells and activate helper T cells and killer T cells as well as B cells by presenting them antigens, alongside non-antigen specific co-stimulatory signals. Thus, dendritic cells can actively induce a T cell- or B cell-related immune response. In some embodiments, the dendritic cells are splenic dendritic cells.

The term "macrophage" refers to a subgroup of phagocytic cells produced by the differentiation of monocytes. Macrophages which are activated by inflammation, immune cytokines or microbial products nonspecifically engulf and kill foreign pathogens within the macrophage by hydrolytic and oxidative attack resulting in degradation of the pathogen. Peptides from degraded proteins are displayed on the macrophage cell surface where they can be recognized by T cells, and they can directly interact with antibodies on the B cell surface, resulting in T and B cell activation and further stimulation of the immune response. Macrophages belong to the class of antigen presenting cells. In some embodiments, the macrophages are splenic macrophages.

By "antigen-responsive CTL" is meant a CD8⁺ T-cell that is responsive to an antigen or a peptide derived from said antigen, which is presented with class I MHC on the surface of antigen presenting cells.

According to the disclosure, CTL responsiveness may include sustained calcium flux, cell division, production of cytokines such as IFN- γ and TNF- α , up-regulation of activation markers such as CD44 and CD69, and specific cytolytic killing of tumor antigen expressing target cells. CTL responsiveness may also be determined using an artificial reporter that accurately indicates CTL responsiveness.

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

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

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

The terms "immune response" and "immune reaction" are used herein interchangeably in their conventional meaning and refer to an integrated bodily response to an antigen and may refer to a cellular immune response, a humoral immune response, or both. According to the disclosure, the term "immune response to" or "immune response against" with respect to an agent such as an antigen, cell or tissue, relates to an immune response such as a cellular response directed against the agent. An immune response may comprise one or more reactions selected from the group consisting of developing antibodies against one or more antigens and expansion of antigen-specific T-lymphocytes, such as CD4⁺ and CD8⁺ T-lymphocytes, e.g. CD8⁺ T-lymphocytes, which may be detected in various proliferation or cytokine production tests *in vitro*.

The terms "inducing an immune response" and "eliciting an immune response" and similar terms in the context of the present disclosure refer to the induction of an immune response, such as the induction of a cellular immune response,

a humoral immune response, or both. The immune response may be protective/preventive/prophylactic and/or therapeutic. The immune response may be directed against any immunogen or antigen or antigen peptide, such as against a pathogen-associated antigen (*e.g.*, an antigen of Mtb). "Inducing" in this context may mean that there was no immune response against a particular antigen or pathogen before induction, but it may also mean that there was a certain level of immune response against a particular antigen or pathogen before induction and after induction said immune response is enhanced. Thus, "inducing the immune response" in this context also includes "enhancing the immune response". In some embodiments, after inducing an immune response in an individual, said individual is protected from developing a disease such as an infectious disease or the disease condition is ameliorated by inducing an immune response.

The terms "cellular immune response", "cellular response", "cell-mediated immunity" or similar terms are meant to include a cellular response directed to cells characterized by expression of an antigen and/or presentation of an antigen with class I or class II MHC. The cellular response relates to cells called T cells or T lymphocytes which act as either "helpers" or "killers". The helper T cells (also termed CD4⁺ T cells) play a central role by regulating the immune response and the killer cells (also termed cytotoxic T cells, cytolytic T cells, CD8⁺ T cells or CTLs) kill cells such as diseased cells.

The term "humoral immune response" refers to a process in living organisms wherein antibodies are produced in response to agents and organisms, which they ultimately neutralize and/or eliminate. The specificity of the antibody response is mediated by T and/or B cells through membrane-associated receptors that bind antigen of a single specificity. Following binding of an appropriate antigen and receipt of various other activating signals, B lymphocytes divide, which produces memory B cells as well as antibody secreting plasma cell clones, each producing antibodies that recognize the identical antigenic epitope as was recognized by its antigen receptor. Memory B lymphocytes remain dormant until they are subsequently activated by their specific antigen. These lymphocytes provide the cellular basis of memory and the resulting escalation in antibody response when re-exposed to a specific antigen.

The term "antibody" as used herein, refers to an immunoglobulin molecule, which is able to specifically bind to an epitope on an antigen. In particular, the term "antibody" refers to a glycoprotein comprising at least two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds. The term "antibody" includes monoclonal antibodies, recombinant antibodies, human antibodies, humanized antibodies, chimeric antibodies and combinations of any of the foregoing. Each heavy chain is comprised of a heavy chain variable region (VH) and a heavy chain constant region (CH). Each light chain is comprised of a light chain variable region (VL) and a light chain constant region (CL). The variable regions and constant regions are also referred to herein as variable domains and constant domains, respectively. The VH and VL regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDRs), interspersed with regions that are more conserved, termed framework regions (FRs). Each VH and VL is composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The CDRs of a VH are termed HCDR1, HCDR2 and HCDR3, the CDRs of a VL are termed LCDR1, LCDR2 and LCDR3. The variable regions of the heavy and light chains contain a binding domain that interacts with an antigen. The constant regions of an antibody comprise the heavy chain constant region (CH) and the light chain constant region (CL), wherein CH can be further subdivided into constant domain CH1, a hinge region, and constant domains CH2 and CH3 (arranged from amino-terminus to carboxy-terminus in the following order: CH1, CH2, CH3). The constant regions of the antibodies may mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system (*e.g.*, effector cells) and the first component (C1q) of the classical complement system. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoactive portions of intact immunoglobulins. Antibodies are typically tetramers of immunoglobulin molecules. Antibodies may exist in a variety of forms including, for example, polyclonal antibodies, monoclonal antibodies, Fv, Fab and F(ab)₂, as well as single chain antibodies and humanized antibodies.

The term "immunoglobulin" relates to proteins of the immunoglobulin superfamily, such as to antigen receptors such as antibodies or the B cell receptor (BCR). The immunoglobulins are characterized by a structural domain, *i.e.*, the immunoglobulin domain, having a characteristic immunoglobulin (Ig) fold. The term encompasses membrane bound immunoglobulins as well as soluble immunoglobulins. Membrane bound immunoglobulins are also termed surface immunoglobulins or membrane immunoglobulins, which are generally part of the BCR. Soluble immunoglobulins are generally termed antibodies. Immunoglobulins generally comprise several chains, typically two identical heavy chains and two identical light chains which are linked via disulfide bonds. These chains are primarily composed of immunoglobulin domains, such as the V_L (variable light chain) domain, C_L (constant light chain) domain, V_H (variable heavy chain) domain, and the C_H (constant heavy chain) domains C_H1, C_H2, C_H3, and C_H4. There are five types of mammalian immunoglobulin heavy chains, *i.e.*, α, δ, ε, γ, and μ which account for the different classes of antibodies, *i.e.*, IgA, IgD, IgE, IgG, and IgM. As opposed to the heavy chains of soluble immunoglobulins, the heavy chains of membrane or surface immunoglobulins comprise a transmembrane domain and a short cytoplasmic domain at their carboxy-terminus. In mammals there are two types of light chains, *i.e.*, lambda and kappa. The immunoglobulin chains comprise a variable region and a constant region. The constant region is essentially conserved within the different isotypes of the immunoglobulins, wherein the variable part is highly divers and accounts for antigen recognition.

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

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

Embodiments of antigen-coding RNA

Generally, 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 RNA (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.

- 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
10 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.

- 15 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 one or more polypeptides which are not endogenous (*i.e.*, it is foreign) to the subject treated.

- In some embodiments, the RNA described herein (*e.g.*, contained in the compositions/formulations of the present disclosure and/or used in the methods of the present disclosure) is single-stranded RNA (in particular, mRNA) that may
20 be translated into the respective protein upon entering cells, *e.g.*, cells of a recipient, *e.g.*, muscle cells or antigen-presenting cells (APCs). In addition to wild-type or codon-optimized sequences encoding an antigen sequence, the RNA may contain one or more structural elements optimized for maximal efficacy of the RNA with respect to stability and translational efficiency (5' cap, 5' UTR, 3' UTR, poly(A)-tail). In 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
25 specific capping structure at the 5'-end of the RNA. 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
30 be used (see WO 2017/060314, herein incorporated by reference). 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. In some embodiments, the 5'-UTR comprises the nucleotide sequence of SEQ ID NO: 56, 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: 56. In some embodiments, the 3'-UTR comprises
35 the nucleotide sequence of SEQ ID NO: 58, 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: 58. In some embodiments, the poly(A) sequence comprises the nucleotide sequence of SEQ ID NO: 59, 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: 59.

- In some embodiments, the RNA described herein is not chemically modified, *i.e.* it solely contains naturally occurring
40 nucleosides, and preferably has the composition of naturally occurring RNA.

In some embodiments, the RNA described herein is modified for optimized efficacy of the RNA (*e.g.*, increased translation efficacy, decreased immunogenicity, and/or decreased cytotoxicity) (*e.g.*, by replacing (partially or completely, preferably completely) naturally occurring nucleosides (in particular uridine) with synthetic nucleosides (*e.g.*, modified nucleosides, *e.g.*, selected from the group consisting of pseudouridine (ψ), N1-methyl-pseudouridine

(m1ψ), and 5-methyl-uridine); and/or codon-optimization). In some embodiments, the RNA comprises a modified nucleoside in place of uridine. In some embodiments, the modified nucleoside replacing (partially or completely, preferably completely) uridine is selected from the group consisting of pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-methyl-uridine. In some embodiments, the RNA encoding the vaccine antigen has a coding sequence

5 (a) which is codon-optimized, (b) the G/C content of which is increased compared to the wild type coding sequence, or (c) both (a) and (b).

In some embodiments, the RNA described herein comprises a 5' cap, a 5' UTR, a 3' UTR, and a poly(A) sequence (e.g., as described above); is modified by replacing (partially or completely, preferably completely) uridine with modified nucleosides, e.g., selected from the group consisting of pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-

10 methyl-uridine; and has a coding sequence which is codon-optimized, and the G/C content of which is increased compared to the wild type coding sequence.

In some embodiments, if the present disclosure provides for a mixture of different RNA molecules, a composition comprising different RNA molecules or an administration of different RNA molecules, these different RNA molecules are present in approximately the same amount. Such different RNA molecules may be formulated in individual particulate

15 formulations, mixed particulate formulations, or combined particulate formulations as described herein.

The present disclosure provides RNA (in particular, mRNA) comprising a nucleic acid sequence encoding an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof. In some embodiments, RNA (in particular, mRNA) described in the present disclosure comprises a nucleic acid sequence encoding an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the

20 immunogenic variant thereof, and is capable of expressing said Mtb antigen, immunogenic variant, or immunogenic fragment, in particular if transferred into a cell or subject, preferably a human cell or subject. Thus, in some embodiments, the RNA (in particular, mRNA) described in the present disclosure contains a coding region (open reading frame (ORF)) encoding an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof.

In some embodiments, RNA comprises a nucleic acid sequence encoding more than one Mtb antigen, immunogenic variant thereof, or immunogenic fragment of the Mtb antigen or the immunogenic variant thereof, e.g., two, three, four or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof. In some embodiments, two or more of such Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof are present as a fusion

25 protein.

In some embodiments, the one or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof 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 Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof, or may comprise the one or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof and may comprise additional sequences such as secretion signals, extended-PK groups, tags and any other sequences. In some embodiments, the additional sequences are fused to the one or more Mtb

30 antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof, in some embodiments, separated by a linker. In these embodiments, the one or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof may be considered the pharmaceutically active peptide or polypeptide even if additional sequences support the function

35

40

or effect of the one or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof.

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

Specific examples of pharmaceutically active peptides and polypeptides include, but are not limited to, antigens for vaccination such as Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof.

Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof described herein can be prepared as fusion or chimeric polypeptides that include a portion which corresponds to one or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof and a heterologous polypeptide (*i.e.*, a polypeptide that is not an Mtb antigen, immunogenic variant thereof, or immunogenic fragment of the Mtb antigen or the immunogenic variant thereof).

According to some embodiments, an amino acid sequence enhancing antigen processing and/or presentation is fused, either directly or through a linker, to an antigenic peptide or polypeptide (antigenic sequence), *e.g.*, one or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof. Accordingly, in some embodiments, the RNA described herein comprises at least one coding region encoding an antigenic peptide or polypeptide and an amino acid sequence enhancing antigen processing and/or presentation.

Such amino acid sequences enhancing antigen processing and/or presentation are preferably located at the C-terminus of the antigenic peptide or polypeptide, without being limited thereto. Amino acid sequences enhancing antigen processing and/or presentation as defined herein preferably improve antigen processing and presentation. In some embodiments, the amino acid sequence enhancing antigen processing and/or presentation as defined herein includes, without being limited thereto, sequences derived from the human MHC class I complex (HLA-B51, haplotype A2, B27/B51, Cw2/Cw3), in particular a sequence comprising the amino acid sequence of SEQ ID NO: 54 or a functional variant thereof.

In some embodiments, an amino acid sequence enhancing antigen processing and/or presentation comprises the amino acid sequence of SEQ ID NO: 54, 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: 54, or a functional fragment of the amino acid sequence of SEQ ID NO: 54, 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: 54. In some embodiments, an amino acid sequence enhancing antigen processing and/or presentation comprises the amino acid sequence of SEQ ID NO: 54.

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

Furthermore, a secretory sequence, e.g., a sequence comprising the amino acid sequence of SEQ ID NO: 53, 78, or 79, may be fused to the N-terminus of the antigenic peptide or polypeptide.

Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof may be fused to an extended-PK group, which increases circulation half-life. Non-limiting examples of extended-PK groups are described herein. It should be understood that other PK groups that increase the circulation half-life of peptides or polypeptides such as Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic 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, or recombinant 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" polypeptide refers to a polypeptide moiety such as an Mtb antigen, immunogenic variant thereof, or immunogenic fragment of the Mtb antigen or the immunogenic variant thereof in combination with an extended-PK group. In some embodiments, the extended-PK polypeptide is a fusion protein in which a polypeptide moiety is linked or fused to an extended-PK group.

In certain embodiments, the serum half-life of an extended-PK polypeptide is increased relative to the polypeptide alone (i.e., the polypeptide not fused to an extended-PK group). In certain embodiments, the serum half-life of the extended-PK polypeptide is at least 20%, at least 40%, at least 60%, at least 80%, at least 100%, at least 120%, at least 150%, at least 180%, at least 200%, at least 400%, at least 600%, at least 800%, or at least 1000% longer relative to the serum half-life of the polypeptide alone. In certain embodiments, the serum half-life of the extended-PK polypeptide 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 polypeptide alone. In certain embodiments, the serum half-life of the extended-PK polypeptide is at least 10 hours, 15 hours, 20 hours, 25 hours, 30 hours, 35 hours, 40 hours, 50 hours, 60 hours, 70 hours, 80 hours, 90 hours, 100 hours, 110 hours, 120 hours, 130 hours, 135 hours, 140 hours, 150 hours, 160 hours, or 200 hours.

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

Pharmacokinetic Analysis: A Practical Approach (1996). Reference is also made to Gibaldi, M. et al., Pharmacokinetics, 2nd Rev. Edition, Marcel Dekker (1982).

In certain embodiments, the extended-PK group includes serum albumin, or fragments thereof or variants of the serum albumin or fragments thereof (all of which for the purpose of the present disclosure are comprised by the term "albumin"). Polypeptides described herein may be fused to albumin (or a fragment or variant thereof) to form albumin fusion proteins. Such albumin fusion proteins are described in U.S. Publication No. 20070048282.

As used herein, "albumin fusion protein" refers to a protein formed by the fusion of at least one molecule of albumin (or a fragment or variant thereof) to at least one molecule of a protein such as a therapeutic protein, in particular an Mtb antigen, immunogenic variant thereof, or immunogenic fragment of the Mtb antigen or the immunogenic variant thereof. 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 the albumin fusion protein may be from a different animal than the therapeutic protein portion.

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

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

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

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

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

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

- 15 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 Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof.

20 In some embodiments, the therapeutic protein(s) is (are) joined to the albumin through (a) peptide linker(s). A peptide linker 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 peptide linker may consist of amino acids such that it is flexible or more rigid. The linker sequence may be cleavable by a protease or chemically.

25 As used herein, the term "Fc region" refers to the portion of a native immunoglobulin formed by the respective Fc domains (or Fc moieties) of its two heavy chains. As used herein, the term "Fc domain" refers to a portion or fragment of a single immunoglobulin (Ig) heavy chain wherein the Fc domain does not comprise an Fv domain. In certain embodiments, an Fc domain begins in the hinge region just upstream of the papain cleavage site and ends at the C-terminus of the antibody. Accordingly, a complete Fc domain comprises at least a hinge domain, a CH2 domain, and a CH3 domain. In certain embodiments, an Fc domain comprises at least one of: a hinge (e.g., upper, middle, and/or lower hinge region) domain, a CH2 domain, a CH3 domain, a CH4 domain, or a variant, portion, or fragment thereof.

30 In certain embodiments, an Fc domain comprises a complete Fc domain (i.e., a hinge domain, a CH2 domain, and a CH3 domain). In certain embodiments, an Fc domain comprises a hinge domain (or portion thereof) fused to a CH3 domain (or portion thereof). In certain embodiments, an Fc domain comprises a CH2 domain (or portion thereof) fused to a CH3 domain (or portion thereof). In certain embodiments, an Fc domain consists of a CH3 domain or portion thereof. In certain embodiments, an Fc domain consists of a hinge domain (or portion thereof) and a CH3 domain (or portion thereof). In certain embodiments, an Fc domain consists of a CH2 domain (or portion thereof) and a CH3 domain. In certain embodiments, an Fc domain consists of a hinge domain (or portion thereof) and a CH2 domain (or portion thereof). In certain embodiments, an Fc domain lacks at least a portion of a CH2 domain (e.g., all or part of a CH2 domain). An Fc domain herein generally refers to a polypeptide comprising all or part of the Fc domain of an immunoglobulin heavy-chain. This includes, but is not limited to, polypeptides comprising the entire CH1, hinge, CH2,

- and/or CH3 domains as well as fragments of such peptides comprising only, e.g., the hinge, CH2, and CH3 domain. The Fc domain may be derived from an immunoglobulin of any species and/or any subtype, including, but not limited to, a human IgG1, IgG2, IgG3, IgG4, IgD, IgA, IgE, or IgM antibody. The Fc domain encompasses native Fc and Fc variant molecules. As set forth herein, it will be understood by one of ordinary skill in the art that any Fc domain may
- 5 be modified such that it varies in amino acid sequence from the native Fc domain of a naturally occurring immunoglobulin molecule. In certain embodiments, the Fc domain has reduced effector function (e.g., FcγR binding). The Fc domains of a polypeptide described herein may be derived from different immunoglobulin molecules. For example, an Fc domain of a polypeptide may comprise a CH2 and/or CH3 domain derived from an IgG1 molecule and a hinge region derived from an IgG3 molecule. In another example, an Fc domain can comprise a chimeric hinge region
- 10 derived, in part, from an IgG1 molecule and, in part, from an IgG3 molecule. In another example, an Fc domain can comprise a chimeric hinge derived, in part, from an IgG1 molecule and, in part, from an IgG4 molecule.
- In certain embodiments, an extended-PK group includes an Fc domain or fragments thereof or variants of the Fc domain or fragments thereof (all of which for the purpose of the present disclosure are comprised by the term "Fc domain"). The Fc domain does not contain a variable region that binds to antigen. Fc domains suitable for use in the
- 15 present disclosure may be obtained from a number of different sources. In certain embodiments, an Fc domain is derived from a human immunoglobulin. In certain embodiments, the Fc domain is from a human IgG1 constant region. It is understood, however, that the Fc domain may be derived from an immunoglobulin of another mammalian species, including for example, a rodent (e.g. a mouse, rat, rabbit, guinea pig) or non-human primate (e.g. chimpanzee, macaque) species.
- 20 Moreover, the Fc domain (or a fragment or variant thereof) may be derived from any immunoglobulin class, including IgM, IgG, IgD, IgA, and IgE, and any immunoglobulin isotype, including IgG1, IgG2, IgG3, and IgG4.
- A variety of Fc domain gene sequences (e.g., mouse and human constant region gene sequences) are available in the form of publicly accessible deposits. Constant region domains comprising an Fc domain sequence can be selected lacking a particular effector function and/or with a particular modification to reduce immunogenicity. Many sequences
- 25 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 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
- 30 WO2009/133208, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is transferrin, as disclosed in US 7,176,278 and US 8,158,579, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is a serum immunoglobulin binding protein such as those disclosed in US2007/0178082, US2014/0220017, and US2017/0145062, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is a fibronectin (Fn)-based scaffold domain
- 35 protein that binds to serum albumin, such as those disclosed in US2012/0094909, which is herein incorporated by reference in its entirety. Methods of making fibronectin-based scaffold domain proteins are also disclosed in US2012/0094909. A non-limiting example of a Fn3-based extended-PK group is Fn3(HSA), i.e., a Fn3 protein that binds to human serum albumin.
- In certain aspects, the extended-PK polypeptide, suitable for use according to the disclosure, can employ one or more
- 40 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 polypeptide moiety, e.g., an Mtb antigen, immunogenic variant thereof, or immunogenic fragment of the Mtb antigen or the immunogenic variant thereof) in a linear amino acid sequence of a polypeptide chain. For example, peptide linkers may be used to connect an Mtb antigen,

immunogenic variant thereof, or immunogenic fragment of the Mtb antigen or the immunogenic variant thereof to a HSA domain.

Linkers suitable for fusing the extended-PK group to, e.g., an Mtb antigen, immunogenic variant thereof, or immunogenic fragment of the Mtb antigen or the immunogenic variant thereof are well known in the art. Exemplary

- 5 linkers include glycine-serine-polypeptide linkers, glycine-proline-polypeptide linkers, and proline-alanine polypeptide linkers. In certain embodiments, the linker is a glycine-serine-polypeptide linker, i.e., a peptide that consists of glycine and serine residues.

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

- 10 **cap:** 5'-cap structure, e.g., selected from the group consisting of $m_2^{7,2'0}G(5')ppSp(5')G$ (in particular its D1 diastereomer), $m_2^{7,3'0}G(5')ppp(5')G$, and $m_2^{7,3'0}Gppp(m_1^{2'-0})ApG$.

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

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

- 20 **Antigen:** Sequences encoding the respective vaccine antigen(s)/epitope(s), i.e., one or more Mtb antigens, immunogenic variants thereof, or immunogenic fragments of the Mtb antigens or the immunogenic variants thereof.

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

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

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

In some embodiments, vaccine RNA described herein has one of the following structures:

- 30 cap-hAg-Kozak-Antigen-FI-A30L70
cap-hAg-Kozak-sec-Antigen-FI-A30L70
cap-hAg-Kozak-sec-Antigen-MITD-FI-A30L70

In some embodiments, vaccine antigen described herein has the structure:

- 35 sec-Antigen
sec-Antigen-MITD

- 40 In some embodiments, hAg-Kozak comprises the nucleotide sequence of SEQ ID NO: 56. In some embodiments, sec of the encoded vaccine antigen/epitope comprises the amino acid sequence of SEQ ID NO: 53, 78, or 79. In some embodiments, MITD of the encoded vaccine antigen/epitope comprises the amino acid sequence of SEQ ID NO: 54. In some embodiments, FI comprises the nucleotide sequence of SEQ ID NO: 58. In some embodiments, A30L70 comprises the nucleotide sequence of SEQ ID NO: 59.

In some embodiments, the different elements (sec, Antigen, MITD) may be linked by one or more GS linkers. In some embodiments, a GS linker of the encoded vaccine antigen/epitope comprises the amino acid sequence of SEQ ID NO: 55.

In some embodiments, the sequence encoding the vaccine antigen/epitope comprises a modified nucleoside replacing (partially or completely, preferably completely) uridine, wherein the modified nucleoside is selected from the group consisting of pseudouridine (ψ), N1-methyl-pseudouridine (m1 ψ), and 5-methyl-uridine.

In some embodiments, the sequence encoding the vaccine antigen/epitope is codon-optimized.

- 5 In some embodiments, the G/C content of the sequence encoding the vaccine antigen/epitope is increased compared to the wild type coding sequence.

In some embodiments, the RNA (in particular, mRNA) described herein comprises:

- a 5' UTR comprising the nucleotide sequence of SEQ ID NO: 56, 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: 56;
- 10 - a 3' UTR comprising the nucleotide sequence of SEQ ID NO: 58, 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: 58; and
- a poly-A sequence comprising the nucleotide sequence of SEQ ID NO: 59.

In some embodiments, the RNA (in particular, mRNA) described herein comprises:

- m₂^{7,3'-O}Gppp(m₁^{2'-O}) ApG as capping structure at the 5'-end of the mRNA;
- 15 - a 5' UTR comprising the nucleotide sequence of SEQ ID NO: 56, 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: 56;
- a 3' UTR comprising the nucleotide sequence of SEQ ID NO: 58, 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: 58; and
- a poly-A sequence comprising the nucleotide sequence of SEQ ID NO: 59.

- 20 In some embodiments, the RNA is unmodified. In some embodiments, the RNA is modified. In some embodiments, the RNA comprises N1-methyl-pseudouridine (m1 ψ) in place of at least one uridine (e.g., in place of each uridine).

In some embodiments, the RNA (in particular, mRNA) described herein comprises:

- m₂^{7,3'-O}Gppp(m₁^{2'-O}) ApG as capping structure at the 5'-end of the mRNA;
- a 5' UTR comprising the nucleotide sequence of SEQ ID NO: 56, or a nucleotide sequence having at least
- 25 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 56;
- a 3' UTR comprising the nucleotide sequence of SEQ ID NO: 58, 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: 58; and
- a poly-A sequence comprising the nucleotide sequence of SEQ ID NO: 59; and
- N1-methyl-pseudouridine (m1 ψ) in place of at least one uridine (e.g., in place of each uridine).

- 30 In some embodiments, a vaccine antigen or epitope described herein is derived from *Mycobacterium tuberculosis*.

In some embodiments, a vaccine antigen or epitope described herein is derived from a *Mycobacterium tuberculosis* protein, an immunogenic variant thereof, or an immunogenic fragment of the *Mycobacterium tuberculosis* protein or the immunogenic variant thereof. Thus, in some embodiments, the RNA, e.g., mRNA, used in the present disclosure encodes an amino acid sequence comprising an Mtb protein, an immunogenic variant thereof, or an immunogenic

- 35 fragment of the Mtb protein or the immunogenic variant thereof.

In some embodiments, a vaccine antigen or epitope described herein is derived from an Mtb protein from the acute phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb protein from the acute phase of the Mtb life cycle or the immunogenic variant thereof.

In some embodiments, the Mtb protein from the acute phase of the Mtb life cycle is Ag85A or ESAT6.

- 40 In some embodiments, a vaccine antigen or epitope described herein is derived from an Mtb protein from the latent phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb protein from the latent phase of the Mtb life cycle or the immunogenic variant thereof.

In some embodiments, the Mtb protein from the latent phase of the Mtb life cycle is VapB47 or Hrp1.

In some embodiments, a vaccine antigen or epitope described herein is derived from an Mtb protein from the resuscitation phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb protein from the resuscitation phase of the Mtb life cycle or the immunogenic variant thereof.

In some embodiments, the Mtb protein from the resuscitation phase of the Mtb life cycle is RpfA or RpfD.

- 5 In some embodiments, RNA (in particular, mRNA) described herein (e.g., contained in the compositions/formulations of the present disclosure and/or used in the methods of the present disclosure) may be presented as a product containing the vaccine RNA as active substance and other ingredients comprising: ALC-0315 ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate), ALC-0159 (2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide), 1,2-Distearoyl-sn-glycero-3-phosphocholine (DSPC), and cholesterol.

- 10 In some embodiments, the RNA (in particular, mRNA) described herein is formulated or is to be formulated as a liquid, a solid, or a combination thereof.

In some embodiments, the RNA (in particular, mRNA) described herein is formulated or is to be formulated for injection.

In some embodiments, the RNA (in particular, mRNA) described herein is formulated or is to be formulated for intramuscular administration.

- 15 In some embodiments, the RNA (in particular, mRNA) described herein is formulated or is to be formulated as a composition, e.g., a pharmaceutical composition.

In some embodiments, the composition comprises a cationically ionizable lipid.

In some embodiments, the composition comprises a cationically ionizable lipid and one or more additional lipids. In some embodiments, the one or more additional lipids are selected from polymer-conjugated lipids, neutral lipids, and combinations thereof. In some embodiments, the neutral lipids include phospholipids, steroid lipids, and combinations thereof. In some embodiments, the one or more additional lipids are a combination of a polymer-conjugated lipid, a phospholipid, and a steroid lipid.

- 20 In some embodiments, the composition comprises a cationically ionizable lipid; a polymer-conjugated lipid which is a PEG-conjugated lipid; cholesterol; and a phospholipid. In some embodiments, the phospholipid is DSPC. In some embodiments, the phospholipid is DOPE.

In some embodiments, the composition comprises a cationically ionizable lipid; a polymer-conjugated lipid which is 2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide; cholesterol; and a phospholipid. In some embodiments, the phospholipid is DSPC. In some embodiments, the phospholipid is DOPE.

- 30 In some embodiments, the composition comprises a cationically ionizable lipid which is ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate); a polymer-conjugated lipid which is 2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide; cholesterol; and a phospholipid. In some embodiments, the phospholipid is DSPC. In some embodiments, the phospholipid is DOPE. In some embodiments, at least a portion of (i) the RNA, (ii) the cationically ionizable lipid, and if present, (iii) the one or more additional lipids is present in particles. In some embodiments, the particles are nanoparticles, such as lipid nanoparticles (LNPs).

- 35 In some embodiments, the composition, in particular the pharmaceutical composition, is a vaccine.

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

In some embodiments, the RNA and/or the composition, in particular the pharmaceutical composition, is/are a component of a kit.

- 40 In some embodiments, the kit further comprises instructions for use of the RNA for inducing an immune response against *Mycobacterium tuberculosis* in a subject.

In some embodiments, the kit further comprises instructions for use of the RNA for therapeutically or prophylactically treating a *Mycobacterium tuberculosis* infection in a subject.

In some embodiments, the subject is a human.

In some embodiments, the RNA (in particular, mRNA), e.g., RNA encoding vaccine antigen, described in the present disclosure is non-immunogenic. RNA encoding an immunostimulant may be administered according to the present disclosure to provide an adjuvant effect. The RNA encoding an immunostimulant may be standard RNA or non-immunogenic RNA.

5

Embodiments of Mycobacterium tuberculosis (Mtb) antigens

The present disclosure describes Mtb antigens, immunogenic variants thereof, and immunogenic fragments of the Mtb antigens or the immunogenic variants thereof (referred to as "Mtb antigens" herein) and RNA encoding these antigens. Mtb antigens described herein may be from any phase of the Mtb life cycle. Such phases of the Mtb life cycle include the acute phase of the Mtb life cycle, the latent phase of the Mtb life cycle, and the resuscitation phase of the Mtb life cycle. Mtb antigens from the acute phase of the Mtb life cycle include Ag85A and ESAT6. Mtb antigens from the latent phase of the Mtb life cycle include VapB47 and Hrp1. Mtb antigens from the resuscitation phase of the Mtb life cycle include RpfA and RpfD. Further Mtb antigens useful herein include Mtb32a, Mtb39a, M72 which is a fusion protein of Mtb32a and Mtb39a, as well as HbhA.

15

Ag85A

Characteristics of the Mtb antigen Ag85A are described in Table 1, below. In some embodiments, the Mtb antigen Ag85A comprises the amino acid sequence according to SEQ ID NO: 2.

In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof.

In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof.

In some embodiments, an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 298 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 positions 2 to 298 of SEQ ID NO: 20, or an immunogenic fragment of the amino acid sequence of positions 2 to 298 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 positions 2 to 298 of SEQ ID NO: 20.

In some embodiments, an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 298 of SEQ ID NO: 20.

In some embodiments, an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 323 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 positions 27 to 323 of SEQ ID NO: 44, or an immunogenic fragment of the amino acid sequence of positions 27 to 323 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 positions 27 to 323 of SEQ ID NO: 44.

In some embodiments, an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 323 of SEQ ID NO: 44.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, or a fragment of the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 894 of SEQ ID NO: 19.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43.

20

ESAT6

Characteristics of the Mtb antigen ESAT6 are described in Table 1, below. In some embodiments, the Mtb antigen ESAT6 comprises the amino acid sequence according to SEQ ID NO: 4.

In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof.

In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof.

In some embodiments, an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 95 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 positions 2 to 95 of SEQ ID NO: 4, or an immunogenic fragment of the amino acid sequence of positions 2 to 95 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 positions 2 to 95 of SEQ ID NO: 4.

In some embodiments, an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 95 of SEQ ID NO: 4.

In some embodiments, an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, or an immunogenic fragment of the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46.

In some embodiments, an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46.

- 5 In some embodiments, an RNA sequence encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, or a fragment of the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21.
- 10 In some embodiments, an RNA sequence encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 285 of SEQ ID NO: 21.
- In some embodiments, an RNA sequence encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, or a fragment of the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45.
- 15 In some embodiments, an RNA sequence encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45.
- 20 In some embodiments, an RNA sequence encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45.

VapB47

- 25 Characteristics of the Mtb antigen VapB47 are described in Table 1, below. In some embodiments, the Mtb antigen VapB47 comprises the amino acid sequence according to SEQ ID NO: 6.
- In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof.
- In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof.
- 30 In some embodiments, an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6, or an immunogenic fragment of the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6.
- 35 In some embodiments, an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 99 of SEQ ID NO: 6.
- 40 In some embodiments, an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, or an immunogenic fragment of the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, or the amino acid sequence having at

least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52.

In some embodiments, an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, or a fragment of the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 297 of SEQ ID NO: 22.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, or a fragment of the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51.

Hrp1

Characteristics of the Mtb antigen Hrp1 are described in Table 1, below. In some embodiments, the Mtb antigen Hrp1 comprises the amino acid sequence according to SEQ ID NO: 8.

In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof.

In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof.

In some embodiments, an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8, or an immunogenic fragment of the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8.

In some embodiments, an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 143 of SEQ ID NO: 8.

In some embodiments, an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 324 to 465

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 positions 324 to 465 of SEQ ID NO: 44, or an immunogenic fragment of the amino acid sequence of positions 324 to 465 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 positions 324 to 465 of SEQ ID NO: 44.

- 5 In some embodiments, an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 324 to 465 of SEQ ID NO: 44.

- In some embodiments, an RNA sequence encoding an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, or a fragment of the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23.

- 10 In some embodiments, an RNA sequence encoding an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 429 of SEQ ID NO: 23.

- 15 In some embodiments, an RNA sequence encoding an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43.

- 20 In some embodiments, an RNA sequence encoding an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43.

RpfA

- Characteristics of the Mtb antigen RpfA are described in Table 1, below. In some embodiments, the Mtb antigen RpfA comprises the amino acid sequence according to SEQ ID NO: 10.

- 30 In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof.

- In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof.

- 35 In some embodiments, an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 407 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 positions 2 to 407 of SEQ ID NO: 10, or an immunogenic fragment of the amino acid sequence of positions 2 to 407 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 positions 2 to 407 of SEQ ID NO: 10.

- 40 In some embodiments, an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 407 of SEQ ID NO: 10.

- In some embodiments, an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, or an immunogenic fragment of the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48.
- In some embodiments, an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48.
- 10 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, or a fragment of the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24.
- 15 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 1221 of SEQ ID NO: 24.
- In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or a fragment of the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47.
- 20 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or a fragment of the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47.
- 25 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47.
- 30 RpfD
Characteristics of the Mtb antigen RpfD are described in Table 1, below. In some embodiments, the Mtb antigen RpfD comprises the amino acid sequence according to SEQ ID NO: 12.
- In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof.
- 35 In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof.
- In some embodiments, an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12, or an immunogenic fragment of the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12.
- 40

In some embodiments, an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 154 of SEQ ID NO: 12.

5 In some embodiments, an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, or an immunogenic fragment of the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46.

10 In some embodiments, an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46.

15 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, or a fragment of the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25.

20 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 462 of SEQ ID NO: 25.

25 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, or a fragment of the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45.

30 In some embodiments, an RNA sequence encoding an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45.

M72

35 Characteristics of the Mtb antigen M72 are described in Table 1, below. In some embodiments, the Mtb antigen M72 comprises the amino acid sequence according to SEQ ID NO: 29.

In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof.

In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof.

40 In some embodiments, an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 723 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 positions 2 to 723 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid

sequence of positions 2 to 723 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 positions 2 to 723 of SEQ ID NO: 29.

In some embodiments, an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 723 of SEQ ID NO: 29.

In some embodiments, an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, or an immunogenic fragment of the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52.

In some embodiments, an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, or a fragment of the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 2169 of SEQ ID NO: 28.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, or a fragment of the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51.

35

HbhA

Characteristics of the Mtb antigen HbhA are described in Table 1, below. In some embodiments, the Mtb antigen HbhA comprises the amino acid sequence according to SEQ ID NO: 18.

In some embodiments, an antigen useful for vaccination herein comprises an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

40

In some embodiments, an RNA useful for vaccination herein encodes an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

In some embodiments, an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to

the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18, or an immunogenic fragment of the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18.

In some embodiments, an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 2 to 199 of SEQ ID NO: 18.

In some embodiments, an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, or an immunogenic fragment of the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48.

In some embodiments, an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, or a fragment of the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 4 to 597 of SEQ ID NO: 30.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, or a fragment of the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47.

In some embodiments, an RNA sequence encoding an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47.

35

Exemplary combination of Mycobacterium tuberculosis (Mtb) antigens for vaccination

In some embodiments, antigens useful for vaccination herein comprising an amino acid sequence comprising an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof are expressed from RNA as fusion molecules. In some embodiments, such antigens are derived from different Mtb antigens and are expressed from RNA as fusion molecules. Thus, in some embodiments, an RNA molecule encodes at least two antigenic amino acid sequences as fusion molecule, wherein each antigenic amino acid sequence comprises an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic

- variant thereof. In some embodiments, the antigenic amino acid sequences are derived from different Mtb antigens. In some embodiments, each RNA molecule encodes at least two antigenic amino acid sequences as fusion molecule. In some embodiments, such fusion molecule comprises an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof.
- In some embodiments, such fusion molecule comprises an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof.
- In some embodiments, such fusion molecule comprises an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.
- In some embodiments, such fusion molecule comprises an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof.
- In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44.
- In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44.
- In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43.
- In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46.
- In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46.

In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48.

In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52.

In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51.

In some embodiments, amino acid sequences described herein comprise a secretory signal peptide. In some embodiments, the secretory signal peptide is fused to the N-terminus of an amino acid sequence described herein.

In some embodiments, a signal peptide (or signal sequence) is fused, either directly or through a linker, to an antigen sequence described herein. Accordingly, in some embodiments, a signal peptide is fused to an amino acid sequence comprising an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof described herein. In some embodiments, a signal peptide is functional in mammalian

- cells. In some embodiments, a signal peptide is heterologous to the amino acid sequence comprising an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof described herein. In some embodiments, a signal peptide has a length of about 15 to 30 amino acids. In some embodiments, a signal peptide is positioned at the N-terminus of an encoded polypeptide as described herein, without
5 being limited thereto. In some embodiments, a signal peptide allows the transport of the polypeptide with which it is associated into a defined cellular compartment, preferably the cell surface, the endoplasmic reticulum (ER) or the endosomal-lysosomal compartment. In some embodiments, a signal peptide is a secretory signal peptide. In some embodiments, a signal peptide comprises the sequence MRVMAPRTLILLSGALALTETWAGS, or a sequence having 1, 2, 3, 4, or at the most 5 amino acid differences relative thereto. In some embodiments, a signal peptide comprises the
10 sequence MGGAAARLGAVILFVVIVGLHGVRG, or a sequence having 1, 2, 3, 4, or at the most 5 amino acid differences relative thereto. In some embodiments, a signal peptide comprises the sequence MFIFLLFLTSTG, or a sequence having 1, 2, 3, 4, or at the most 5 amino acid differences relative thereto.
- In some embodiments, a secretory signal peptide comprises the amino acid sequence of positions 1 to 26 of SEQ ID NO: 44, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the
15 amino acid sequence of positions 1 to 26 of SEQ ID NO: 44, or a functional fragment of the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44.
- In some embodiments, a secretory signal peptide comprises the amino acid sequence of positions 1 to 26 of SEQ ID NO: 44.
- 20 In some embodiments, an RNA sequence encoding a secretory signal peptide comprises the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43.
- 25 In some embodiments, an RNA sequence encoding a secretory signal peptide comprises the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43.
- In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the
30 immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 44, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 44.
- In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the
35 immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 44.
- In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 1448 of SEQ ID NO: 43, or a
40 nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 1448 of SEQ ID NO: 43.
- In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or

the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 43, 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: 43.

5 In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 46, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 46.

10 In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 46.

15 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 872 of SEQ ID NO: 45, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 872 of SEQ ID NO: 45.

20 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 45, 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: 45.

25 In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 48, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 48.

30 In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 48.

35 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 1943 of SEQ ID NO: 47, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 1943 of SEQ ID NO: 47.

40 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 47, 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: 47.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 52, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 52.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 52.

In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 54 to 2591 of SEQ ID NO: 51, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 2591 of SEQ ID NO: 51.

In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 51, 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: 51.

In some embodiments, a secretory signal peptide comprises the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, or a functional fragment of the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63.

In some embodiments, a secretory signal peptide comprises the amino acid sequence of positions 1 to 25 of SEQ ID NO: 63.

In some embodiments, an RNA sequence encoding a secretory signal peptide comprises the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, or a fragment of the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62.

In some embodiments, an RNA sequence encoding a secretory signal peptide comprises the nucleotide sequence of positions 1 to 75 of SEQ ID NO: 62.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 63, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 63.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid

sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 63.

5 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 62, 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: 62.

10 In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 65, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 65.

15 In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 65.

20 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 64, 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: 64.

25 In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 67, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 67.

30 In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 67.

35 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 66, 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: 66.

40 In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 69, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 69.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 69.

- 5 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 68, 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: 68.

- 10 In some embodiments, a secretory signal peptide comprises the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, or a functional fragment of the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71.

15 In some embodiments, a secretory signal peptide comprises the amino acid sequence of positions 1 to 13 of SEQ ID NO: 71.

- In some embodiments, an RNA sequence encoding a secretory signal peptide comprises the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, or a fragment of the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70.

20 In some embodiments, an RNA sequence encoding a secretory signal peptide comprises the nucleotide sequence of positions 1 to 39 of SEQ ID NO: 70.

- 25 In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 71, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 71.

- 30 In some embodiments, a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 71.

- In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 70, 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: 70.

- 35 In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 73, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 73.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 73.

- 5 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 72, 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:

10 72.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 75, or an amino acid sequence having at least 99%,

- 15 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 75.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 75.

- 20 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 74, 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:

25 74.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 77, or an amino acid sequence having at least 99%,

- 30 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 77.

In some embodiments, a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of SEQ ID NO: 77.

- 35 In some embodiments, RNA encoding a fusion molecule comprising an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of SEQ ID NO: 76, 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:

40 76.

The antigens used in a TB mRNA vaccine designated herein as "RNA mix 3" were selected based on their performance in preclinical and clinical studies in the field of TB vaccine research (Table 1). The vaccine comprises eight *Mtb* antigens:

(1) Ag85A, (2) ESAT6, (3) VapB47, (4) Hrp1, (5) RpfA, (6) RpfD, (7) M72, and (8) HbhA. Selection of antigens 1 to 6 was based on published pre-clinical studies showing the immunogenic potential of these candidates and suggesting their use in a TB vaccine. Moreover, the eight antigens cover all phases (acute, latent, and resuscitation phases) of *Mtb* infection within the host. When combined in a polyprotein encoded within a cytomegalovirus (CMV) vector, these six antigens were shown to induce strong CD4⁺ and CD8⁺ T-cell responses and confer reduction of *Mtb* infection and disease by ~68% in a rhesus macaque challenge model (Hansen SG, et al. Nat Med, 2018; 24(2):130–143).

Antigen 7 (M72) is a fusion protein comprised of the *Mtb* antigens PPE18 (Mtb39a) and PepA (Mtb32a) in combination with the adjuvant AS01E (Baldrige J, et al. In: Hackett C.J., Harn D.A. (eds) Vaccine Adjuvants. Infectious Disease. Humana Press. p. 235–255; Skeiky YAW, et al. J Immunol, 2004; 172(12):7618–7628). Clinical data showed this subunit vaccine to provide 54% protection from pulmonary TB for at least 3 years, without safety concerns (Tait DR, et al. N Engl J Med, 2019; 381(25):2429–2439). HbhA was selected as an additional antigen for its localization on the *Mtb* surface, making it a suitable target for antibodies and, hence, prevention of dissemination of bacteria within the host. Further, HbhA was shown to induce potent cellular immunity in animals and in LTBI infected patients (Parra M, et al. Infect.Immun., 2004; 72(12):6799–6805; Aerts L, et al. J. Immunol., 2019; 202(2):421–427).

Table 1: Description of antigens included in the RNA mix 3 vaccine candidate

Antigen	Predicted function, localization in <i>Mtb</i> , phase association	Immunogenicity and protective activity
Ag85A (Rv3804c)	Lipid synthesis, secreted, active/acute phase	- Human: induction of Ag85A-specific humoral and cellular responses characterized by polyfunctional CD4⁺ (IFN-γ, TNFα, and IL-2-positive cells) and CD8⁺ T-cell responses when administered as boost immunization to healthy BCG vaccinated adults (Wilkie M, et al. Vaccine, 2020; 38(4):779–789). - Mouse: induction Ag85A-specific antibodies and <i>Mtb</i>-specific Th1 and CTL responses; Ag85A DNA vaccine mediate protection against <i>Mtb</i> and <i>Mycobacterium bovis</i> challenge (Huygen K, et al. Nat Med, 1996; 2(8):893–898).
ESAT-6 (Rv3875)	Virulence factor, secreted, active/acute phase	- Human: CAF01-adjuvanted Ag85B/ESAT-6 fusion protein (H1) vaccination induced long-lasting CD4⁺ T-cell response (van Dissel JT, et al. Vaccine, 2010; 28(20):3571–3581) - Mouse: induction of CD4⁺ and CD8⁺ T-cell responses to ESAT-6 in lung mucosa. Intranasal delivery of BCG as prime and MVA-AE as boost vaccination mediates protection against <i>Mtb</i> challenge (Khanna M, et al. Vaccine, 2021; 39(12):1780–1787). Deletion of ESAT-6 and CFP10 from MTBVAC (<i>Mtb</i> live-attenuated vaccine candidate) reduced its efficacy to BCG levels after <i>Mtb</i> intranasal challenge (Aguilo N, et al. Nat Commun, 2017; 8(1):16085).
M72 (Rv0125 and Rv1196)	Fusion protein consisting of antigens Mtb32A/Rv0125 (PepA) and Mtb39A/Rv1196 (PPE18; (Skeiky YAW, et al. J Immunol, 2004; 172(12):7618–7628)) PepA: Metabolism, secreted and membrane protein, active/acute phase PPE18: function unknown, membrane bound, detected in late stages of infection (Kruh NA, et al. PLOS ONE, 2010; 5(11): e13938)	- Human: induction of M72-specific antibodies and CD4⁺ T-cell responses sustained for 3 years; 49.7% (95% confidence interval [CI]: 2.1 to 74.2) vaccine efficacy in preventing progression to pulmonary TB for at least 3 years (Tait DR, et al. N Engl J Med, 2019; 381(25):2429–2439). - Mouse: induction of IFN-γ and antibody responses to Mtb32A and Mtb39A components. Vaccination mediated protection against <i>Mtb</i> aerosol challenge (Skeiky YAW, et al. J Immunol, 2004; 172(12):7618–7628).

VapB47 (Rv3407)	Possible antitoxin, localization unknown, dormancy/latency associated	Human: VapB47-specific IFN-γ responses observed in peripheral blood mononuclear cells from active TB patients (Slayden RA, et al. <i>Pathog Dis</i>, 2018; 76(4)). Mouse: induction of Rv3407-specific serum antibodies and IFN-γ secreting T cells; Rv3407-encoding DNA vaccine deliver as BCG boost mediates protection against <i>Mtb</i> (Mollenkopf HJ, et al. <i>Infect Immun</i>, 2004; 72(11):6471–6479).
Hrp1 (Rv2626)	Function unknown, secreted and membrane bound, dormancy/latency associated	Human: Hrp1-specific antibody detected in active TB patients (Davidow A, et al. <i>Infect Immun</i>, 2005; 73(10):6846–6851). Mouse: plasmid DNA vaccination induced humoral and Th1-type cellular response (Roupie V, et al. <i>Infect Immun</i>, 2007; 75(2):941–949). Hrp1 when administered with ESAT-6-Ag85B-MPT64_(190–198)-Mtb8.4 mediates protection against <i>Mtb</i> challenge up to 30 weeks post-vaccination (Liu X, et al. <i>Hum Vaccin Immunother</i>, 2016; 12(7):1670–1677).
RpfA (Rv0867c)	Function unknown, secreted, resuscitation associated and detected in late stages of infection (Kruh NA, et al. <i>PLOS ONE</i> , 2010; 5(11): e13938)	Human: induction of Rv0867c-specific polyfunctional memory CD4⁺ and CD8⁺ T-cell responses in long-term latently infected individuals that did not develop TB disease (Commandeur S, et al. <i>Clin Vaccine Immunol</i>, 2011; 18(4):676–683). IFN-γ production by T cells in response to RpfD and RpfA is higher in LTBI than in subjects with active pulmonary TB (Arroyo L, et al. <i>BMC Infect Dis</i>, 2018; 18(1):26). Mouse: Immunogenicity observed after <i>Mtb</i> infection and BCG vaccination. Recombinant RpfA protein induced cellular and humoral immune responses in C57BL/6 mice (Yeremeev V.V., et al. <i>Infect. Immun.</i>, 2003; 71(8):4789–4794). Moderate protection from TB was induced by immunization with a family member (RpfB; Romano M, et al. <i>Microbes Infect</i>, 2012; 14(1):86–95).
RpfD (Rv2389c)	Function unknown, localization unknown, resuscitation associated	Human: induction of Rv2389c-specific polyfunctional memory CD4⁺ and CD8⁺ T-cell responses in long-term latently infected individuals that did not develop TB disease (Commandeur S, et al. <i>Clin Vaccine Immunol</i>, 2011; 18(4):676–683). IFN-γ production by T cells in response to RpfD and RpfA is higher in LTBI than in subjects with active pulmonary TB (Arroyo L, et al. <i>BMC Infect Dis</i>, 2018; 18(1):26). Mouse: Immunogenicity observed after <i>Mtb</i> infection and BCG vaccination. Recombinant RpfD protein induced cellular and humoral immune responses in C57BL/6 mice (Yeremeev V.V., et al. <i>Infect. Immun.</i>, 2003; 71(8):4789–4794). Moderate protection from TB was induced by immunization with a family member (RpfB; Romano M, et al. <i>Microbes Infect</i>, 2012; 14(1):86–95).
HbhA (Rv0475)	Mediate adherence to epithelial cells to help extrapulmonary dissemination, membrane bound, expressed under high temperature and favorable conditions	Human: induction of HbhA-specific IFN-γ secreting T-cell responses that discriminate active TB from LTBI (Tang J, et al. <i>PLoS One</i>, 2021; 16(7):e0254571). Mouse: induction of HbhA-specific antibodies and IFN-γ secreting T cells. HbhA vaccine mediates protection against <i>Mtb</i> challenge (Parra M, et al. <i>Infect.Immun.</i>, 2004; 72(12):6799–6805).

Abbreviations: Ag85A = antigen 85A; CTL = cytotoxic T lymphocyte; ESAT-6 = 6 kilodalton early secretory antigenic target; M72 = antigen Mtb72F (recombinant protein derived from two *Mtb* proteins, Mtb32A and Mtb39A); HbhA = heparin-binding hemagglutinin; Hrp1 = hypoxic response protein 1; IFN = interferon; LTBI = latent tuberculosis infection; *Mtb* = *Mycobacterium tuberculosis*; RpfA = resuscitation-promoting factor A; RpfD =

resuscitation-promoting factor D; Th1 = type 1 helper T cell; TNF = tumor necrosis factor; VapB47 = virulence associated protein B47.

5 For optimal expression of the antigen encoding- mRNAs in eukaryotic cells, a codon optimization was performed, in which nucleotides were exchanged without alteration of the amino acid (aa) sequence of the synthesized protein (the native protein sequences of the *Mtb* antigens is listed in Table 2, the codon optimized sequences are listed in Table 3). Codon optimization was performed based on a protocol that favors the use of the most frequent eukaryotic codons for the respective amino acid without alteration of the guanine-cytosine (GC-) content of the overall nucleotide sequence. Antigen Ag85A contains an N-terminal signal peptide, predicted to be potentially functional in eukaryotic cells utilizing
10 the SignalP5.0 prediction tool (Almagro Armenteros JJ, et al. Nat. Biotechnol., 2019; 37(4):420–423). The predicted cleavage site is after aa 41 therefore, aa 1–41 of Ag85A were deleted from the sequence included in preclinical evaluation as well as in the TB vaccine candidate of the invention. Alteration of the N-terminus of Ag85A by depletion of one more amino acid (aa 1–42) has been shown to maintain immunogenicity when included in an antigen cassette (Hansen SG, et al. Nat Med, 2018; 24(2):130–143).

15

[illegible]

RpfD	Resuscitation- promoting factor D
Rv2389c;	PepA Rv0125, Mtb32a; PepA
M72 ^a	
Rv1196, Mtb39a; PPE18	

[illegible]

^a M72 is a fusion protein consisting of Mtb32a and Mtb39a

Codon optimized (opt 10) nucleotide sequence

[illegible]

[illegible]

HbhA

TGVIDPNGLVLTNNHVTAGATDINAFSVSGQTYGVVDWVG
DRTPQDAVLQLRGAGGLPSAIGGGVAVGEVPVAMGNSSGG
QGGTPRAVPGRWALGQTVQASDSLTAEEETLNLGIQTDAAI
QPDSGGPVMNGLGQVVGMNTAA (SEQ ID NO: 29)

MAENSNIDDIKAPILAAALGAADIALATVNWELTNLRERAETTR
TDTRSVEESRRLTKLQEDLPQLTELREKTAEELRKAAEG
YLEAATSRYNELVERGEALERLSQSOFEEVSARAEGYVDQ
AVELTQEALGTVASQTRAVGERAAKLVGIELPKKAAPAANKAP
AKKAAPAANKAAKKAPAKKAAKKVTKQ (SEQ ID NO: 18)

M72 is a fusion protein consisting of *Mycobacterium tuberculosis* antigens Mtb32a and Mtb39a.

Every antigen sequence designed was further modified by N-terminal addition of a human major histocompatibility complex (MHC) class I signal peptide fragment (sec) (Kreiter S, et al. J Immunol, 2008; 180(5):309–318). This facilitates guidance of the *de novo* synthesized protein into the lumen of the endoplasmic reticulum in turn resulting in export of the protein via the Golgi network to the extracellular space. There it can induce humoral immune responses or be taken up by antigen-presenting cells, inducing T helper immune cell responses.

- Initial preclinical evaluation of the developed Mtb antigen constructs (detailed in Examples) led to the generation of three different RNA mixes, each comprising 4 modRNAs. Each RNA mix expresses 4 to 8 *Mtb* antigens (Figure 1) and is manufactured by an *in vitro* transcription (IVT) reaction. RNA mix 1 comprises the antigens Ag85A(Δ 1-41), M72, ESAT6, and HbhA as four single modRNA-encoded antigens. Ag85A(Δ 1-41), M72, and ESAT6 were selected as immunodominant antigens, whereas HbhA induced antibody responses beneficial to prevent *Mtb* infection. RNA mix 2 comprises Ag85A(Δ 1-41) and M72 as two single modRNA-encoded antigens together with two modRNAs each encoding a fusion of two *Mtb* antigens (ESAT6-Hrp1 and RpfA-HbhA). The two fusion antigen constructs were added to Ag85A(Δ 1-41) and M72 to target *Mtb* in different phases of infection. RNA mix 3 consists of four RNAs, each encoding a fusion protein (Ag85A(Δ 1-41)-Hrp1, ESAT6-RpfD, RpfA-HbhA and M72-vapB47). With RNA mix 3 it was aimed to have the broadest protection against TB, since all phases of infection are targeted with antigens that are known to be immunogenic and protective in at least an animal model. Six out of eight antigens were shown to confer protection in a non-human primate model (Hansen SG, et al. Nat Med, 2018; 24(2):130–143), M72 has been shown to be protective in humans (Tait DR, et al. N Engl J Med, 2019; 381(25):2429–2439) and HbhA was added for generating protective antibody and T-cell responses in mice (Parra M, et al. Infect.Immun., 2004; 72(12):6799–6805) and for eliciting T cell immune response in humans with LTBI (Tang J, et al.; 7). Immunogenicity testing of the three mixes *in vivo* (described in Example 5) demonstrated that immunization with RNA mix 3 confers broad protection against infection and recurrence of active TB. The tables below list the nucleotide and amino acid sequences of the mRNA constructs included in RNA mix 1 (Table 4), RNA mix 2 (Table 5), and RNA mix 3 (Table 6).

110

AGATAaactagtaattcttctgtccccaagactgagagaaacccgccacatgagadgtagtggcccccagaacacctgatactctgctgctgtgtggccgacctgtgacctgacagagacatggcggaagcACACCATTACCATCACAGCGCCTCTTGACAACTTCCAGCTGCTCAATCCCAAGTGCACATTGGACCACAGCTTCTGGGACTGGGAGTGGTGGACAACAAATGGAAATGGAGCCAGAGT

MRVYMAPRTJLLSGALALTETWAGSHH
HHHHHTAASDNFQLSQGQGFAPIGQA
MALAQIRSGSGSPTVHGPTAFILGLW
DNINGANGVRQVRVGSAPALNGHSTGD
TVDVADGAPNSATAMADANLSHGPGDV
TSVTVWTKSGGTRTGNVTLAEGPPAEFM
VDFGALPPENSARMYAGPSASLVAAA
QMWDSVASDLSAASAFQSVWVWGLTVG
SWIGSSAGLWMAAASPYAVWMSVTAQG
AEITAAQVRVAAAAYTAYGLTVPPVIA
ENRAELMILATNLGQNTPAJAVNEAEY
GEMWQAQDAAMFYAAATATATATLLP
FEAEAPMTSAGGLELQAAAEEASDTAA
QMLNNVQALQLOQTOQTGTPSSK
GGLWKTVPSPRSPISNVSMANNHMS
MTNNGSVMTNLTSSMLKGFAPAARQA
VTQAQNGVRANSSLSGSLSGGLGG
VAANLGRAASVGSLSVQAWAANQAV
TTPAARALPLTSLTSAERGRQMLGLLP
VGQMGARAGGGLSVLRVPPRPMYMPH
SPQAGDIAPPALSQDRFADPLDPASA
VMAAQGPQWANTIKLGYNNAVAGTG
VIDPNGVLTNNHVATGATDINAFVGS
GGQTYGVWGVDRYTDQDVAVLQLRGAGG
PSAIGGGSVAVGEPVWAMGNSGGQGG
TRPVRPGRVVALQTVQASDNLGAEET
NGLITFQADATPGDSGPGVNLGQWV
GMNTAAS (SEQ ID NO: 38)

[illegible]

sec-M72 [2726]

MRVYMAPRTLILLISGALALITETWAGSTT
ARDIMNAGVTCVGEHETILTAQAQYMR
HDIGALPGTGGDDRLHGMLTRDRIKIG
LAAGLDPNTATAGELARSDIYVDANASI
QEMLVNMEEHQVRRPVPVISEHRLVGIVT
EADJIRHLPEHATVIFYKATCSPNALAST
EEQQWNFAGIEAASAIAQGNVTSIHS
LLDDEGKQSLTKLAAWGGSGSEAY
QGVQKQWDATATELNNALQNLRAT
ISEAGQAMASTEGNVTGMFA (SEQ
ID NO: 40)

sec-Hrp1-ESAT6 [1250]

[illegible]

sec-RpfD-HbhA [1595]

[illegible]

Table 6: Nucleotide and amino acid sequences of fusion-antigens modRNA constructs as included in RNA mix 3. AGA: T7-translation start; low case letters in italic: 3'- and 5'-untranslated regions, and poly-adenosine tail; underlined: sec signal peptide fragment; capital letters: first antigen, capital letters bold: second antigen.

[illegible]

aaaaaaaaaaaaaaaa (SEQ ID NO: 45)

(SEO ID NO: 47)

sec-M72-vapB47 [3020]

MRVYMAPRTILLISGALALTEIWAGSSG
MAAQATAATDGEWDQVARGESGGNWS
TNNVSVAKIAFTGAVLGGGGIA
TNGVGLGQJFTQSTWAAHGGGEFA
SQAQSLAQREQJGAVERVLTQGRGAW
VPCVGRGLSNATPREVLPSAAMDAPLDA
AAVANGEPAPLAPPPADPAPVLEANDLP
LEPLPAAPDAPLAPPELADAPPAPADVA
PPVLEAVYNDLPAPLGEPLADPADPAPPA
LDLAPLAPADLAPPAIDLAPPAIDLAPP
VELAVYNDLPAPLGEPLPAAPAEAPPADL
APASADLAPPADLADPPAPAPAPAPPA
DLAPPAAVNEQTAPGQDPATAPGPPVGL
ATDLELPDQPADAPPPGDVTEAPAE
TPQVSNIAVTKLWOAIRAQDVCNDAL
DLSLAQPVVIGAENSNDIDIKAPLLAAL
GAADLALATVNLITLNLRERAEETR
DTSRVEESRARLTKLQEDLEPQLT
DTRLEKFTAEELRKAEGYEAATSRY
NNELVERGEALERLSQQSFEFVSA
RAEGYVDQAVELTQEALGTVASQTR
AVGERRAAKLVGIELPKKAAPAKKAA
PAPKKAAPAKKAAKAPAKKAAK
KVTK (SEQ ID NO: 48)

MRVVMAPRTLLLLSGALALTETWAGSHH
HHHHHHHTAASDNFQLSQGGQFAIPTQA
MALAGQJRSGGSPVTHGPTAFLGLGW
DNNNGARQVRVWGSPAASLGISTGD
VTVTVADGAPNTAMADALNGHPGDV
VTVSWQTKSGTTRTGNVTLAEGPPEFM
VDFGLPPEINARMYAGPSASLVA
QMWDSVASDLFSAASFOVWVGLTVG

117

118

The table below lists nucleotide and amino acid sequences of mRNA constructs similar to those included in RNA mix 3 (Table 6) but having different signal peptides (MGGAAARLGAVILFVIVGLHGVRG and MFJLLFLTLTSG, respectively).

Table 7: Nucleotide and amino acid sequences of fusion-antigens RNA constructs similar to those included in RNA mix 3. bold: signal peptide fragment.	
Construct	nt sequence
pST4-T7-AGA-dEarI-hAg-SP(HSV-1 gD)-hsopt10_Ag85A-hsopt10_Hrp1-FI-A30LA70	atgggggggggctcgccagggtggggcggtgattttgttgcgtcatagtgggcc tccatgggggtccgcggcgagccttcagcagacctgagctgctggaatactacacgagtg ccagccccagcatgggcagagacatcaaaagtcagttccagctctggaggagccaactctctgccc tgtacctgtgtagtgccctgagagcccgagatgatttctctgctgggacatcaacacccctgctt gaaatggtatgacgtctggactctgtgtagtgctgtggaggagacagcagcttactctg attgtaccagcctgctgtgaaagccgagtcagaccctacacacaaatggaaaccttctgacct ctgaactgctgtgctgctgcagccacagacatgtgaaacccacacaggtctgtgtgtggac tgagcatggctgttcttctgtgacctggccatctaccacctcagcagttgtgtatgctggag ccatgtctgagctgtggaccccgagccatgggacccacctgattgactggccatgggag atgccggaggatatacaagcctctctgacatgtgggacccaagaaagacccagcctgtgcagagaaa tgacctctgtgaaatgtggcacaactgtgccaacacacagagtggtgtgactgtggaat ggaaacctctgtatctggggaacaaacctgctgccaagtcttctggaaaggctttgtgagacc agcaacatcaaatctcagtgatgcctacaatgcggaggaggacacaaatggagtggttgcattctctg actctggaacccacttgggaatactggggagcccgactgaatgccatgaacctgtatctgcaga gagccctgggagccacacacacagagcctgcccctcagggagccacacacagccagagacat catgaatgccggtgacatgtggcgaacatgaacacactgacagctgtctcagtcacatgag agaacatgacattggagccctgcccattgtggagacgacgacactggtgcatgtgtgacag acagagacatctgtatcaaggcctggccgcccgtggaccccaacacacagccagccgagga actgcccagagacagcatctactatgtgatgccaatgcccagcatccaggaatgtgaatgtgat ggaaagaacacacaggtgagggggtgctgtgactctgaacacagactgtgtggcattgtgacag aagccgacattgcagacacctgctgaaacatgcccattgtgagttgtgaaagccattgtgagccc catgcccctggccagc (SEQ ID NO: 62)
	MGGAAARLGAVILFVIVGLHGVRG GAFSRPGLPVEYLQVPSPSM GRDIKVFQSSGGANSPALYLLDGLRAQDDDFSGWDINTPAFEWYDQSG LSWMPVGGQSSFYSDWYQACGKAGCQTYKWETFLTSELPGWLQA NRHVKPTGSVVGLSMAASSALTATYHPQQFVYAGAMSGLLDPSQAM GPTLIGLAMGDAGGYKASDMWGPKEPAWQRNDPLLNVGKLIANNIT RVWVYCGNGKPSDLGGNNLPAKFLGFRVTSNIKFDQDAYNAGGGHNG VFDFPDSGTHSWEYWGALNNAKPDQLQRALGATPNTGPAPQGATTA RDIMNAGVTCVEHETLTAQAQYMRHEDIGALPICGDDDDRLHGMILT RDIVIKGLAAGLDPNATATAGELARDSIYYVDANASIQEMLNMVEEHQV RRVPVISEHRLVIGIVTEADJARHLPEHATVQFVKAICSPMALAS (SEQ ID NO: 63)

120

121

122

123

	agtgggaagccctgcccgcagccctcagcagcagattgagggtgcccgaacaacatcatgaag accagggaccagaggagctggccaaagtcagcagctgcagccagggagagatgctctctctgggaa gcctgaccacatctgaccccttctgctgtgagacagagagatgctctctggaagcagagatgat (SEQ ID NO: 72)	
pST4-T7-AGA-dEarI-hAg- SP(SARS-COV-1 S)- hsopt10_RpfA-hsopt10_hbhA-Fl- A30LA70	atgtttattttcttaattttcttactctcactagtggttcttggaagacacagaaaaaccaccac cagcaatgtgtctgtggccaaattgctctcacagagcgtgtctggagagagggaattgccaat ggccgccagggccacagcgcgcacagatggagaatgggatcaggtggccagatgtgaatctgga ggaactgtggagcaitcaacacacaggaatgatactgtggagactgcagttccaccagagacact ggccgcccatgtggagagaggaattgccccttctccacgtgcccagcagagagaacagcagatt ggcgtgggaagaagtgctggccaccagggcagagcgccctggtgctgtgtgtgtggcagaggcc tgagcaatggccaccacagagaagtgctgctgtctgcccgaatggaatgcccctctgtgagtcgcg cgccgtgaatggagaacctgcccctctgtggccctctctctgcgcatctgcccctctgtggaaactgg ccgccaatgatctgcccctctgtggagaacctctgcccctgcccctgcccctgcccctctgcccctct ggcagctggcccctctgcccctgcccctgcccctctgtggaactgtggaactgtggaatgtctgccc tgcccctctgggaacactctgcccctgcccctgcccctctgcccctctgcccctctgcccctctcc tgcccctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgccc tctgtggaactggccgtgaatgtatctgcccctctgtgggaacactctgcccctgcccctgccc aactggcccctctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgccc cccctctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgcccctgccc gaacagacagcccctggaatcagctgcccagcccctggaagacactgtgggactgtggccacag atctggaactgctgaacctgatctcagcctgtgtgtctctctctctgagagatgtgacagaagct ctgtgtaaacacactcaggtgagcaacatgtcctacacaagaagctgtgaggccatcagagc ccaggatgtgtgtggaatgtgcccctggaatctgtgcccagccttattgtgtgtgagctgaaac agcaacatcagacatcaaggcccctgtgcccctgcccctgcccctgcccctgcccctgcccctg cacagtgaatgagctgatacacaatctgagagagagagctgagagacagaagaacagacacaag aagcagagtgaggagagcagagcagactgacaaactgcaggagaagacccctgtgaaacagct gacagaactgagagaaaaattcacagccgaagaactgagaaaggccgcccgaaggctactctgga agccgccacacagacatacaatgaactgtgtggaagaggagaagctgctctgtgaaagactgaga	MFIFLLFLTLTSGSGRHRKPTTSNVSVAKIAFTGAVLGGGIAMAAQ ATAATDGEWDQVARCESGGNWSINTNGYLGGLQFTQSTWAAHGG GEFAPSAQLASREQQIAVGERVLATQGRGAWPVCGRGLSNATPREVL ASAAMDAPLDAAVNGEPAPLAPPADPAPPVELAANDLPAPLGEPLPA APADPAPPADLAPPADVAPPVELAVNDLPAPLGEPLPAAPADPAPPA DLAPPADLAPPADLAPPADLAPPADLAPPVELAVNDLPAPLGEPLPAAPA ELAPPADLAPASADLAPPADLAPPADLAPPADLAPPADLAPPAAVNEQT APGDQPATAPGGPVGLATDLELPEPDQPADAPPPGDVTEAPAEPTQV SNIAYTKLWQAIRAQDVCGNDALDSLAQQPVYVJGAENSNIDDIKAPLLA ALGAADLALATVNLITNLREAAETRTDTSRVEESRRLTKLQEDLP EQLTELREKFTAEELRKAEGYLEAATSRYNELVERGEAALERLRSQQS FEEVSARAEGVVDQAVELTQEALGTVASQTRAVGERAAKLVGIELPKKA APAKKAAPAKKAAPAKKAAPAKKAAPAKKAAPAKKAAPAKKAAPAKKA 75)

[illegible]

126

RNA delivery

RNA described herein may be delivered for therapeutic applications described herein using any appropriate methods known in the art, including, e.g., delivery as naked RNA, or delivery mediated by delivery vehicles.

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

- 5 In some embodiments, after administration of the RNA (in particular, mRNA) compositions/formulations 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 (in particular, mRNA) is translated by the target cell to produce the encoded peptide or polypeptide. In some embodiments, the target cell is a muscle cell. In some embodiments, the target cell is a cell in the liver. In some embodiments, the target cell is a cell in the lung. In some embodiments, the disclosure involves targeting the lymphatic system, in particular secondary lymphoid organs, more specifically spleen. In some embodiments, the target cell is a cell in the lymph nodes. In some embodiments, the target cell is a spleen cell. In some embodiments, the target cell is an antigen presenting cell such as a professional antigen presenting cell in the spleen. In some embodiments, the target cell is a dendritic cell in the spleen. Thus, RNA (in particular, mRNA) compositions/formulations described
- 10 herein may be used for delivering RNA to such target cell. The "lymphatic system" is part of the circulatory system and an important part of the immune system, comprising a network of lymphatic vessels that carry lymph. The lymphatic system consists of lymphatic organs, a conducting network of lymphatic vessels, and the circulating lymph. The primary or central lymphoid organs generate lymphocytes from immature progenitor cells. The thymus and the bone marrow constitute the primary lymphoid organs. Secondary or peripheral lymphoid organs, which include lymph nodes and the spleen, maintain mature naive lymphocytes and initiate an adaptive immune response. Lipid-based RNA delivery systems have an inherent preference to the liver, where, depending on the composition of the RNA delivery systems used, RNA expression in the liver can be obtained. 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 for RNA expression is liver and the target tissue is liver tissue.
- 20 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.

Delivery vehicles

- 30 To overcome the barriers to safe and effective RNA delivery, RNA may be administered with one or more delivery vehicles that protect the RNA from degradation, maximize delivery to on-target cells and minimize exposure to off-target cells. Such RNA delivery vehicles may complex or encapsulate RNA and include a range of materials, including polymers and lipids. In some embodiments, such RNA delivery vehicles may form particles with RNA.
- 35 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.
- 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
- 40

vehicles, nanoparticle encapsulation of RNA physically protects RNA from degradation and, depending on the specific chemistry, can aid in cellular uptake and endosomal escape.

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

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

RNA particles described herein include lipid nanoparticle (LNP)-based and lipoplex (LPX)-based formulations.

A lipoplex (LPX) described herein 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, 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 optionally additional lipids, such as DOPE or DSPC. In some embodiments, an RNA (especially mRNA) lipoplex particle is a nanoparticle.

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/cationically ionizable lipid and helper lipids such as phospholipids, cholesterol, and/or polymer-conjugated lipids (e.g., polyethylene glycol (PEG) lipids). In some embodiments, in the RNA LNPs described herein the RNA (in particular, mRNA) is bound by cationically ionizable lipid that occupies the central core of the LNP. In some embodiments, polymer-conjugated lipid forms the surface of the LNP, along with phospholipids. In some embodiments, the surface comprises a bilayer. In some embodiments, cholesterol and cationically ionizable lipid in charged and uncharged forms can be distributed throughout the LNP.

In some embodiments, RNA (e.g., mRNA) described herein 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)).

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 about 1900 nm (e.g., at most about 1800 nm, at most about 1700 nm, at most about 1600 nm, at most about 1500 nm, at most about 1400 nm, at most about 1300 nm, at most about 1200 nm, at most about 1100 nm, at most about 1000 nm, at most about 950 nm, at most about 900 nm, at most about 850 nm, at most about 800 nm, at most about 750 nm, at most about 700 nm, at most about 650 nm, at most about 600 nm, at most about 550 nm, or at most about 500 nm), such as in the range of about 20 to about 1500 nm, such as about 30 to about 1200 nm, about 40 to about 1100 nm, about 50 to about 1000 nm, about 60 to about 900 nm, about 70 to about 800 nm, about 80 to about 700 nm, about 90 to about 600 nm, or about 50 to about 500 nm or about 100 to about 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, 70 to 150 nm, or 80 to 150 nm. In some embodiments, the particles (e.g., LNPs and LPXs) described herein have a size (such as a diameter) in the range of from about 40 nm to about 200 nm, such as from about 50 nm to about 180 nm, from about 60 nm to about 160 nm, from about 80 nm to about 150 nm or from about 80 nm to about 120 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, from about 200 nm to about 250 nm, or from about 80 to about 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 40 nm to about 200 nm, such as from about 50 nm to about 180 nm, from about 60 nm to about 160 nm, from about 80 nm to about 150 nm or from about 80 nm to about 120 nm.

10 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 (especially mRNA particles) described herein may exhibit a polydispersity index (PDI) less than about 15 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 20 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 25 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 30 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 35 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 40 involves formation of a water-in-oil emulsion between an aqueous phase and an organic phase containing lipids. A

brief sonication of this mixture is required for system homogenization. The removal of the organic phase under reduced pressure yields a milky gel that turns subsequently into a liposomal suspension.

The term "ethanol injection technique" refers to a process, in which an ethanol solution comprising lipids is rapidly injected into an aqueous solution through a needle. This action disperses the lipids throughout the solution and promotes lipid structure formation, for example lipid vesicle formation such as liposome formation. Generally, the RNA (especially 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 (especially mRNA) lipoplex particles described herein are obtainable without a step of extrusion.

The term "extruding" or "extrusion" refers to the creation of particles having a fixed, cross-sectional profile. In particular, it refers to the downsizing of a particle, whereby the particle is forced through filters with defined pores. Other methods having organic solvent free characteristics may also be used according to the present disclosure for preparing a colloid.

In some embodiments, LNPs comprise four components: cationically ionizable 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 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 PEG lipids, or do not comprise any polymer-conjugated 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 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 diluting and filtrating, such as tangential flow filtrating, can follow.

The present disclosure describes compositions comprising RNA (especially mRNA) and at least one cationic or cationically ionizable lipid which associates with the RNA to form RNA particles and formulations comprising such particles. The RNA particles may comprise RNA which is complexed in different forms by non-covalent interactions to the particle. The particles described herein are not viral particles, in particular infectious viral particles, *i.e.*, they are not able to virally infect cells.

Suitable cationic or cationically ionizable lipids are those that form 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 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 elements such as molecular architecture, capping, coding regions or other features,

- 5 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 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
- 10 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. Mixed particulate formulations according to the present disclosure are obtainable by forming, separately, individual particulate formulations, followed by a step of mixing of the individual particulate
- 15 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. 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
- 20 agent, thereby allowing the formation of particles. As opposed to a mixed particulate formulation, a combined particulate formulation will typically comprise particles which comprise more than one RNA species. In a combined particulate composition different RNA species are typically present together in a single particle.

Polymers

- 25 Given their high degree of chemical flexibility, polymers are commonly used materials for nanoparticle-based delivery. Typically, cationic polymers are used to electrostatically condense the negatively charged 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 occurring polymers
- 30 such as chitosan have all been applied to nucleic acid delivery and are suitable as cationic polymers herein. In addition, some investigators have synthesized polymers specifically for nucleic acid delivery. Poly(β -amino esters), in particular, have gained widespread use in nucleic acid delivery owing to their ease of synthesis and biodegradability. Such synthetic polymers are also suitable as cationic polymers herein.

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

- If more than one type of repeat unit is present within the polymer, then the polymer is said to be a "copolymer."
- 40 It is to be understood that the polymer being employed herein can be a copolymer. The repeat units forming the copolymer can be arranged in any fashion. For example, the repeat units can be arranged in a random order, in an

alternating order, or as a "block" copolymer, i.e., comprising one or more regions each comprising a first repeat unit (e.g., a first block), and one or more regions each comprising a second repeat unit (e.g., a second block), etc. Block copolymers can have two (a diblock copolymer), three (a triblock copolymer), or more numbers of distinct blocks.

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

In certain embodiments, polymer may be protamine or polyalkyleneimine.

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

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

- In some embodiments, 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).

- 25 Cationic polymers (including polycationic polymers) contemplated for use herein include any cationic polymers which are able to electrostatically bind nucleic acid. In some embodiments, 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.

30

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
- 35
- 40

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 monovalent radical of a hydrocarbon is referred to as hydrocarbonyl herein. The hydrophobic group can have functional groups (e.g., ether, ester, halide, etc.) and atoms other than carbon and hydrogen as long as the group satisfies the condition of being substantially immiscible or insoluble in aqueous solution.

The term "hydrocarbon" includes non-cyclic, e.g., linear (straight) or branched, hydrocarbonyl groups, such as alkyl, alkenyl, or alkynyl as defined herein. It should be appreciated that one or more of the hydrogen atoms in alkyl, alkenyl, or alkynyl may be substituted with other atoms, e.g., halogen, oxygen or sulfur. Unless stated otherwise, hydrocarbon groups can also include a cyclic (alkyl, alkenyl or alkynyl) group or an aryl group, provided that the overall polarity of the hydrocarbon remains relatively nonpolar.

The term "alkyl" refers to a saturated linear or branched monovalent hydrocarbon moiety which may have one to thirty, typically one to twenty, often six to eighteen carbon atoms. Exemplary nonpolar alkyl groups include, but are not limited to, methyl, ethyl, propyl, isopropyl, hexyl, decyl, dodecyl, tetradecyl, hexadecyl, octadecyl, and the like.

The term "alkenyl" refers to a linear or branched monovalent hydrocarbon moiety having at least one carbon-carbon double bond in which the total carbon atoms may be six to thirty, typically six to twenty often six to eighteen. Generally, the maximal number of carbon-carbon double bonds in the alkenyl group can be equal to the integer which is calculated by dividing the number of carbon atoms in the alkenyl group by 2 and, if the number of carbon atoms in the alkenyl group is uneven, rounding the result of the division down to the next integer. For example, for an alkenyl group having 9 carbon atoms, the maximum number of carbon-carbon double bonds is 4. Preferably, the alkenyl group has 1 to 6 (such as 1 to 4), i.e., 1, 2, 3, 4, 5, or 6, carbon-carbon double bonds.

The term "alkynyl" refers to a linear or branched monovalent hydrocarbon moiety having at least one carbon-carbon triple bond in which the total carbon atoms may be six to thirty, typically six to twenty, often six to eighteen. Alkynyl groups can optionally have one or more carbon-carbon double bonds. Generally, the maximal number of carbon-carbon triple bonds in the alkynyl group can be equal to the integer which is calculated by dividing the number of carbon atoms in the alkynyl group by 2 and, if the number of carbon atoms in the alkynyl group is uneven, rounding the result of the division down to the next integer. For example, for an alkynyl group having 9 carbon atoms, the maximum number of carbon-carbon triple bonds is 4. Preferably, the alkynyl group has 1 to 6 (such as 1 to 4), i.e., 1, 2, 3, 4, 5, or 6, more preferably 1 or 2 carbon-carbon triple bonds.

The term "alkylene" refers to a saturated linear or branched divalent hydrocarbon moiety which may have one to thirty, typically two to twenty, often four to twelve carbon atoms. Exemplary nonpolar alkylene groups include, but are not limited to, methylene, ethylene, trimethylene, hexamethylene, decamethylene, dodecamethylene, tetradecamethylene, hexadecamethylene, octadecmethylenes, and the like.

The term "alkenylene" refers to a linear or branched divalent hydrocarbon moiety having at least one carbon-carbon double bond in which the total carbon atoms may be two to thirty, typically two to twenty, often four to twelve. Generally, the maximal number of carbon-carbon double bonds in the alkenylene group can be equal to the integer which is calculated by dividing the number of carbon atoms in the alkenylene group by 2 and, if the number of carbon atoms in the alkenylene group is uneven, rounding the result of the division down to the next integer. For

example, for an alkenylene group having 9 carbon atoms, the maximum number of carbon-carbon double bonds is 4. Preferably, the alkenylene group has 1 to 6 (such as 1 to 4), i.e., 1, 2, 3, 4, 5, or 6, carbon-carbon double bonds.

The term "cycloalkyl" represents cyclic non-aromatic versions of "alkyl" and "alkenyl" with preferably 3 to 14 carbon atoms, such as 3 to 12 or 3 to 10 carbon atoms, i.e., 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 carbon atoms (such as 3, 4, 5, 6, 7, 8, 9, or 10 carbon atoms), more preferably 3 to 7 carbon atoms. Exemplary cycloalkyl groups include cyclopropyl, cyclopropenyl, cyclobutyl, cyclobutenyl, cyclopentyl, cyclopentenyl, cyclohexyl, cyclohexenyl, cycloheptyl, cycloheptenyl, cyclooctyl, cyclooctenyl, cyclononyl, cyclononenyl, cyclodecyl, cyclodecenyl, and adamantyl. The cycloalkyl group may consist of one ring (monocyclic), two rings (bicyclic), or more than two rings (polycyclic).

The term "aryl" refers to a monoradical of an aromatic cyclic hydrocarbon. Preferably, the aryl group contains 3 to 14 (e.g., 5, 6, 7, 8, 9, or 10, such as 5, 6, or 10) carbon atoms which can be arranged in one ring (e.g., phenyl) or two or more condensed rings (e.g., naphthyl). Exemplary aryl groups include cyclopropenyl, cyclopentadienyl, phenyl, indenyl, naphthyl, azulenyl, fluorenyl, anthryl, and phenanthryl. Preferably, "aryl" refers to a monocyclic ring containing 6 carbon atoms or an aromatic bicyclic ring system containing 10 carbon atoms. Preferred examples are phenyl and naphthyl. Aryl does not encompass fullerenes.

The term "aromatic" as used in the context of hydrocarbons means that the whole molecule has to be aromatic. For example, if a monocyclic aryl is hydrogenated (either partially or completely) the resulting hydrogenated cyclic structure is classified as cycloalkyl for the purposes of the present disclosure. Likewise, if a bi- or polycyclic aryl (such as naphthyl) is hydrogenated the resulting hydrogenated bi- or polycyclic structure (such as 1,2-dihydronaphthyl) is classified as cycloalkyl for the purposes of the present disclosure (even if one ring, such as in 1,2-dihydronaphthyl, is still aromatic).

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

The term "lipid-like material", "lipid-like compound" or "lipid-like molecule" relates to substances, in particular amphiphilic substances, that structurally and/or functionally relate to lipids but may not be considered as lipids in a strict sense. For example, the term includes compounds that are able to form amphiphilic layers as they are present in vesicles, multilamellar/unilamellar liposomes, or membranes in an aqueous environment and includes surfactants, or synthesized compounds with both hydrophilic and hydrophobic moieties. Generally speaking, the term 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 comprising two long alkyl chains and a polar head group are generally cylindrical. The area occupied by the two alkyl chains is similar to the area occupied by the polar head group. Such lipids have low solubility as monomers and tend to aggregate into planar bilayers that are water insoluble. Traditional surfactant monomers comprising only one linear alkyl chain and a hydrophilic head group are generally cone shaped. The hydrophilic head group

tends to occupy more molecular space than the linear alkyl chain. 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 *cis* 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, GPCo or lecithin), phosphatidylethanolamine (PE or GPEtn) and phosphatidylserine (PS or GPSer).

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

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

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

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

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

Cationic/Cationically ionizable lipids

In some embodiments, the RNA compositions and formulations and 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 able to electrostatically bind nucleic acid. In some embodiments, cationic or cationically ionizable lipids contemplated for use herein can be associated with nucleic acid, *e.g.* by forming complexes with the nucleic acid or forming vesicles in which the nucleic acid is enclosed or encapsulated.

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

In some embodiments, a cationic lipid has a net positive charge only at certain pH, in particular acidic pH, while it has preferably no net positive charge, preferably has no charge, *i.e.*, it is neutral, at a different, preferably higher pH such as physiological pH. This ionizable behavior is thought to enhance efficacy through helping with endosomal escape and reducing toxicity as compared with particles that remain cationic at physiological pH.

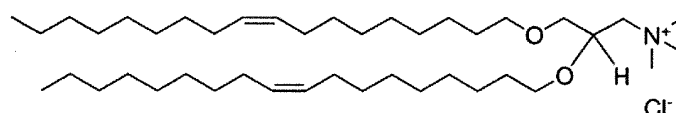
As used herein, a "cationically ionizable lipid" refers to a lipid or lipid-like material which has a net positive charge or is neutral, *i.e.*, which is not permanently cationic. Thus, depending on the pH of the composition in which the cationically ionizable lipid is solved, the cationically ionizable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

In some embodiments, the cationic or cationically ionizable lipid comprises a head group which includes at least one nitrogen atom (N) which is positive charged or capable of being protonated, *e.g.*, under physiological conditions.

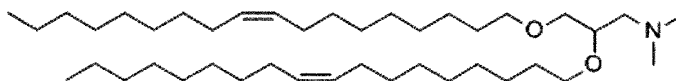
Examples of cationic or cationically ionizable lipids include, but are not limited to N,N-dimethyl-2,3-dioleoyloxypropylamine (DODMA), 1,2-dioleoyl-3-trimethylammonium propane (DOTAP); 1,2-di-O-octadecenyl-3-trimethylammonium propane (DOTMA), 3-(N—(N',N'-dimethylaminoethane)-carbamoyl)cholesterol (DC-Chol), dimethyldioctadecylammonium (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 ammonium bromide (DORIE), and 2,3-dioleoyloxy-N-[2(spermine carboxamide)ethyl]-N,N-dimethyl-1-propanaminium trifluoroacetate (DOSPA), 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLinDMA), 1,2-dilinolenyloxy-N,N-dimethylaminopropane (DLenDMA), dioctadecylamidoglycyl spermine (DOGS), 3-dimethylamino-2-(cholest-5-en-3-beta-oxybutan-4-oxy)-1-(cis,cis-9,12-oc-tadecadienoxy)propane (CLinDMA), 2-[5'-(cholest-5-en-3-beta-oxy)-3'-oxapentoxy]-3-dimethyl-1-(cis,cis-9',12'-octadecadienoxy)propane (CpLinDMA), N,N-dimethyl-3,4-dioleoyloxybenzylamine (DMOBA), 1,2-N,N'-dioleylcarbamyl-3-dimethylaminopropane (DOcarbDAP), 2,3-Dilinoleyloxy-N,N-dimethylpropylamine (DLinDAP), 1,2-N,N'-Dilinoleyloxy-3-dimethylaminopropane (DLincarbDAP), 1,2-Dilinoleyloxy-3-dimethylaminopropane (DLinCDAP), 2,2-dilinoleyloxy-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA), 2,2-dilinoleyloxy-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-K-XTC2-DMA), 2,2-dilinoleyloxy-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-tetradecenyl)-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-aminium (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-propane (DMDAP), 1,2-dipalmitoyl-3-dimethylammonium-propane (DPDAP), N1-[2-((1S)-1-[(3-aminopropyl)amino]-4-[di(3-amino-propyl)amino]butylcarboxamido)ethyl]-3,4-di[oleoyloxy]-benzamide (MVL5), 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (DOEPC), 2,3-bis(dodecyloxy)-N-(2-hydroxyethyl)-N,N-dimethylpropan-1-aminium bromide (DLRIE), N-(2-aminoethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)propan-1-aminium bromide (DMORIE), di((Z)-non-2-en-1-yl) 8,8'-(((2(dimethylamino)ethyl)thio)carbonyl)azanediyl)dioctanoate (ATX), N,N-dimethyl-2,3-bis(dodecyloxy)propan-1-amine (DLDMA), N,N-dimethyl-2,3-bis(tetradecyloxy)propan-1-amine (DMDMA), Di((Z)-non-2-en-1-yl)-9-((4-(dimethylaminobutanoyl)oxy)heptadecanedioate (L319), N-Dodecyl-3-((2-dodecylcarbamoyl-ethyl)-{2-[(2-dodecylcarbamoyl-ethyl)-2-[(2-dodecylcarbamoyl-ethyl)-[2-(2-dodecylcarbamoyl-ethylamino)-ethyl]-amino}-ethylamino)propionamide (lipidoid 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 embodiments, the cationic or cationically ionizable lipid is DODMA.

DOTMA is a cationic lipid with a quaternary amine headgroup. The structure of DOTMA may be represented as follows:



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



5

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.

10

Additional lipids

The RNA compositions and formulations and RNA particles described herein may also comprise lipids (including lipid-like materials) other than cationic or cationically ionizable lipids (also collectively referred to herein as cationic lipids), *i.e.*, non-cationic lipids (including non-cationic or non-cationically ionizable lipids or lipid-like materials). Collectively, anionic and neutral lipids or lipid-like materials are referred to herein as non-cationic lipids. Optimizing the formulation of 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.

15

One or more additional lipids may or may not affect the overall charge of the RNA particles. In some embodiments, the one 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.

20

In some embodiments, the RNA compositions and formulations and RNA particles described herein comprise a cationic or cationically ionizable lipid and one or more additional lipids.

25

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 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.

30

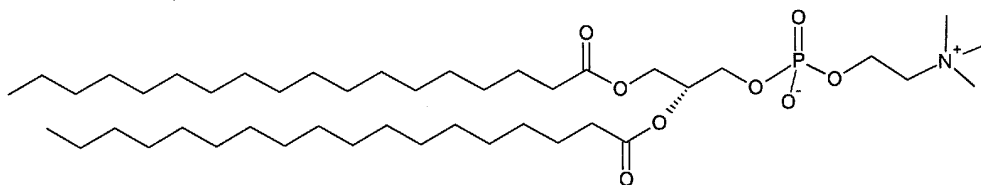
In some embodiments, the one or more additional lipids comprised in the RNA compositions and formulations and RNA particles described herein comprise one or more of the following: neutral lipids, steroids, and combinations thereof.

In some embodiments, the one or more additional lipids comprise a neutral lipid which is a phospholipid. In some embodiments, the phospholipid is selected from the group consisting of phosphatidylcholines, phosphatidylethanolamines, phosphatidylglycerols, phosphatidic acids, phosphatidylserines and sphingomyelins. Specific phospholipids that can be used include, but are not limited to, phosphatidylcholines,

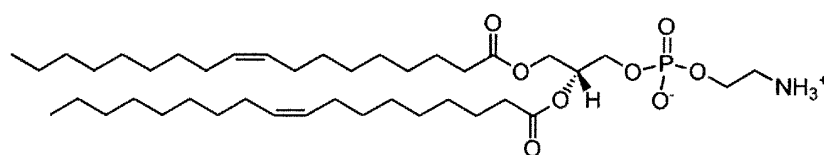
35

- phosphatidylethanolamines, phosphatidylglycerols, phosphatidic acids, phosphatidylserines or sphingomyelin. Such phospholipids include in particular diacylphosphatidylcholines, such as distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dimyristoylphosphatidylcholine (DMPC), dipentadecanoylphosphatidylcholine, dilauroylphosphatidylcholine, dipalmitoylphosphatidylcholine (DPPC), diarachidoylphosphatidylcholine (DAPC),
- 5 dibehenoylphosphatidylcholine (DBPC), ditricosanoylphosphatidylcholine (DTPC), dilignoceroylphosphatidylcholine (DLPC), palmitoyloleoyl-phosphatidylcholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemSPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC) and phosphatidylethanolamines, in particular diacylphosphatidylethanolamines, such as dioleoylphosphatidylethanolamine (DOPE), distearoyl-
- 10 phosphatidylethanolamine (DSPE), dipalmitoyl-phosphatidylethanolamine (DPPE), dimyristoyl-phosphatidylethanolamine (DMPE), dilauroyl-phosphatidylethanolamine (DLPE), diphytanoyl-phosphatidylethanolamine (DPyPE), 1,2-di-(9Z-octadecenyl)-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
- 15 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 DSPC. In some embodiments, the neutral lipid is DOPE.
- 20 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.
- 25 Thus, in some embodiments, the RNA compositions and formulations and RNA particles described herein comprise (1) a cationic or cationically ionizable lipid, and a phospholipid such as DSPC or DOPE or (2) a cationic or cationically ionizable lipid and a phospholipid such as DSPC or 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.

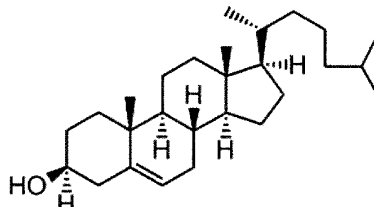
- 30 DSPC is a neutral phospholipid. The structure of DSPC may be represented as follows:



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



The structure of cholesterol may be represented as follows:



- 5 In some embodiments, RNA compositions and formulations and RNA particles described herein do not include a polymer conjugated lipid such as a pegylated lipid. The term "pegylated lipid" refers to a molecule comprising both a lipid portion and a polyethylene glycol portion. Pegylated lipids are known in the art.
- In some embodiments, the additional lipid (e.g., one or more phospholipids and/or cholesterol) may comprise from about 0 mol % to about 90 mol %, from about 0 mol % to about 80 mol %, from about 2 mol % to about 80 mol %,
- 10 %, 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)
- 15 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.

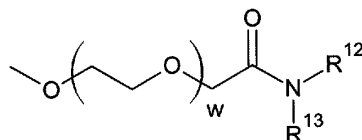
Polymer-conjugated lipids

- 20 In some embodiments, RNA compositions and formulations and RNA particles described herein 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. The term "pegylated lipid" refers to a molecule comprising both a lipid portion and a polyethylene glycol portion. Pegylated lipids are known in the art. In some embodiments,
- 25 a polymer-conjugated lipid is a polysarcosine-conjugated lipid, also referred to herein as sarcosinylated lipid or pSar-lipid. The term "sarcosinylated lipid" refers to a molecule comprising both a lipid portion and a polysarcosine portion.
- 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
- 30 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

In some embodiments, RNA compositions/formulations and RNA particles described herein comprise a PEG-conjugated lipid.

In some embodiments, the PEG-conjugated lipid (pegylated lipid) is a lipid having the structure of the following general formula:



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

each of R^{12} and R^{13} is each independently a straight or branched, alkyl or alkenyl chain containing from 10 to 30 carbon atoms, wherein the alkyl/alkenyl chain is optionally interrupted by one or more ester bonds; and w has a mean value ranging from 30 to 60.

In some embodiments of this formula, each of R^{12} and R^{13} is independently a straight alkyl chain containing from 10 to 18 carbon atoms, preferably from 12 to 16 carbon atoms.

In some embodiments of this formula, R^{12} and R^{13} are identical. In some embodiments, each of R^{12} and R^{13} is a straight alkyl chain containing 12 carbon atoms. In some embodiments, each of R^{12} and R^{13} is a straight alkyl chain containing 14 carbon atoms. In some embodiments, each of R^{12} and R^{13} is a straight alkyl chain containing 16 carbon atoms.

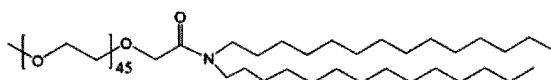
In some embodiments of this formula, R^{12} and R^{13} are different. In some embodiments, one of R^{12} and R^{13} is a straight alkyl chain containing 12 carbon atoms and the other of R^{12} and R^{13} is a straight alkyl chain containing 14 carbon atoms.

In some embodiments of this formula, w has a mean value ranging from 40 to 50, such as a mean value of 45.

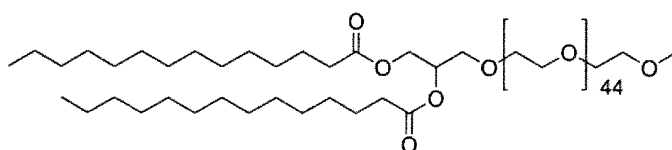
In some embodiments of this formula, w is within a range such that the PEG portion of the pegylated lipid has an average molecular weight of from about 400 to about 6000 g/mol, such as from about 1000 to about 5000 g/mol, from about 1500 to about 4000 g/mol, or from about 2000 to about 3000 g/mol. In some embodiments, each of R^{12} and R^{13} is a straight alkyl chain containing 14 carbon atoms and w has a mean value of 45.

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 dialkoxypylcarbamate such as ω -methoxy(polyethoxy)ethyl-N-(2,3-di(tetradecanoxyl)propyl)carbamate or 2,3-di(tetradecanoxyl)propyl-N-(ω -methoxy(polyethoxy)ethyl)carbamate, and the like.

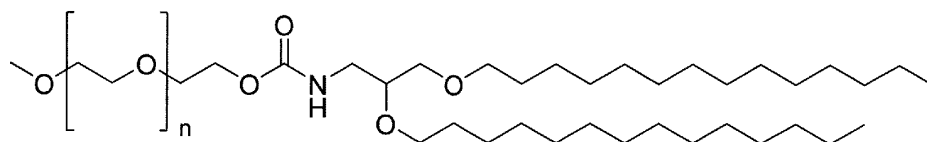
In some embodiments, the PEG-conjugated lipid (pegylated lipid) is or comprises 2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide. In some embodiments, the pegylated lipid has the following structure:



In some embodiments, the PEG-conjugated lipid (pegylated lipid) is DMG-PEG 2000, e.g., having the following structure:



In some embodiments, the PEG-conjugated lipid (pegylated lipid) has the following structure:



wherein n has a mean value ranging from 30 to 60, such as about 50. In some embodiments, the PEG-conjugated lipid (pegylated lipid) is PEG₂₀₀₀-C-DMA which preferably refers to 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).

In some embodiments, RNA compositions/formulations described herein 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.

In some embodiments, the pegylated lipid comprises from about 1 mol % to about 10 mol %, preferably from about 1 mol % to about 5 mol %, more preferably from about 1 mol % to about 2.5 mol % of the total lipid present in the RNA compositions/formulations and RNA particles described herein.

Embodiments of Lipoplex Particles

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.

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

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

RNA lipoplex particles described herein have an average diameter that in 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 675 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 some embodiments, the RNA lipoplex particles have an average diameter that ranges from about 250 nm to about 700 nm. In some embodiments, 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.

- 5 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. Spleen targeting RNA lipoplex particles are described in WO 2013/143683, herein incorporated by reference. It has been found that RNA lipoplex particles having a net negative charge may be used to preferentially target spleen tissue or spleen cells such as antigen-presenting cells, in particular dendritic cells. Accordingly, following
- 10 administration of the RNA lipoplex particles, RNA accumulation and/or RNA expression in the spleen occurs. Thus, RNA lipoplex particles of the disclosure may be used for expressing RNA in the spleen. In an embodiment, after administration of the RNA lipoplex particles, no or essentially no RNA accumulation and/or RNA expression in the lung and/or liver occurs. In some embodiments, after administration of the RNA lipoplex particles, RNA accumulation and/or RNA expression in antigen presenting cells, such as professional antigen presenting cells in the spleen
- 15 occurs. Thus, RNA lipoplex particles of the disclosure may be used for targeting RNA, e.g., RNA encoding an antigen or at least one epitope, to the lymphatic system, in particular secondary lymphoid organs, more specifically spleen. Targeting the lymphatic system, in particular secondary lymphoid organs, more specifically spleen is in particular preferred if the RNA administered is RNA encoding vaccine antigen. In some embodiments, the target cell is a spleen cell. In some embodiments, the target cell is an antigen presenting cell such as a professional antigen
- 20 presenting cell in the spleen. In some embodiments, the target cell is a dendritic cell in the spleen. The electric charge of the RNA lipoplex particles of the present disclosure is the sum of the electric charges present in the at least one cationic lipid and the electric charges present in the RNA. The charge ratio is the ratio of the positive charges present in the at least one cationic lipid to the negative charges present in the RNA. The charge ratio of the positive charges present in the at least one cationic lipid to the negative charges present in the RNA is calculated by the following equation: charge ratio = [(cationic lipid concentration (mol)) * (the total number of
- 25 positive charges in the cationic lipid)] / [(RNA concentration (mol)) * (the total number of negative charges in RNA)]. The concentration of RNA and the at least one cationic lipid amount can be determined using routine methods by one skilled in the art.
- In some embodiments, at physiological pH the charge ratio of positive charges to negative charges in the RNA
- 30 lipoplex particles is from about 1.6:2 to about 1:2, or about 1.6:2 to about 1.1:2. In specific embodiments, the charge ratio of positive charges to negative charges in the RNA lipoplex particles at physiological pH is about 1.6:2.0, about 1.5:2.0, about 1.4:2.0, about 1.3:2.0, about 1.2:2.0, about 1.1:2.0, or about 1:2.0.

Embodiments of Lipid nanoparticles (LNPs)

- 35 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: cationically ionizable lipid, 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
- 40 dissolved in ethanol with RNA in an aqueous buffer.

In some embodiments, in the RNA LNPs described herein the RNA is bound by cationically ionizable lipid that occupies the central core of the LNP. Polymer-conjugated lipid forms the surface of the LNP, along with phospholipids. In some embodiments, the surface comprises a bilayer. In some embodiments, cholesterol and cationically ionizable lipid in charged and uncharged forms can be distributed throughout the LNP.

5 In some embodiments, the LNP comprises one or more cationically ionizable lipids, and one or more stabilizing lipids. Stabilizing lipids include neutral lipids and polymer-conjugated lipids.

In some embodiments, the LNP comprises a cationically ionizable lipid, a neutral lipid, a steroid, a polymer-conjugated lipid; and the RNA, encapsulated within or associated with the lipid nanoparticle.

10 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 cationically ionizable 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.

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.

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

In some embodiments, the LNP comprises from 45 to 55 mol percent of a cationically ionizable 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.

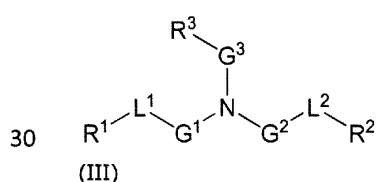
20 In some embodiments, the mol percent is determined based on total mol of lipid present in the lipid nanoparticle. In some embodiments, the mol percent is determined based on total mol of cationically ionizable lipid, neutral lipid, steroid and polymer-conjugated lipid present in the lipid nanoparticle.

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

In some embodiments, the steroid is cholesterol.

In some embodiments, the polymer conjugated lipid is a pegylated lipid, e.g., a pegylated lipid as described above.

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



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

35 one of L¹ or L² is -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, -S(O)_x-, -S-S-, -C(=O)S-, SC(=O)-, -NR^aC(=O)-, -C(=O)NR^a-, NR^aC(=O)NR^a-, -OC(=O)NR^a- or -NR^aC(=O)O-, and the other of L¹ or L² is -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, -S(O)_x-, -S-S-, -C(=O)S-, SC(=O)-, -NR^aC(=O)-, -C(=O)NR^a-, NR^aC(=O)NR^a-, -OC(=O)NR^a- or -NR^aC(=O)O- or a direct bond;

G¹ and G² are each independently unsubstituted C₁-C₁₂ alkylene or C₁-C₁₂ alkenylene;

G³ is C₁-C₂₄ alkylene, C₁-C₂₄ alkenylene, C₃-C₈ cycloalkylene, C₃-C₈ cycloalkenylene;

R^a is H or C₁-C₁₂ alkyl;

R¹ and R² are each independently C₆-C₂₄ alkyl or C₆-C₂₄ alkenyl;

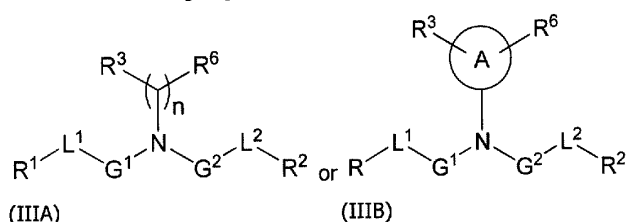
R³ is H, OR⁵, CN, -C(=O)OR⁴, -OC(=O)R⁴ or -NR⁵C(=O)R⁴;

R⁴ is C₁-C₁₂ alkyl;

5 R⁵ is H or C₁-C₆ alkyl; and

x is 0, 1 or 2.

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



10 wherein:

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

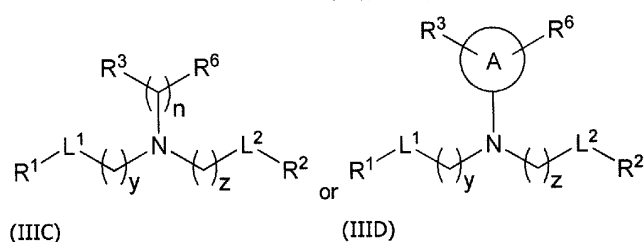
R⁶ is, at each occurrence, independently H, OH or C₁-C₂₄ alkyl;

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,

15 the lipid has structure (IIIB).

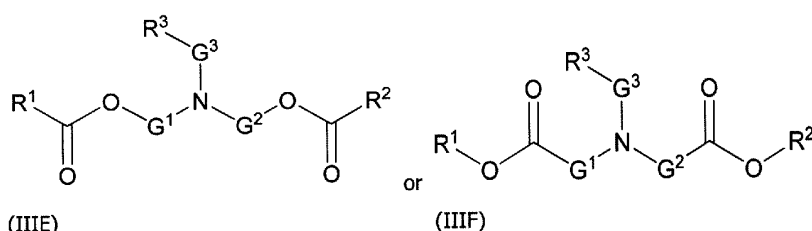
In other embodiments of Formula (III), the lipid has one of the following structures (IIIC) or (IIID):



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

20 In any of the foregoing embodiments of Formula (III), one of L¹ or L² is -O(C=O)-. For example, in some embodiments each of L¹ and L² are -O(C=O)-. In some different embodiments of any of the foregoing, L¹ and L² are each independently -(C=O)O- or -O(C=O)-. For example, in some embodiments each of L¹ and L² is -(C=O)O-.

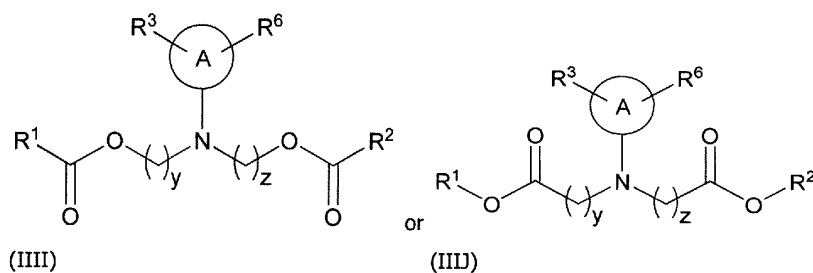
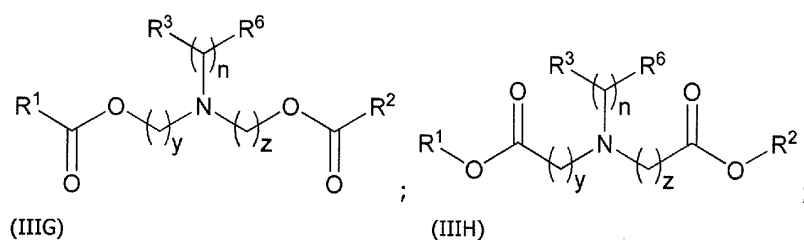
In some different embodiments of Formula (III), the lipid has one of the following structures (IIIE) or (IIIF):



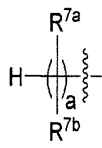
25 (IIIE)

(IIIF)

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

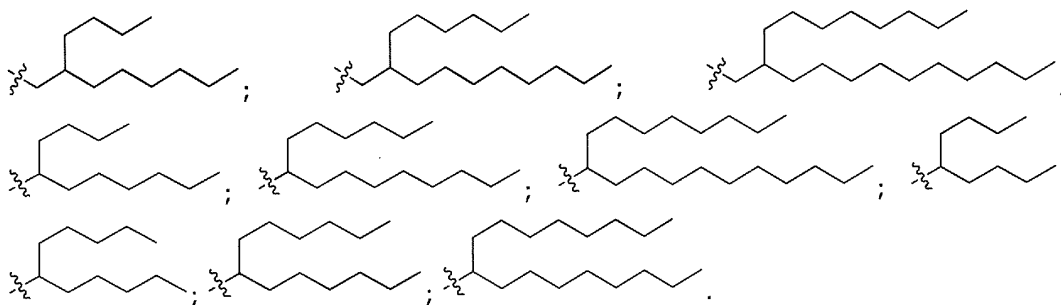


- 5 In some of the foregoing embodiments of Formula (III), n is an integer ranging from 2 to 12, for example from 2 to 8 or from 2 to 4. For example, in some embodiments, n is 3, 4, 5 or 6. In some embodiments, n is 3. In some embodiments, n is 4. In some embodiments, n is 5. In some embodiments, n is 6.
- In some other of the foregoing embodiments of Formula (III), y and z are each independently an integer ranging from 2 to 10. For example, in some embodiments, y and z are each independently an integer ranging from 4 to 9 or from 4 to 6.
- 10 In some of the foregoing embodiments of Formula (III), R^6 is H. In other of the foregoing embodiments, R^6 is C_1 - C_{24} alkyl. In other embodiments, R^6 is OH.
- In some embodiments of Formula (III), G^3 is unsubstituted. In other embodiments, G^3 is substituted. In various different embodiments, G^3 is linear C_1 - C_{24} alkylene or linear C_1 - C_{24} alkenylene.
- 15 In some other foregoing embodiments of Formula (III), R^1 or R^2 , or both, is C_6 - C_{24} alkenyl. For example, in some embodiments, R^1 and R^2 each, independently have the following structure:



wherein:

- R^{7a} and R^{7b} are, at each occurrence, independently H or C_1 - C_{12} alkyl; and
- 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 example, in some embodiments, R^{7a} is H at each occurrence. In other different embodiments of the foregoing, at least one occurrence
- 25 of R^{7b} is C_1 - C_8 alkyl. For example, in some embodiments, C_1 - C_8 alkyl is methyl, ethyl, *n*-propyl, iso-propyl, *n*-butyl, iso-butyl, *tert*-butyl, *n*-hexyl or *n*-octyl.
- In different embodiments of Formula (III), R^1 or R^2 , or both, has one of the following structures:



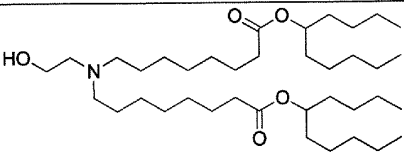
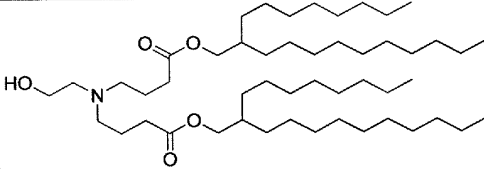
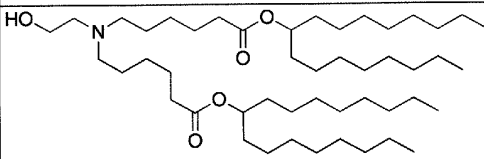
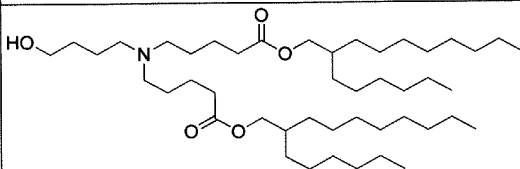
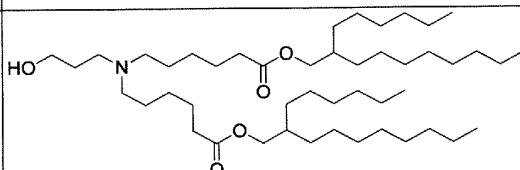
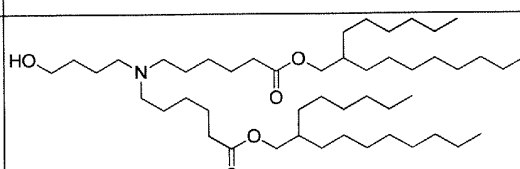
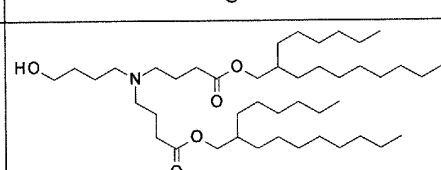
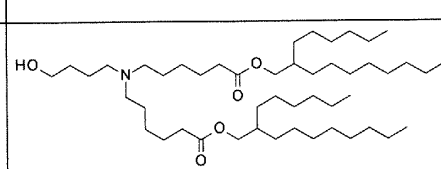
In some of the foregoing embodiments of Formula (III), R^3 is OH, CN, $-C(=O)OR^4$, $-OC(=O)R^4$ or $-NHC(=O)R^4$. In some embodiments, R^4 is methyl or ethyl.

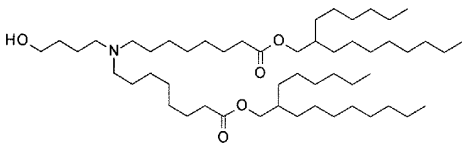
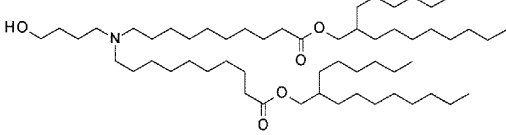
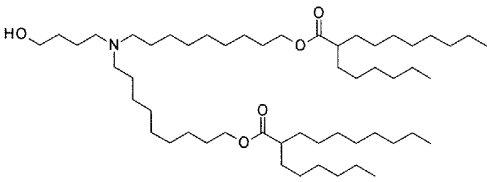
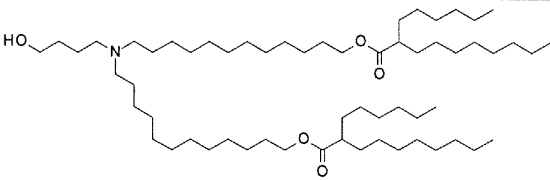
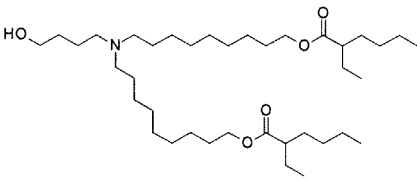
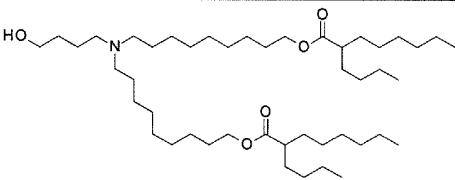
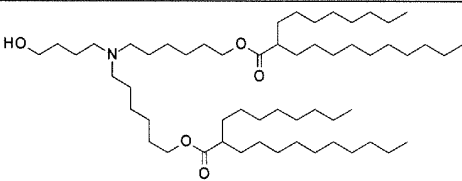
In various different embodiments, the cationic lipid of Formula (III) has one of the structures set forth in the table below.

Representative Compounds of Formula (III).

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

No.	Structure
III-5	<chem>CCCCCCCC(=O)OCCN(CCCOC(=O)CCCCCCCC)CCOC(=O)CCCCCCCC</chem>
III-6	<chem>CCCCCCCC(=O)OCCN(CCCOC(=O)CCCCCCCC)CCOC(=O)CCCCCCCC</chem>
III-7	<chem>CCCCCCCC(=O)OCCN(CCCOC(=O)CCCCCCCC)CCCCCO</chem>
III-8	<chem>CCCCCCCC(=O)OCCN(CCCOC(=O)CCCCCCCC)CCCCCO</chem>
III-9	<chem>CCCCCCCC(=O)OCCN(CCCOC(=O)CCCCCCCC)C1(CCCCC1)O</chem>
III-10	<chem>CCCCCCCC(=O)OCCN(CCCOC(=O)CCCCCCCC)CCOC(=O)CCCCCCCC</chem>
III-11	<chem>CCCCCCCC(=O)OCCN(CCCOC(=O)CCCCCCCC)CCOC(=O)CCCCCCCC</chem>

No.	Structure
III-12	
III-13	
III-14	
III-15	
III-16	
III-17	
III-18	
III-19	

No.	Structure
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	

No.	Structure
III-49	

Further representative cationically ionizable lipids are as follows:

No.	Structure
A	
B	
C	
D	
E	
F	

- 5 In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid, e.g., a cationically ionizable lipid as shown above, a neutral lipid, a steroid, and a polymer conjugated lipid. In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid of Formula III, a neutral lipid, a steroid, and a polymer conjugated lipid. In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid shown in the above tables, a neutral lipid, a steroid, and a polymer conjugated lipid.
- 10 In some embodiments, RNA described herein is formulated in an LNP composition comprising 3D-P-DMA, a neutral lipid, a steroid, and a polymer conjugated lipid.

- In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0366, a neutral lipid, a steroid, and a polymer conjugated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0315, a neutral lipid, a steroid, and a polymer conjugated lipid.
- 5 In some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the polymer conjugated lipid is a pegylated lipid, e.g., DMG-PEG 2000, PEG₂₀₀₀-C-DMA, or ALC-0159. In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid, e.g., a cationically ionizable lipid as shown above, a neutral lipid, a steroid, and a pegylated lipid.
- 10 In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid of Formula III, a neutral lipid, a steroid, and a pegylated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid shown in the above tables, a neutral lipid, a steroid, and a pegylated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising 3D-P-DMA, a neutral lipid, a steroid, and a pegylated lipid.
- 15 In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0366, a neutral lipid, a steroid, and a pegylated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0315, 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
- 20 embodiments, the pegylated lipid is DMG-PEG 2000, PEG₂₀₀₀-C-DMA, or ALC-0159.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid, e.g., a cationically ionizable lipid as shown above, DSPC, cholesterol, and a pegylated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid of Formula III, DSPC, cholesterol, and a pegylated lipid.
- 25 In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid shown in the above tables, DSPC, cholesterol, and a pegylated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising 3D-P-DMA, DSPC, cholesterol, and a pegylated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0366, DSPC,
- 30 cholesterol, and a pegylated lipid.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0315, DSPC, cholesterol, and a pegylated lipid.
- In some embodiments, the pegylated lipid is DMG-PEG 2000, PEG₂₀₀₀-C-DMA, or ALC-0159.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable
- 35 lipid, e.g., a cationically ionizable lipid as shown above, DSPC, cholesterol, and DMG-PEG 2000.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid of Formula III, DSPC, cholesterol, and DMG-PEG 2000.
- In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid shown in the above tables, DSPC, cholesterol, and DMG-PEG 2000.
- 40 In some embodiments, RNA described herein is formulated in an LNP composition comprising 3D-P-DMA, DSPC, cholesterol, and DMG-PEG 2000.

In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0366, DSPC, cholesterol, and DMG-PEG 2000.

In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0315, DSPC, cholesterol, and DMG-PEG 2000.

- 5 In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid, e.g., a cationically ionizable lipid as shown above, DSPC, cholesterol, and PEG₂₀₀₀-C-DMA.

In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid of Formula III, DSPC, cholesterol, and PEG₂₀₀₀-C-DMA.

- 10 In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid shown in the above tables, DSPC, cholesterol, and PEG₂₀₀₀-C-DMA.

In some embodiments, RNA described herein is formulated in an LNP composition comprising 3D-P-DMA, DSPC, cholesterol, and PEG₂₀₀₀-C-DMA.

In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0366, DSPC, cholesterol, and PEG₂₀₀₀-C-DMA.

- 15 In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0315, DSPC, cholesterol, and PEG₂₀₀₀-C-DMA.

In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid, e.g., a cationically ionizable lipid as shown above, DSPC, cholesterol, and ALC-0159.

- 20 In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid of Formula III, DSPC, cholesterol, and ALC-0159.

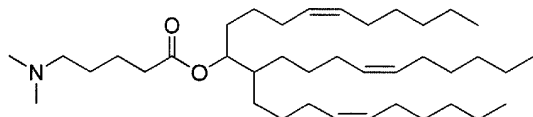
In some embodiments, RNA described herein is formulated in an LNP composition comprising a cationically ionizable lipid shown in the above tables, DSPC, cholesterol, and ALC-0159.

In some embodiments, RNA described herein is formulated in an LNP composition comprising 3D-P-DMA, DSPC, cholesterol, and ALC-0159.

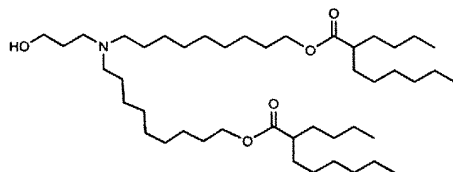
- 25 In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0366, DSPC, cholesterol, and ALC-0159.

In some embodiments, RNA described herein is formulated in an LNP composition comprising ALC-0315, DSPC, cholesterol, and ALC-0159.

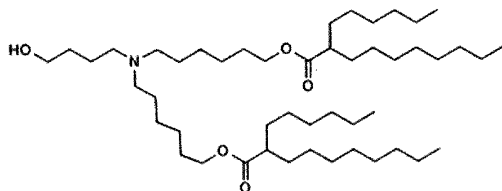
- 30 3D-P-DMA: (6Z,16Z)-12-((Z)-dec-4-en-1-yl)docosa-6,16-dien-11-yl 5-(dimethylamino)pentanoate



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

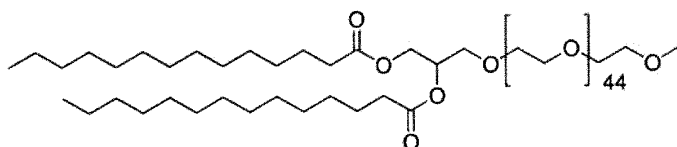


ALC-0315: ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate) / 6-[N-6-(2-hexyldecanoyloxy)hexyl-N-(4-hydroxybutyl)amino]hexyl 2-hexyldecanoate



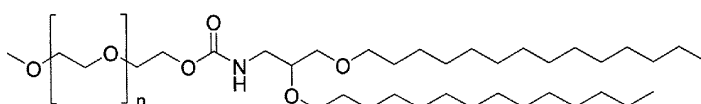
5

DMG-PEG 2000:



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))

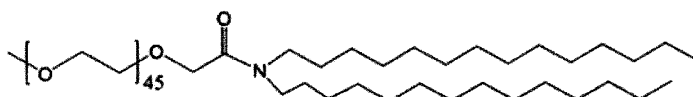
10



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

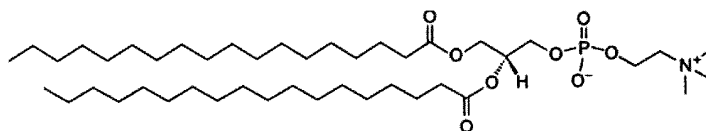
ALC-0159: 2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide / 2-[2-(ω-methoxy (polyethyleneglycol2000) ethoxy]-N,N-ditetradecylacetamide

15

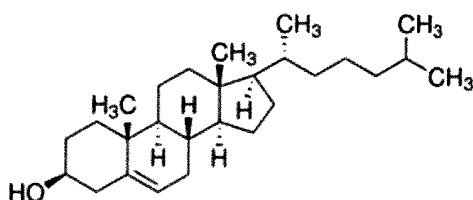


DSPC: 1,2-Distearoyl-*sn*-glycero-3-phosphocholine

20



Cholesterol:



- 5 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.

Doses

- The term "dose" as used herein refers in general to a "dose amount" which relates to the amount of RNA administered per administration, i.e., per dosing.
- 10 In some embodiments, administration of RNA of the present disclosure may be performed by single administration or boosted by multiple administrations.
- In some embodiments, an amount the RNA described herein from 0.1 µg to 300 µg, 0.5 µg to 200 µg, or 1 µg to 100 µg, such as about 1 µg, about 3 µg, about 10 µg, about 30 µg, about 50 µg, or about 100 µg may be administered per dose.
- 15 In some embodiments, a regimen described herein includes at least one dose. In some embodiments, a regimen includes a first dose and at least one subsequent dose. In some embodiments, a regimen includes a first dose and two subsequent doses. 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 regimen consists of two doses. In some embodiments, a regimen comprises three doses. In some embodiments, a regimen consists of three doses.
- 20 In some embodiments, the disclosure envisions administration of a single dose. In some embodiments, the disclosure envisions administration of a priming dose followed by one or more booster doses. The booster dose or the first booster dose may be administered 7 to 90 days, 14 to 60 days, or 30 to 60 days following administration of the priming dose. For example, the booster dose or the first booster dose may be administered about 56 days following administration of the priming dose. The second booster dose may be administered 120 to 270 days, or 150 to 210 days following administration of the priming dose. For example, the second booster dose may be administered about 180 days following administration of the priming dose.
- 25 In some embodiments, an amount of the RNA described herein of 60 µg or lower, 50 µg or lower, or 40 µg or lower may be administered per dose.
- In some embodiments, an amount of the RNA described herein of at least 0.25 µg, at least 0.5 µg, at least 1 µg, at least 2 µg, at least 3 µg, at least 4 µg, at least 5 µg, at least 10 µg, at least 20 µg, at least 30 µg, or at least 40 µg may be administered per dose.
- 30 In some embodiments, an amount of the RNA described herein of 0.25 µg to 60 µg, 0.5 µg to 55 µg, 1 µg to 50 µg, 5 µg to 40 µg, or 10 µg to 30 µg may be administered per dose.

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

In some embodiments, a first dose and a second dose (and/or other subsequent doses) may be administered by intramuscular injection. In some embodiments, a first dose and a second dose (and/or other subsequent doses) may be administered in the deltoid muscle. In some embodiments, a first dose and a second dose (and/or other subsequent doses) may be administered in the same arm. In some embodiments, an mRNA composition described herein is administered (e.g., by intramuscular injection) as a series of three doses. In some embodiments, each dose is about 30 μ g. In some embodiments, each dose is about 10 μ g. In some embodiments, each dose is about 3 μ g. In some embodiments, each dose is about 1 μ g.

In some such embodiments, an mRNA composition described herein is administered to subjects of age 12 or older. In some such embodiments, an mRNA composition described herein is administered to subjects of age 5 to 11. In some such embodiments, an mRNA composition described herein is administered to subjects of age 2 to less than 5.

In some such embodiments, an mRNA composition described herein is administered to subjects of age 12 or older and each dose is about 30 μ g. In some such embodiments, an mRNA composition described herein is administered to subjects of age 5 to 11 and each dose is about 10 μ g. In some such embodiments, an mRNA composition described herein is administered to subjects of age 2 to less than 5 and each dose is about 3 μ g.

30 Compositions comprising nucleic acid

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. In some embodiments, the buffering substance is HEPES. In some embodiments, the buffering substance is Tris.

Compositions (in particular, RNA compositions/formulations) 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 some embodiments, the cryoprotectant is a carbohydrate. The term "carbohydrate", as used herein, refers to and encompasses monosaccharides, disaccharides, trisaccharides, oligosaccharides and polysaccharides.

In some embodiments, the cryoprotectant is a monosaccharide. The term "monosaccharide", as used herein refers to a single carbohydrate unit (*e.g.*, a simple sugar) that cannot be hydrolyzed to simpler carbohydrate units.

Exemplary monosaccharide cryoprotectants include glucose, fructose, galactose, xylose, ribose and the like.

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

In some embodiments, the cryoprotectant is sucrose.

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

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

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

Other exemplary cyclic oligosaccharide cryoprotectants include compounds which include a cyclodextrin moiety in a larger molecular structure, such as a polymer that contains a cyclic oligosaccharide moiety. A cyclic oligosaccharide can be oxidized or reduced, for example, oxidized to dicarbonyl forms. The term "cyclodextrin moiety", as used herein refers to cyclodextrin (*e.g.*, an α , β , or γ cyclodextrin) radical that is incorporated into, or a part of, a larger

molecular structure, such as a polymer. A cyclodextrin moiety can be bonded to one or more other moieties directly, or through an optional linker. A cyclodextrin moiety can be oxidized or reduced, for example, oxidized to dicarbonyl forms.

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

An exemplary cryoprotectant is a polysaccharide. The term "polysaccharide", as used herein refers to a compound or a chemical moiety formed by at least 16 monosaccharide units that are bonded together through glycosidic linkages, for example through 1-4 linkages or 1-6 linkages, to form a linear, branched or cyclic structure, and includes polymers that comprise polysaccharides as part of their backbone structure. In backbones, the polysaccharide can be linear or cyclic. Exemplary polysaccharide cryoprotectants include glycogen, amylose, cellulose, dextran, maltodextrin and the like.

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 limitation, trehalose and glucose. In a specific embodiment, an alternative stabilizer to sucrose is trehalose or a mixture of sucrose and trehalose.

A preferred cryoprotectant is selected from the group consisting of sucrose, trehalose, glucose, and a combination thereof, such as a combination of sucrose and trehalose. In a preferred embodiment, the cryoprotectant is sucrose. Some embodiments of the present disclosure contemplate the use of a chelating agent in 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 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'-tetraacetic acid. In some embodiments, the chelating agent is EDTA or a salt of EDTA. In some embodiments, 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 involving an antigen, in particular tuberculosis.

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.

- 5 The pharmaceutical compositions of the present disclosure may comprise one or more adjuvants or may be administered with one or more adjuvants. The term "adjuvant" relates to a compound which prolongs, enhances or accelerates an immune response. Adjuvants comprise a heterogeneous group of compounds such as oil emulsions (*e.g.*, Freund's adjuvants), mineral compounds (such as alum), bacterial products (such as Bordetella pertussis toxin), or immune-stimulating complexes. Examples of adjuvants include, without limitation, LPS, GP96, 10 CpG oligodeoxynucleotides, growth factors, and cytokines, such as monokines, lymphokines, interleukins, chemokines. The chemokines may be IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, INF α , INF- γ , GM-CSF, LT- α . Further known adjuvants are aluminum hydroxide, Freund's adjuvant or oil such as Montanide® ISA51. Other suitable adjuvants for use in the present disclosure include lipopeptides, such as Pam3Cys, as well as lipophilic components, such as saponins, trehalose-6,6-dibehenate (TDB), monophosphoryl lipid-A (MPL), 15 monomycoloyl glycerol (MMG), or glucopyranosyl lipid adjuvant (GLA).

- The pharmaceutical compositions of the present disclosure may be in a storable form (*e.g.*, in a frozen or lyophilized/freeze-dried form) or in a "ready-to-use form" (*i.e.*, in a form which can be immediately administered to a subject, *e.g.*, without any processing such as diluting). Thus, prior to administration of a storable form of a pharmaceutical composition, this storable form has to be processed or transferred into a ready-to-use or 20 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".

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

- The term "pharmaceutically effective amount" refers to the amount which achieves a desired reaction or a desired effect alone or together with further doses. In some embodiments relating to the treatment of a particular disease, the desired reaction may relate to inhibition of the course of the disease. This comprises slowing down the progress 30 of the disease and, in some embodiments, interrupting or reversing the progress of the disease. The desired reaction in a treatment of a disease may also be delay of the onset or a prevention of the onset of said disease or said condition, or symptoms thereof. 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 35 therapy (if present), the specific route of administration and similar factors. Accordingly, the doses administered of the pharmaceutical compositions described herein may depend on various of such parameters. In the case that a reaction in a patient is insufficient with an initial dose, higher doses (or effectively higher doses achieved by a different, more localized route of administration) may be used.

- The pharmaceutical compositions of the present disclosure may contain buffers, preservatives, and optionally other 40 therapeutic agents. In some embodiments, the pharmaceutical compositions of the present disclosure comprise one or more pharmaceutically acceptable carriers, diluents and/or excipients.

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

The term "excipient" as used herein refers to a substance which may be present in a pharmaceutical composition of the present disclosure but is not an active ingredient. Examples of excipients, include without limitation, carriers, binders, diluents, lubricants, thickeners, surface active agents, preservatives, stabilizers, emulsifiers, buffers, flavoring agents, or colorants

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

The term "carrier" refers to a component which may be natural, synthetic, organic, inorganic in which the active component is combined in order to facilitate, enhance or enable administration of the pharmaceutical composition. A carrier as used herein may be one or more compatible solid or liquid fillers, diluents or encapsulating substances, which are suitable for administration to subject. Suitable carriers include, without limitation, sterile water, Ringer, Ringer lactate, sterile sodium chloride solution, isotonic saline, polyalkylene glycols, hydrogenated naphthalenes and, in particular, biocompatible lactide polymers, lactide/glycolide copolymers or polyoxyethylene/polyoxypropylene copolymers. In some embodiments, the pharmaceutical composition of the present disclosure includes isotonic saline.

Pharmaceutically acceptable carriers, excipients or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (A. R. Gennaro edit. 1985).

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

Routes of administration of pharmaceutical compositions

In some embodiments, the pharmaceutical compositions described herein may be administered intravenously, intraarterially, subcutaneously, intradermally, dermally, intranodally, or intramuscularly. In some embodiments, the pharmaceutical compositions described herein may be administered intramuscularly. In some embodiments, the pharmaceutical composition is formulated for local administration or systemic administration. Systemic administration may include enteral administration, which involves absorption through the gastrointestinal tract, or parenteral administration. As used herein, "parenteral administration" refers to the administration in any manner other than through the gastrointestinal tract, such as by intravenous injection. In some embodiments, the pharmaceutical compositions are formulated for systemic administration. In some embodiments, the systemic administration is by intravenous administration. In some embodiments, the pharmaceutical compositions are formulated for intramuscular administration.

In some embodiments, intramuscular administration comprises administration into the upper arm, in particular into the musculus deltoideus. If more than one dose, e.g., three doses, of a pharmaceutical composition described herein is administered, the different administrations may be into the same arm.

Use of compositions

Compositions described herein may be used in the therapeutic or prophylactic treatment of diseases wherein provision of one or more peptides or polypeptides, i.e., vaccine antigens, described herein to a subject results in a therapeutic or prophylactic effect. In some embodiments, the disease is infection with *Mycobacterium tuberculosis*. In some embodiments, the disease is tuberculosis.

- 5 The term "disease" (also referred to as "disorder" herein) refers to an abnormal condition that affects the body of an individual. A disease is often construed as a medical condition associated with specific symptoms and signs. A disease may be caused by factors originally from an external source, such as infectious disease, or it may be caused by internal dysfunctions, such as autoimmune diseases. In humans, "disease" is often used more broadly to refer to any condition that causes pain, dysfunction, distress, social problems, or death to the individual afflicted, or
- 10 similar problems for those in contact with the individual. In this broader sense, it sometimes includes injuries, disabilities, disorders, syndromes, infections, isolated symptoms, deviant behaviors, and atypical variations of structure and function, while in other contexts and for other purposes these may be considered distinguishable categories. Diseases usually affect individuals not only physically, but also emotionally, as contracting and living with many diseases can alter one's perspective on life, and one's personality.
- 15 The term "disease involving an antigen" refers to any disease which implicates an antigen, e.g. a disease which is characterized by the presence of an antigen. The disease involving an antigen can be an infectious disease. The antigen may be a disease-associated antigen, such as a bacterial antigen. In some embodiments, a disease involving an antigen is a disease involving cells comprising and/or expressing an antigen, and preferably presenting the antigen on the cell surface, e.g., in the context of MHC.
- 20 The term "infectious disease" refers to any disease which can be transmitted from individual to individual or from organism to organism, and is caused by a microbial agent (e.g. common cold). Infectious diseases are known in the art and include, for example, a viral disease, a bacterial disease, or a parasitic disease, which diseases are caused by a virus, a bacterium, and a parasite, respectively. In this regard, the infectious disease can be, for example, hepatitis, sexually transmitted diseases (e.g. chlamydia or gonorrhea), tuberculosis, HIV/acquired immune
- 25 deficiency syndrome (AIDS), diphtheria, hepatitis B, hepatitis C, cholera, severe acute respiratory syndrome (SARS), the bird flu, and influenza.
- In the present context, the term "treatment", "treating" or "therapeutic intervention" relates to the management and care of a subject for the purpose of combating a condition such as a disease. The term is intended to include the full spectrum of treatments for a given condition from which the subject is suffering, such as administration of
- 30 the therapeutically effective compound to alleviate the symptoms or complications, to delay the progression of the disease, disorder or condition, to alleviate or relieve 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.
- 35 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, elderly, children, and newborns. In some embodiments of the present disclosure, the "individual" or "subject" is a "patient". In some embodiments, the terms "individual" and "subject" relate to pregnant women and immunocompromised persons.

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

RNA described herein may be administered to a subject for delivering the RNA to cells of the subject.

RNA described herein may be administered to a subject for delivering a therapeutic or prophylactic peptide or polypeptide (e.g., a pharmaceutically active peptide or polypeptide) to the subject, wherein the RNA encodes a therapeutic or prophylactic peptide or polypeptide.

RNA described herein may be administered to a subject for treating or preventing a disease in a subject, wherein delivering the RNA to cells of the subject is beneficial in treating or preventing the disease.

RNA described herein may be administered to a subject for treating or preventing a disease in a subject, wherein the RNA encodes a therapeutic or prophylactic peptide or polypeptide and wherein delivering the therapeutic or prophylactic peptide or polypeptide to the subject is beneficial in treating or preventing the disease.

In some embodiments of the disclosure, the aim is to induce an immune response by providing RNA described herein.

A person skilled in the art will know that one of the principles of immunotherapy and vaccination is based on the fact that an immunoprotective reaction to a disease is produced by immunizing a subject with an antigen or an epitope, which is immunologically relevant with respect to the disease to be treated. Accordingly, RNA described herein is applicable for inducing or enhancing an immune response. RNA described herein is thus useful in a prophylactic and/or therapeutic treatment of a disease involving an antigen or epitope.

In some embodiments of the disclosure, the aim is to provide an immune response against cells comprising an antigen, e.g., Mtb antigen.

In some embodiments of the disclosure, the aim is to prophylactically or therapeutically treat tuberculosis by vaccination.

In some embodiments of the disclosure, the aim is to provide an immune response against *Mycobacterium tuberculosis* and to prevent or treat tuberculosis.

In some embodiments of the disclosure, the aim is to provide an immune response against *Mycobacterium tuberculosis* and to prevent or treat an infection with *Mycobacterium tuberculosis*.

In some embodiments of the disclosure, the aim is to provide an immune response against *Mycobacterium tuberculosis* and to prevent or treat disease symptoms caused by an infection with *Mycobacterium tuberculosis*.

In some embodiments of the disclosure, the aim is to provide protection against an infection with *Mycobacterium tuberculosis* by vaccination.

In some embodiments of the disclosure, the aim is to provide protection against an outbreak of disease in a subjected infected with *Mycobacterium tuberculosis*. In some embodiments of the disclosure, the aim is to provide protection against symptoms of tuberculosis in a subjected infected with *Mycobacterium tuberculosis*.

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

5 In some embodiments, the RNA is administered in a pharmaceutically effective amount.

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

In some embodiments, the subject treated had been exposed to *Mycobacterium tuberculosis*. In some embodiments, the subject treated had not been exposed to *Mycobacterium tuberculosis*.

10 In some embodiments, the treatments described herein involve pre- or post-exposure vaccination against *Mycobacterium tuberculosis*, or a combination thereof.

Citation of documents and studies referenced herein is not intended as an admission that any of the foregoing is pertinent prior art. All statements as to the contents of these documents are based on the information available to the applicants and do not constitute any admission as to the correctness of the contents of these documents.

15 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
20 intended to be limited to the examples described herein and shown, but are to be accorded the scope consistent with the claims.

Examples

25 Methods

Constructs and immunization:

To select the best mRNA platform for delivery of the *Mtb* antigens *in vivo*, different mRNA constructs or mixes thereof were designed and cloned. In addition to the sequence encoding the antigens (i.e., the open reading frame [ORF]), each RNA contains common structural elements optimized for maximal RNA stability and translational efficiency (5'-cap, 5'-untranslated region [UTR], 3'-UTR, and poly(A)-tail). The cap at the 5'-end of the RNA is a so-called cap0 (self-amplifying mRNA and unmodified mRNA; [m₂^{7,2'-O}GpppG]) or cap1 structure (nucleoside-modified mRNA; [m₂^{7,3'-O}Gppp(m₁^{2'-O}) ApG]). All antigen-encoding mRNA sequences were fused at the N-terminus
30 of the cell translated antigen into the endoplasmic reticulum (Figure 3, A-E; Figure 11; Figure 37). To one construct (Figure 3B) an MHC class I transmembrane and cytoplasmic domain (MITD), a cell trafficking-signal for cell membrane anchoring of the translated protein, was included in the antigen C-terminus. The construct named "6-antigen cassette" encodes a fusion protein of antigens Ag85A, ESAT6, vapB47, Hrp1, RpfA, and RpfD (Figure 3A). The construct named "6-antigen cassette with MITD" (Figure 3B) encodes the same antigens as the 6-antigen
35 cassette and has additionally a C-terminal sequence encoding MITD. The mixture of constructs named "6-antigen mix" (Figure 3C) encompasses a mixture of 6 mRNA constructs, each encoding one antigen: Ag85A, ESAT6, vapB47,

Hrp1, RpfA, and RpfD. The mixture of constructs named "2-antigen mix" (Figure 3D) encompasses a mixture of 2 mRNA constructs, each encoding one antigen: M72 and HbhA. The mixture of constructs named "6-antigen cassette + 2-antigens" (Figure 3E) is composed of a mixture of the 6-antigen cassette and the 2-antigen mix.

Unmodified mRNA (uRNA), nucleoside-modified mRNA (modRNA), and self-amplifying mRNA (saRNA) were produced by a cell free *in vitro* transcription (IVT) procedure. For efficient delivery of the mRNAs, already existing and proven technologies were used. The constructs encoded by modRNA and saRNA were formulated in lipid nanoparticles (LNP) for efficient intramuscular (i.m.) delivery to the gastrocnemius muscle. uRNA was formulated in lipoplexes (LPX) for intravenous (i.v.) injection into the tail vein. Five C57BL/6J0laHsd (C57BL/6) mice per group were immunized with 4 µg modRNA or saRNA on days 0 and 21 or with 20 µg uRNA on days 0, 7, and 21. As controls, mice were injected i.m. with 20 µL of phosphate-buffered saline (buffer). To compare the effects of the developed mRNA vaccine candidates to the effects of the only licensed TB vaccine, mice in a reference group were subcutaneously (s.c.) injected with 10⁶ colony forming units (CFU; 100 µL dose volume; BCG vaccine "SSI" – BCG Danish Strain 1331, AJ Vaccines) of bacillus Calmette-Guerin (BCG) once, on day 0. In mice a reference group of 5 animals was immunized with BCG subcutaneously (s.c.) in indicated studies. Animal health was monitored according to internal standards and protocols.

Assessment of cellular immune responses:

At 42 days after the first immunization mice were euthanized, and spleens were collected. To evaluate the induction of cellular immune responses to the *Mtb* antigens encoded by the various mRNA platforms, total splenocyte preparations, and splenocyte-derived CD4⁺ and CD8⁺ T cells (in presence of bone-marrow derived dendritic cells) were treated overnight (12 to 16 h) with 10 µg/mL purified protein-derivative (PPD, a non-defined mixture of secreted *Mtb* proteins from bacterial cultures) or with 2 µg/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens. As negative assay control, cells were treated with 2 µg/mL of a peptide from tyrosinase related protein 1 (TRP1, amino acid sequence TAPDNLGYA), unless indicated otherwise. Cells treated with culture medium only or with 2 µg/mL of concanavalin A served as technical controls (not shown). The cellular responses were analyzed by quantification of interferon gamma (IFNγ) secretion by enzyme-linked immunospot assay (ELISpot) according to the manufacturer's protocol (MABTech). ELISpot assays on total splenocytes were performed in duplicate, and in triplicate on treatment group pooled CD4⁺ or CD8⁺ T cells.

Assessment of humoral immune responses:

If not indicated otherwise at days 14, 28, and 42 after the first injection (Figure 2), blood was collected for serum analysis of antigen-binding total immunoglobulin G (IgG) by enzyme-linked immunosorbent assay (ELISA). For the challenge study, humoral responses were assessed at study days 43, 87 and 117. In short, wells of a 96-well plate were coated with 100 ng of recombinant protein for each of the analyzed antigens separately and incubated with serial dilutions of mouse serum samples. Protein-bound IgG was detected by addition of anti-mouse IgG conjugated to horse-radish peroxidase (HRP). To measure bound IgG, the HRP substrate 3,3',5,5'-tetramethylbenzidine (TMB) is added to wells, and the reaction gives a blue product (measurable at 620 nm). Addition of a strong acid stops the HRP-TMB reaction and turns its product to a super-oxidized soluble yellow product measurable at 450 nm. ELISA results are expressed as the subtraction of the absorbance at 620 nm from the absorbance at 450 nm [ΔOD (450-620 nm)].

Intracellular cytokine staining:

To assess cytokine responses by flow cytometry, 5×10^5 splenocytes were stimulated with an overlapping peptide mix in the presence of co-stimulatory antibodies (CD28 and CD49d). One hour later, Brefeldin A and Monensin were added. After another incubation of 4 to 5 h, the plates were stored at 4°C overnight. Cells were then harvested and stained for CD3, CD4 and CD8 with fluorophore-conjugated primary antibodies to identify T-cell subsets. In addition, to assess functionality of these T cells an intracellular staining for cytokines IFN- γ , IL-2 and TNF α was performed. Samples were analyzed using a FACS Celesta (BD Biosciences).

Mycobacterium tuberculosis challenge and enumeration of viable bacterial burdens:

Animals were aerosol-infected with *M. tuberculosis* strain H37Rv at study day 57. Organs were harvested from mice at study days 87 (n=10 animals) and d117 (n=5 animals) homogenized in 4.5 mL of saline using glass tube and pestles. An 1:10 serial dilution was performed including seven dilution steps, starting with 0.5 mL homogenate in 4.5 mL normal saline. 100 μ L of each dilution step were transferred to a 100 mm petri dish and dishes were incubated for 18 days at 37°C, 6% CO₂ in 60% humidity. After the incubation time, bacterial concentration was measured and CFU from serial dilutions were determined based on total volume of tissue homogenate.

Transfection of HEK293T cells for in vitro expression analyses:

HEK293T cells were transfected with uRNA or modRNA using Ribojuice transfection reagent according to the manufacturers protocol (Merck Millipore). At 16-20 h post transfection, cells were lysed for subsequent SDS-PAGE and Western blot or used for assessment of expression by antibody staining and flow cytometry.

Assessment of in vitro expression by sodium-dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot:

Total protein extract from non-transfected or transfected HEK293T cells or recombinant protein controls were denatured under reducing conditions and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was performed. Upon separation, proteins were transferred (semi-dry method) to a nitrocellulose membrane. Membranes were blocked and incubated with primary antigen-specific monoclonal (mAb) or polyclonal (pAb) antibody dilutions. Primary antibodies were detected using anti-isotype specific secondary antibodies conjugated to Horseradish Peroxidase (HRP). Membranes were developed with chemiluminescent substrate and analyzed.

Assessment of in vitro expression by flow cytometry:

RNA-transfected or untreated cells were staining for viability and fixed using commercial reagents. Cells were permeabilized and incubated with antigen-specific primary antibodies that in turn were stained with fluorophore-conjugated secondary antibodies to detect intracellular expressed antigen. Samples were measured on a FACS Celesta (BD Biosciences) and gated according to internal protocols. Only viable cells were included to assess the mean fluorescence intensities for each antigen.

Example 1: Pseudouridine-modified mRNA induces the highest immune responses in vivo

To evaluate the best performing mRNA platform, cellular immune responses were analyzed in splenocytes isolated from mice immunized with the mRNA constructs/mixtures of constructs with the different platforms (uRNA, modRNA, or saRNA) and from mice vaccinated with BCG.

- 5 C57BL/6 mice (5 animals per group) were immunized with the indicated mRNA constructs/mix of mRNA constructs (6-antigen cassette, 6-antigen cassette with MITD, 6-antigen mix, 2-antigen mix, 6-antigen cassette + 2 antigens) as unmodified mRNA (uRNA), pseudouridine-modified mRNA (modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg (100 µL dose volume) uRNA or intramuscularly twice (days 0 and 21) with 4 µg (in 20 µL dose volume) modRNA or saRNA. Mice in a reference group were subcutaneously (s.c.) injected with 10⁶ colony forming units (CFU; 100 µL dose volume) of bacillus Calmette-Guerin (BCG) once, on day 0. Mice in the control group were injected intramuscularly with 20 µL phosphate buffer saline (buffer). On day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5 × 10⁵ cells) in culture were treated overnight (12-16 h) with 10 µg/mL purified protein-derivative (PPD) and interferon gamma (IFNγ) secretion was assessed by ELISpot assay.
- 10 Splenocytes were stimulated with PPD and IFN-γ secretion was detected by ELISpot.
- 15 For all mRNA and BCG vaccinated mice a PPD-specific response was detected (Figure 4). As indicated by the highest IFNγ spot counts in each of the immunized mouse groups, the best performing mRNA platform was modRNA. Further, the highest IFNγ secretion by splenocytes were observed after immunization with the "6-antigen mix". This response was also significantly higher than that from the BCG immunized group. Surprisingly, the addition of
- 20 antigens M72 and HbHA to the 6-antigen cassette (6-antigen cassette + 2 antigens) induced only weak PPD-specific immune responses. These findings suggest that RNA, in particular the modRNA platform, is suitable for the development of a Mtb vaccine.

25 Example 2: A mixture of pseudouridine-modified mRNAs encoding 6 Mtb antigens induces cellular and humoral immune responses against three antigens

- Mice were immunized with the mRNA construct/mixture of mRNA constructs described in Figure 3 and the cellular and humoral immune responses were analyzed to select the most immunogenic one.
- 30 C57BL/6 mice (5 animals per group) were immunized with the 6-antigen cassette mRNA construct (encoding *Mycobacterium tuberculosis* antigens Ag85A, ESAT6, vapB47, Hrp1, RpfA, and RpfD) as unmodified mRNA (uRNA), or pseudouridine-modified mRNA (modRNA), or self-amplifying mRNA (saRNA). uRNA was formulated with lipoplexes, and modRNA and saRNA were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intravenously three times (days 0, 7, and 21) with 20 µg uRNA (100 µL dose volume) or intramuscularly twice (days 0 and 21) with 4 µg modRNA or saRNA (20 µL dose volume). Mice in the control group were injected
- 35 intramuscularly with 20 µL phosphate buffer saline (buffer).
- To determine the total splenocyte response, on day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5 × 10⁵ cells) in culture were treated overnight (12-16 h) with 2 µg/mL overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1 and interferon gamma (IFNγ) secretion was assessed by ELISpot assay.

To determine CD4⁺ and CD8⁺ T cell responses, CD4⁺ and CD8⁺ T cells (10⁵ cells) were selected from mouse splenocytes pooled by treatment group and treated as described above.

To determine antibody responses, blood samples were collected from mice on days 14, 28, and 42 after the first immunization to obtain the serum. Antigen-specific IgG was determined in serum samples by ELISA.

- 5 **6-antigen cassette:** Analyses of cellular responses from mice immunized with the 6-antigen cassette (Figure 3A) showed induction of antigen-specific IFN γ responses against Ag85A, ESAT6, and RpfD in total splenocyte preparations (Figure 5A). The highest secretion of IFN γ by CD8⁺ T cells were observed in response to treatment with Ag85A and RpfD, whereas only weak IFN γ secretion by CD4⁺ T-cells were seen in response to Ag85A and ESAT6 (Figure 5B). Overall, the highest responses were observed in mice immunized with the unmodified mRNA platform, whereas modRNA was also inducing antigen-specific T cells. For vapB47, Hrp1, and RpfA no cellular responses were detected. Antibody responses (Figure 5C) were observed against Ag85A and, to our surprise, also against antigens vapB47, Hrp1 and RpfA, for which no cellular responses were detected.

- 10 **6-antigen cassette with MITD:** Immunization with the signal peptide- and transmembrane domain-fused 6-antigen cassette (Figure 3B) encoded in the different RNA platforms resulted in induction cellular responses against Ag85A, ESAT6 and RpfD (Figure 6A). IFN γ secretion by CD4⁺ T cells (Figure 6B) was observed in response to Ag85A and ESAT6, being more pronounced in uRNA and modRNA-immunized animals. RpfD-specific CD8⁺ T-cell responses were found after vaccination with modRNA and saRNA. Similarly to immunization with the 6 antigen cassette, no cellular responses were seen against Hrp1, vapB47, and RpfA. Antibodies against Ag85A, as well as Hrp1, vapB47 and RpfA could be detected (Figure 6C).

- 15 **6-antigen mix:** Immunization with the 6-antigen mix induced the highest IFN γ secretion by splenocytes (Figure 7A) in response to Ag85A, ESAT6 and Hrp1, which are reflected mainly by CD4⁺ T cells (Figure 7B). The best T-cell response was seen using the uRNA platform, followed by the modRNA platform. Contrary to what was seen in response to immunization with the 6-antigen cassette, here no CD8⁺ T-cell response was detected against RpfD (Figure 7B). Antigen-specific IgG was induced against Ag85A and Hrp1.

- 20 **2-antigen mix:** Immunization with the 2-antigen mix (Figure 3D) induced a strong IFN γ secretion by splenocytes in response to M72 using the uRNA and saRNA platforms (Figure 8A). This response was fully reflected by CD8⁺ T cells (Figure 8B). No HbhA-specific T-cell responses were detected. Antibodies were detected against M72 when the uRNA and saRNA platforms were used, and against HbhA when encoded by modRNA. The lack of response to antigen M72 as modRNA was later attributed to a non-functional, faulty manufactured modRNA. After new manufacturing of the modRNA encoding the M72 construct as contained in the 2-antigen mix, another immunogenicity study was performed and strong induction of cellular and humoral responses against M72 were observed (Figure 9).

- 25 **6-antigen cassette + 2-antigens:** Immunization with the 6-antigen cassette + 2-antigens (Figure 3E) induced IFN γ secretion by splenocytes in response to Ag85A and M72 when encoded by the uRNA platform (Figure 10A and B). Antibodies were induced against Hrp1 and M72, whereas the previously observed high potential of Ag85A to induce antibodies was absent in this RNA mixture (Figure 10C).

Taken together, the highest cellular responses were observed when mice were immunized with a mixture of separately encoded *Mtb* antigens.

- 30 Example 3: Human HLA-I signal peptide fusion-antigens induce the most potent immune responses

To increase immune responses towards weak immunogens three antigens from the 6-antigen cassette (Ag85A(Δ 1-41), RpfA and vapB47) were selected and modified by fusion of signal peptide sequences (SP; SP1 or SP2) in presence or absence of transmembrane domains (TMD1, TMD2, or TMD3). Those modifications alter the localization of antigens by redirection into the secretory pathway and subsequent secretion (SP-fused antigens) or display on the cellular surface (SP/TMD-fused antigens).

C57BL/6 mice (5 animals per group) were immunized with codon optimized (opt1 or opt10) pseudouridine-modified mRNA (modRNA) construct encoding *Mycobacterium tuberculosis* antigen Ag85A(Δ 1-41) alone or with a secretion signal peptide sequence (sec, SP1, or SP2) fused to its N-terminus, as described in Figure 11. The modRNAs were formulated with C12 lipid nanoparticles (LNP-C12). Mice were injected intramuscularly twice (days 0 and 21) with 4 μ g modRNA (20 μ L dose volume). Mice in the control group were injected intramuscularly with 20 μ L phosphate buffer saline (buffer).

To determine the total splenocyte response, on day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 μ g/mL overlapping peptide pools covering the Ag85A(Δ 1-41) antigen, or RpfA antigen, or VapB47 antigen, or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay

To determine CD4⁺ and CD8⁺ T cell responses, CD4⁺ and CD8⁺ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as described above.

To determine antibody responses, blood samples were collected from mice on day 42 after the first immunization to obtain the serum. Ag85A-specific IgG, or RpfA-specific IgG, or VapB47-specific IgG was determined in serum samples by ELISA.

Ag85A(Δ 1-41) was selected based on its observed strong induction of cellular and humoral responses in mice. RpfA and vapB47 were selected based on their weak observed immunogenicity and their different molecular weight (RpfA: 40 kDa, vapB47: 11 kDa), representing a cross-section of the overall tested antigens.

Constructs encoding antigens fused to different SPs or combinations of SP/TMD were generated (Figure 11). Antigenic nucleotide sequences were codon-optimized by two different optimization methods (opt1 or opt10). All antigens were encoded by modRNA, which was formulated with LNP C12, and their immunogenicity in mice was tested. Constructs encoding the antigen (Ag85A(Δ 1-41), RpfA, or vapB47) with sec fused to its N-terminal (as tested in Examples 1 and 2) served as a benchmark.

Investigation of the immunogenicity of Ag85A(Δ 1-41) in mice as non-modified antigen or modified with different SPs revealed the plain opt1 codon-optimized antigen (without addition of an SP) to induce the highest T-cell responses (splenocytes and CD4⁺ T-cells), whereas no IFN γ -secreting CD8⁺ T-cells were detected (Figure 12). In comparison, sec-fused Ag85A(Δ 1-41) induced significantly higher production of Ag85A-specific IgG than the non-modified antigen (Figure 12C). Strikingly, for Ag85A the novel tested SP1 and SP2 performed equally good to sec for the induction of humoral responses (Figure 12C). Ag85A(Δ 1-41) in combination with all tested SP- and TMD-combinations resulted in less T-cell induction in comparison to the sec-Ag85A(Δ 1-41) (Figure 13A and B), whereas for the combination of SP1-Ag85A-TMD1 the production of antigen-specific IgG was comparable to sec-Ag85A(Δ 1-41) (Figure 13C).

Different codon- and SP-modified RpfA induced lower total splenocyte and CD4⁺ T-cell responses as compared to sec-RpfA whereas significantly higher IFN γ secretion was detected upon treatment of CD8⁺ T-cells isolated from mice immunized with SP1- or SP2-fused RpfA (Figure 14A and B). Antibody responses were highest for sec-RpfA

and, to our surprise, almost absent for RpfA without fusion to an SP (Figure 14C). Modification of RpfA with combinations of SPs and TMDs resulted in none to very weak induction of T-cell responses, whereas immunization with sec-RpfA induced high IFN γ secretion by splenocytes solely reflected by CD4 $^{+}$ T-cells (Figure 15A and B). The highest induction of RpfA-specific IgG production was observed for sec-RpfA (Figure 15C). Fusion of SP1 with one of the tested TMDs also induced antibody responses, although the response was very heterogeneous among mouse group samples (Figure 15C).

Optimization of vapB47-specific immune responses did not yield any improved splenocyte or T-cell responses (Figures 16 and Figure 17, panels A and B). Fusion of vapB47 with the SP1 and with the SP1/TMD1 combination yielded heterogeneous IgG responses that were, overall, similar to sec-fused vapB47 (Figures 16 and Figure 17, panel C).

In conclusion, despite the lack of effect from inclusion of the tested SPs (SP1, SP2) alone or in combination with TMDs (TMD1, TMD2, TMD3) to the sequence of the vapB47 antigen, results from the tested constructs for Ag85A(Δ 1-41) and RpfA have shown that the benchmark modification with SP sec generally offers improved induction of cellular and humoral immune responses.

Example 4: A mixture of four modRNAs encoding eight *Mtb* antigens induced strong and broad immune responses

Three modRNA mixes comprising 4 to 8 of the described *Mtb* antigens were generated (Figure 1) based on i. their potential to induce cellular and/or humoral immune responses (as assessed in Examples 1-3) and ii. by the most feasible combinations for mRNA molecules facilitating adequate RNA analytics of the final drug substance. The modRNA mixtures (RNA mix 1, RNA mix 2, and RNA mix 3) were formulated in LNPs using the ionizable amino lipid ALC-0315 for electrostatic interaction with the RNA, the polyethylene glycol (PEG)ylated lipid ALC-0159, which sterically stabilizes the particle, as well as 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) and cholesterol, which act as structural lipids.

C57BL/6 mice (5 animals per group) were immunized with mixtures of codon optimized (opt1) pseudouridine-modified mRNA (modRNA) constructs as described in Figure 1. Mice were injected intramuscularly twice (days 0 and 21) with 4 μ g modRNA (20 μ L dose volume). Mice in a reference group were subcutaneously (s.c.) injected with 10 6 colony forming units (CFU; 100 μ L dose volume) of bacillus Calmette-Guerin (BCG) once, on day 0. Mice in the control group were injected intramuscularly with 20 μ L phosphate buffer saline (buffer).

To determine the total splenocyte response, on day 42 after the first immunization, mice were sacrificed, and the spleens were dissected to isolate splenocytes. Splenocytes (1.25 $\times$ 10 5 cells) in culture were treated overnight (12-16 h) with 10 μ g/mL purified protein-derivative (PPD), and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. To determine CD4 $^{+}$ and CD8 $^{+}$ T cell responses, CD4 $^{+}$ and CD8 $^{+}$ T cells (10 5 cells) were selected from mouse splenocytes pooled by treatment group and treated as described above.

Upon treatment of splenocytes isolated from immunized mice with antigen-specific peptide pools, all mice groups immunized with the developed modRNA mixes showed a significantly higher induction of PPD-specific responses in comparison to the BCG immunized mice in the reference group, which was reflected exclusively by CD4 $^{+}$ T cells (Figure 18).

Additionally, splenocytes (5×10^5 cells) in culture were treated overnight (12-16 h) with 2 $\mu\text{g/mL}$ overlapping peptide pools covering the full length of each of the construct-encoded antigens or with the unspecific peptide TRP1, and interferon gamma (IFN γ) secretion was assessed by ELISpot assay. Because spot counts from treatment of splenocytes with Ag85A- and M72-specific peptide pools were too high when 5×10^5 splenocytes were used, response to these antigens was analyzed using 1.25×10^5 splenocytes, treated as described before. CD4 $^+$ and CD8 $^+$ T cells (10^5 cells) were selected from mouse splenocytes pooled by treatment group and treated as described above.

Antigen-specific cellular immune responses as determined from secretion of IFN γ by total splenocyte preparations were detected against all antigens, except vapB47 (Figure 19A). This was expected, since vapB47 was reported to only induce cellular responses in a different mouse strain than the one used in our experiments (Liang Y, et al. Autoimmunity, 2018; 51(8):417–422). Ag85A(Δ 1-41) and M72 induced similar responses in total splenocytes, independent of the modRNA mix (Figure 19B). Note that the antigens Ag85A(Δ 1-41) and M72 were analyzed using less splenocytes (1.25×10^5), since the detected spot counts using 5×10^5 cells exceeded the limit of quantification (Figure 19B). High CD4 $^+$ T-cell IFN γ responses against Ag85A(Δ 1-41), M72, ESAT6, and HbhA were induced upon immunization of mice with all three modRNA mixes. Hrp1-specific cellular responses (total splenocytes and CD4 $^+$ T cells) were induced with both Hrp1-containing RNA mixes (RNA mix 2 and RNA mix 3). With RNA mix 3 RpfA-specific CD4 $^+$ responses were detected (Figure 19C). Immunization with RNA mixes containing RpfD (RNA mix 2 and RNA mix 3) and vapB47 (RNA mix 3) did not induce relevant CD4 $^+$ T cell responses (Figure 19C). CD8 $^+$ T-cell responses against antigens Ag85A(Δ 1-41), M72, ESAT6, and RpfD were detected (Figure 19D). CD8 $^+$ T cell responses against Ag85A(Δ 1-41) and RpfD were higher after immunization with RNA mix 3.

Determination of humoral responses by antigen-specific IgG by ELISA revealed a high production of Ag85A(Δ 1-41)-specific IgG (Figure 20). Immunization with RNA mixes encoding antigens M72, Hrp1 and RpfA also induced high antibody production.

T cell and antibody responses against the immunodominant antigens Ag85A(Δ 1-41) and M72 was similar in all three RNA mixes tested.

The change of the type of formulation from LNP C-12 to Acuitas LNP 315 greatly improved immunogenicity. Taking, for example, the IFN γ response to the antigen Ag85A in the 6-antigen mix (Figure 7) in comparison to RNA mix 3, the response from total splenocytes almost doubled (average spot counts: 132 vs. 76 spots/ 10^5 cells), CD4 $^+$ response increased by 8-fold (669 vs. 83 spots/ 10^5 cells), and a CD8 $^+$ response was elicited (317 vs. 4 spots/ 10^5 cells).

Using RNA mix 3 cellular immune responses (although of varying magnitude) against the six other antigens (ESAT6, RpfD, Hrp1, RpfA, HbhA, and vapB47) were also detected. In conclusion, the broadest protection against infection and recurrence of active TB is anticipated after immunization with the RNA mix 3 vaccine candidate.

Example 5: Expression and immunogenicity analysis of uRNA Mix 3 and modRNA Mix 3

In the following, RNA mix 3 as unmodified RNA (uRNA) or modified RNA (modRNA) was further evaluated with respect to expression and immunogenicity. The cap used in both platforms was a cap1 structure, specifically m $_2^{7,3'-}$ - $^{\text{O}}$ Gppp(m $_1^{2'-}$) ApG.

To assess expression of uRNA and modRNA encoded RNA Mix 3, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot analyses of lysates from differentially transfected HEK293T/17 cells were performed.

Briefly, HEK293T/17 cells were transfected with unmodified mRNA or nucleoside-modified mRNA as single mRNA (single, 0.25 µg mRNA), comprised within a mixture of equal amounts of each of the four fusion-mRNAs (mix 1; 0.25 µg of each mRNA, total of 1 µg) or with drug substance containing the four fusion-mRNAs (mix 2; 1 µg total mRNA). Non-transfected HEK293T/17 cells served as control (NT). Whole cell lysates of non-treated and transfected cells were generated and separated by SDS-PAGE.

Western blots showed equal or increased expression of the respective antigen when the fusion antigen-encoding RNAs were mixed as compared to the single fusion RNA (Figure 21A to D). Note that the polyclonal antiserum against Hrp1 detected dimers and monomers of the fusion construct whereas the monoclonal antibody against Ag85A exclusively detected the dimer (Figure 21A).

To investigate the cellular responses induced by uRNA Mix 3, an immunogenicity study was designed in which both modRNA Mix 3 and uRNA Mix 3 were evaluated in a murine model (Figure 22 and 23).

Briefly, splenocytes were isolated on study Day 42 from C57BL/6 mice injected with uRNA Mix 3, modRNA Mix 3, or a control (saline). Cells were stimulated for 18 hours with antigen-specific overlapping peptide pools at 2 µg/mL per peptide and responses were assessed by an IFN-γ ELISpot assay.

Assessment of IFN-γ secretion from stimulated splenocytes and isolated T cells from Mix 3-injected mice showed no significant differences between modRNA Mix 3 and uRNA Mix 3 (Figure 22A-D). Antibody responses were overall higher for modRNA Mix 3 or equal between the two RNA platforms, although low antibody responses were observed for ESAT-6 and RpfD upon immunization with uRNA Mix 3.

Antigen-specific immunoglobulin G (IgG) antibodies in sera from C57BL/6 mice immunized with uRNA Mix 3, modRNA Mix 3, or a saline control were assessed by ELISA. The represented data is from Day 42 after the first immunization (Figure 23). No IgG was detected in the sera from mice in the control group.

In addition to IFN-γ secretion measured by ELISpot using total splenocytes and IgG levels measured by ELISA on sera, intracellular cytokine staining was performed to evaluate the functionality of CD4⁺ and CD8⁺ T cells. For this assay, splenocytes were stimulated with a pool of overlapping peptides covering all antigens used for immunization, in the presence of the co-stimulatory antibodies anti-CD28 and anti-CD49d. Cells were stained for CD3, CD4, and CD8 to separate T-cell subtypes and with IFN-γ, IL-2, and TNFα specific antibodies. The results showed significantly higher frequencies of CD4⁺ and CD8⁺ T cells producing single cytokines in the group that was injected with uRNA Mix 3. Frequencies of polyfunctional T cells were low but augmented in the CD8⁺ T-cell population upon immunization with uRNA Mix 3 (Figure 24A and B). Moreover, both vaccine candidates induced Mtb32a tetramer positive CD8⁺ T cells (Figure 24C).

Since the antigen VapB47 was not shown to induce cellular immune responses in the C57BL/6 mouse strain we tested uRNA Mix 3 and modRNA Mix 3 in BALB/c mice. In this strain, immunogenicity of VapB47 is reported (Mollenkopf HJ, et al. Infect Immun., 2004;72(11):6471-6479). This mouse strain shows another MHC haplotype (H2^d) than C57BL/6 mice (H2^b).

To determine the total splenocyte response, splenocytes were isolated on study Day 42 from BALB/c mice injected with uRNA Mix 3, modRNA Mix 3, or a control (saline). Total splenocytes were stimulated for 18 hours with antigen-specific overlapping peptide pools at 2 µg/mL per peptide and responses were assessed by an IFN-γ ELISpot assay. Injection of uRNA Mix 3 and modRNA Mix 3 elicited cellular responses for all antigens except ESAT-6 (Figure 25).

Antigen-specific immunoglobulin G (IgG) antibodies in sera from BALB/c mice immunized with uRNA Mix 3, modRNA Mix 3, or a saline control were assessed by ELISA. The represented data is from Day 42 after the first immunization (Figure 26).

- 5 Immunogenicity of RNA-encoded VapB47 was shown for both uRNA Mix 3 and modRNA Mix 3 whereas the former induced significantly higher responses when compared to modRNA Mix 3. The uRNA Mix 3 elicited cellular responses against RpfA significantly differed from responses induced by modRNA Mix 3. Antibodies were induced against all antigens except ESAT-6 and RpfD (Figure 26).

10 Example 6: RNA Mix 3 induces potent T-cell responses in HLA transgenic A2.1/DR1 mice

- To evaluate the immune response of RNA Mix 3 as unmodified mRNA (uRNA) or modified mRNA (modRNA) in a model as close as possible to the human setting, a non-clinical study was conducted in a mouse strain which was depleted for any murine major histocompatibility complex (MHC) molecules (Pajot et al., 2004, *Eur J Immunol.* 34(11):3060-3069) but transgenic for human HLA-A2 and HLA-DR1.
- 15 Following the same dose and immunization schedule as described in Figure 2A, the cellular and humoral responses were evaluated from uRNA Mix 3 and modRNA Mix 3 immunized mice compared to a buffer control. Briefly, to determine the total splenocyte response, splenocytes were isolated on Day 42 from humanized mice (transgenic for HLA alleles A2.1/DR1) injected with 4 µg uRNA Mix 3, 4 µg modRNA Mix 3, or a saline control (Buffer). Splenic sorted CD4⁺ and CD8⁺ T cells were stimulated with antigen-specific peptide pools and IFN-γ production was
- 20 evaluated by ELISpot assay. uRNA Mix 3 and modRNA Mix 3 immunization induced CD4⁺ and/or CD8⁺ T-cell responses against all the encoded antigens (Figure 27).

Example 7: Analyzing immune interference between RNAs comprised by modRNA Mix 3

- 25 Combining several different antigens may result in immune interference and obliterate the response to one or more individual antigens. To assess whether the mix of eight antigens encoded by the four RNAs in modRNA Mix 3 resulted in specifically compromised immune responses towards one of the antigens, C57BL/6 mice were immunized with modRNA Mix 3 (4 µg/mouse), one of the single-formulated four individual RNAs of modRNA Mix 3 (1 µg/mouse), or saline.
- 30 To determine the total splenocyte response, splenocytes were isolated on study Day 42 post prime from C57BL/6 mice injected with 4 µg modRNA Mix 3 or with 1 µg RNA-LNP (Ag85A-Hrp1, ESAT6-RpfD, RpfA-HbhA, or M72-VapB47). Control group received saline (Buffer). Cells were stimulated with antigen-specific peptide pools and responses were assessed by an IFN-γ ELISpot assay after ~18 h incubation.
- The presence in the mix had different effects on different antigens compared to their presence in a single RNA-LNP, ranging from a decreased response, to a similar response, to an increased response (see Figure 28). Cellular
- 35 responses against HbhA and VapB47 remained low; strong responses against M72 and RpfD were detected which was in line with previous observations in the C57BL/6 mouse model (compare Fig. 19).
- Antigen-specific IgG antibodies in sera from C57BL/6 mice immunized with 4 µg modRNA Mix 3 or 1 µg of RNA-LNP (Ag85A-Hrp1, ESAT6-RpfD, RpfA-HbhA, or M72-VapB47) were assessed by ELISA to determine humoral
- 40 responses. Control group received saline (Buffer). The data shown are from Day 42 post prime. Humoral responses

were either comparable upon immunization with modRNA Mix 3 or with the respective RNA-LNPs or decreased in the modRNA Mix 3 immunized group compared to their respective RNA-LNPs (Figure 29); the observed responses are in line with previous observations in the C57BL/6 mouse model (compare Fig. 20).

- Each RNA of RNA Mix 3 translates into one fusion protein, which contains two Mtb antigens. To further assess potential interference of immune responses in RNA Mix 3, cellular and humoral responses in animals injected with modRNA Mix 3 were assessed in comparison to mice that received doses of single mRNAs encoding only one of the fusion antigens or the respective single Mtb antigens. Following the immunization schedule as described in Figure 2A, C57BL/6 mice were immunized with modRNA Mix 3 (4 µg dose) or with a single RNA encoding the fusion protein Ag85A-Hrp1 (1 µg dose), or a single RNA encoding a single antigen (Ag85A(Δ1-41) [1 µg dose] or Hrp1 [1 µg dose]). Animals in the control group received buffer. Splenocytes and blood were isolated on Day 42 post prime. Splenocytes were stimulated with antigen-specific peptide pools and responses were assessed by an IFN-γ ELISpot assay. Antigen-specific immunoglobulin G (IgG) antibodies in sera were assessed by ELISA. Cellular immune responses measured by IFN-γ ELISpot against Ag85A were comparable in all test groups. Cellular IFN-γ responses against Hrp1 were decreased when Hrp1 was translated as a fusion product from Ag85A(Δ1-41)-Hrp1 or from modRNA Mix 3 compared to Hrp1 alone (Figure 30A). However, observations of humoral responses assessed by ELISA were inverse. Ag85A(Δ1-41), encoded on a single RNA, induced higher IgG titers than Ag85A(Δ1-41) comprised in a single LNP-formulated fusion-mRNA or modRNA Mix 3. For Hrp1, the antibody response was not altered between the test items (Figure 30B).
- In summary, while some antigen dependent immune interference was observed in modRNA Mix 3, all antigens showed at least one immune response (humoral, cellular or both) and the induction of immune responses against the eight Mtb antigens is expected to increase the breadth of response against Mtb.

Example 8: RNA Mix 3 is immunogenic in rats

- Wistar Han rats were administered four i.m. injections of 30 µg uRNA Mix 3 or modRNA Mix 3 or saline as control seven days apart. Seven days after the last injection, the study was terminated (Fig. 31 A). Humoral responses were assessed by total IgG ELISA 28 days post first immunization, cellular responses were assessed by ELISpot assay of total splenocytes. Briefly, IFN-γ ELISpot assay was performed using splenocytes isolated on Day 28 and stimulated for ~36 h with each of eight Mtb antigen-specific overlapping peptide pools or the respective recombinant *Mtb* proteins.
- Antigen-specific antibodies are shown in Figure 31B. In ELISpot, stimulation with overlapping peptide-pools or proteins were tested to identify T-cell responses via IFN-γ secretion from splenocytes. The data from ~36 h incubation of stimulated splenocytes is shown in Figure 32 for two antigens M72 and HbhA. The responses against the other antigens were statistically non-significant.
- In summary, data of humoral and cellular responses demonstrate induction of immune responses against uRNA Mix 3 and modRNA Mix 3 in Wistar Han rats as a further animal model.

Example 9: RNA Mix 3 induces protective immunity in a Mtb challenge model

The mouse challenge study was conducted to assess the protective efficacy of uRNA Mix 3 and modRNA Mix 3 vaccines in controlling bacterial burden following infection with Mtb. Mice were vaccinated by i.m. injection with saline, a non-translatable RNA control (NTL), uRNA Mix 3, or modRNA Mix 3. The protective efficacy of the RNA Mix 3 was also bench-marked against BCG administered subcutaneously. The study examined if vaccination using a "prime-boost" (two-dose) regimen improves the reduction in bacterial burden over a "prime" (one-dose) regimen. The study design is described in Figure 33.

Antibody titers were measured at several time points (Days 21, 43, 64, 87, and 117) to confirm the induction of immune responses following RNA Mix 3 vaccination and understand the differences over time. Selected time points and sera dilutions are shown in Figure 34 (Day 43: 3 weeks after the second injection and 2 weeks prior to Mtb challenge; Day 87 and Day 117 to estimate the levels of antigen-specific antibodies at time points after Mtb challenge). Antibody responses against six antigens (except ESAT-6 and RpfD) were detected. The antibody-responses remained until Day 117 in groups which received a two-dose immunization of uRNA Mix 3 or modRNA Mix 3 whereas a decline over time was observed. Animals that were injected only once with modRNA Mix 3 did not evolve high antibody titers. The BCG-immunized group was not tested for humoral responses.

At day 57 post-immunization, the mice were aerosol infected with a low dose of Mtb H37Rv (approximately 100 colony forming units [CFU]). Estimation of bacterial load was performed using five random mice one day post-infection which confirmed the infection dose to be approximately 85 CFU/mouse. Reduction of viable bacteria was observed for mice vaccinated with BCG compared to the saline group in lungs and spleen at 30 days post-infection. As expected, mice vaccinated with the NTL control had no reduction in bacterial burden in the lung or the spleen and remained comparable with the saline group (Figure 35). However, uRNA Mix 3 and modRNA Mix 3 immunized mice showed statistically significant reductions in bacterial load from the lung and spleen. The reduction in bacterial burden was seen across all three RNA Mix 3 immunized groups compared to saline control (Figure 35A). At 60 days after challenge, the total bacterial burden was comparable or slightly decreased as compared to Day 87 (Figure 35B). Although differences between control and immunization groups decreased, reductions in CFU counts were observed in the lung for BCG- and all RNA Mix 3 -immunized animals as compared to the saline group. In the spleen, immunization with both RNA Mix 3 significantly reduced bacterial burden (Figure 35B).

In summary, both modRNA Mix 3 vaccines confer protection in mice towards infection with the virulent Mtb-strain H37Rv.

Example 10: Antigen-specific humoral immune responses can be increased with a third immunization

To investigate the longevity of responses against the antigens comprised by uRNA Mix 3 and modRNA Mix 3 an immunogenicity study with an additional immunization was performed. For this, C57BL/6J0laHsd (C57BL/6) mice were injected twice i.m (d0 and d21) with 4 µg of uRNA Mix 3 or modRNA Mix 3 respectively. The antigen-specific antibody titers against all encoded antigens were followed by biweekly serum collection and IgG ELISA as depicted in the study scheme in Figure 36A.

After day 42 upon first immunization Antibody titers declined for all antigens whereas only weak antibody responses against RpfD and no responses against ESAT-6 were observed (Figure 36B; only data for d42 onward are shown). At day 133 post first immunization antibody titers against antigen HbhA (expressed as ΔOD (450-620 nm)) in the uRNA immunized group declined to baseline. Hence a third immunization was conducted one day later (d134) which

resulted in an increase in antibody titers against ESAT-6, RpfD, RpfA, HbhA, M72 and VapB47. Ag85A and Hrp1 humoral responses were not increased upon third immunization.

In summary, all antigen-specific humoral responses that decreased over a time course of 13 weeks increased with a third immunization.

5

Example 11: Fusion of alternative signal peptides show similar expression and immunogenicity of antigens encoded by uRNA Mix 3 and modRNA Mix 3

10 RNAs comprising RNA Mix 3 were fused to alternative signal peptides (SP) to assess whether an increased expression and secretion of the fusion antigens could be achieved. To this end the sec SP was exchanged for the two candidate sequences SP1 or SP2 for all RNAs in uRNA Mix 3 and modRNA Mix 3 (Figure 37).

In vitro expression analyses were performed by intracellular staining of transfected HEK293T cells with antigen-specific monoclonal or polyclonal antibodies and detection via flow cytometry. Gates were set to identify viable cells and expression is represented as mean-fluorescence intensity (MFI) of this viable population (Figures 38 and 39).

15 To evaluate humoral and cellular immune responses of alternative SP combined with the RNA Mix 3 fusion-antigens, C57BL/6JOLaHsd (C57BL/6) mice were immunized twice i.m. with 4 µg of uRNA or modRNA mixes (Figure 37) according to the schedule depicted in Figure 2A. Humoral responses as assessed by total IgG ELISA showed similar responses at the last study day (d42 post first immunization) for all differentially SP-fused antigens (Figure 40). As expected, no ESAT-6 and RpfD specific IgG was detected. Cellular responses were investigated by stimulation of
20 isolated splenocytes with antigen-specific peptide pools and IFN-γ ELISpot. IFN-γ production was detected for all SP-fused RNA mixes whereas IFN-γ secretion was low for vapB47 stimulated cells (Figure 41). To assess T-cell responses in more detail CD4+ and CD8+ T cells were isolated by magnetic-activated cell sorting (MACS) and stimulated with antigen-specific overlapping peptide pools. IFN-γ secretion by CD4+ T-cells (Figure 42) and CD8+ T cells (Figure 43) was induced against all antigens.

25 In summary, all alternative SP led to expression of fusion antigens and induced immune responses.

Example 12: No systemic toxicity of uRNA Mix 3 or modRNA Mix 3 after repeated i.m. dosing of Wistar Han rats

A non-GLP repeat-dose study was conducted to assess the immunogenicity and tolerability of uRNA Mix 3 and modRNA Mix 3 in rats. Female Wistar Han rats were administered two i.m. injections of uRNA Mix 3, modRNA Mix
30 3, non-translatable control (NTL), or saline (control) 21 d apart. Study endpoints included mortality, clinical observations, body weight changes, clinical pathology, organ weights, necropsy observations, and immunogenicity assessment. Administration of 30 µg uRNA Mix 3, modRNA Mix 3, and NTL to female Wistar Han rats on Days 1 and 22 was tolerated without evidence of systemic toxicity. Expected inflammatory responses to the vaccine were evident. No significant difference in the reactogenicity and tolerability profile was observed between NTL, uRNA
35 Mix 3, and modRNA Mix 3, suggesting that these effects are independent of the encoded antigens and the sequence of RNA components.

An additional GLP-compliant repeat-dose toxicity study was conducted in female and male Wistar Han rats covering 4 weekly i.m. administrations of uRNA Mix 3, modRNA Mix 3, and NTL followed by a 3-week recovery period. This study aimed to obtain information on the toxicity and immunotoxicity of uRNA Mix 3 and modRNA Mix 3 including
40 their encoded antigens and showed no vaccine-related systemic clinical signs or mortalities. There was no sign of

swelling, redness, or edema at the injection site. The animals had regular increments in body weight and comparable food/water consumption during the study. A test article-related inflammatory response was evident as typical changes in blood parameters associated with inflammation such as acute phase reactants, increase in WBC counts, local injection site reactions, transient increases in body temperature compared with controls, and
5 microscopic inflammation at the injection site, which occasionally extended into surrounding tissues. Such reactions are indicators of the pharmacology of vaccines. Effects considered secondary to immune activation, and the inflammatory response included reversible transient and early decreases in reticulocytes and reduced platelets. Similar reticulocyte and platelet changes have been observed in rats treated with other RNA vaccines but have not
10 been observed in humans. The effect is therefore considered species-specific.

Claims

1. A composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes a set of antigenic amino acid sequences, wherein the set of antigenic amino acid sequences comprises (i) at least one Mtb antigen from the acute phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof, (ii) at least one Mtb antigen from the latent phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof, and (iii) at least one Mtb antigen from the resuscitation phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof.
2. The composition or medical preparation of claim 1, wherein the at least one Mtb antigen from the acute phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof comprises
 - (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof; and/or
 - (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof.
3. The composition or medical preparation of claim 1 or 2, wherein the at least one Mtb antigen from the latent phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof comprises
 - (i) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof; and/or
 - (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof.
4. The composition or medical preparation of any one of claims 1 to 3, wherein the at least one Mtb antigen from the resuscitation phase of the Mtb life cycle, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof comprises
 - (i) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof; and/or
 - (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof.
5. A composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes a set of antigenic amino acid sequences, wherein each antigenic amino acid sequence comprises an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof and each RNA molecule encodes at least two of the antigenic amino acid sequences as fusion molecule.
6. The composition or medical preparation of claim 5, wherein each RNA molecule encodes two of the antigenic amino acid sequences as fusion molecule.

7. The composition or medical preparation of claim 5 or 6, wherein the Mtb antigens, immunogenic variants, or immunogenic fragments in a fusion molecule are not linked by a linker comprising a sequence which is heterologous to the Mtb antigens, immunogenic variants, or immunogenic fragments.

8. The composition or medical preparation of any one of claims 1 to 7, wherein the set of antigenic amino acid sequences comprises two or more, three or more, four or more, five or more, six or more, seven or more, or eight or more of the following:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof; and
- (ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

9. A composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes at least one antigenic amino acid sequence comprising an Mtb antigen, an immunogenic variant thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof, wherein the RNA encodes an amino acid sequence comprising a secretory signal peptide at the N-terminus of the encoded amino acid sequence and wherein

- (i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44, or a functional fragment of the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44; and/or
- (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43.

10. A composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes at least one antigenic amino acid sequence comprising an Mtb antigen, an immunogenic variant

thereof, or an immunogenic fragment of the Mtb antigen or the immunogenic variant thereof, wherein the RNA comprises modified uridines and/or is formulated in lipid nanoparticles.

11. The composition or medical preparation of any one of claims 1 to 10, wherein the at least one RNA molecule encodes two or more, three or more, four or more, five or more, six or more, seven or more, or eight or more of the following:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof; and
- (ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

12. The composition or medical preparation of any one of claims 1 to 11, wherein the at least one RNA molecule encodes the following amino acid sequences:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof; and

(ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

13. A composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes the following amino acid sequences:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof;
- (viii) an amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof; and
- (ix) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

14. The composition or medical preparation of any one of claims 8, and 11 to 13, wherein the amino acid sequence comprising Mtb32a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb32a or the immunogenic variant thereof and the amino acid sequence comprising Mtb39a, an immunogenic variant thereof, or an immunogenic fragment of the Mtb39a or the immunogenic variant thereof are present as a fusion protein.

15. The composition or medical preparation of claim 14, wherein the fusion protein comprises an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof.

16. The composition or medical preparation of any one of claims 1 to 15, wherein the at least one RNA molecule encodes the following amino acid sequences:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;

- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof; and
- (viii) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

17. A composition or medical preparation comprising at least one RNA molecule, wherein the at least one RNA molecule encodes the following amino acid sequences:

- (i) an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof;
- (ii) an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof;
- (iii) an amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof;
- (iv) an amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;
- (v) an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof;
- (vi) an amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;
- (vii) an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof; and
- (viii) an amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof.

18. The composition or medical preparation of any one of claims 8, and 11 to 17, wherein

- (i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 323 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 positions 27 to 323 of SEQ ID NO: 44, or an immunogenic fragment of the amino acid sequence of positions 27 to 323 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 positions 27 to 323 of SEQ ID NO: 44; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1022 of SEQ ID NO: 43.

19. The composition or medical preparation of any one of claims 8, and 11 to 18, wherein

- (i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, or an immunogenic fragment of the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 120 of SEQ ID NO: 46; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, or a fragment of the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 413 of SEQ ID NO: 45.

20. The composition or medical preparation of any one of claims 8, and 11 to 19, wherein

- (i) the amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, or an immunogenic fragment of the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 749 to 846 of SEQ ID NO: 52; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, or a fragment of the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 2298 to 2591 of SEQ ID NO: 51.

21. The composition or medical preparation of any one of claims 8, and 11 to 20, wherein

- (i) the amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 324 to 465 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 positions 324 to 465 of SEQ ID NO: 44, or an immunogenic fragment of the amino acid sequence of positions 324 to 465 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 positions 324 to 465 of SEQ ID NO: 44; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, or a fragment

of the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1023 to 1448 of SEQ ID NO: 43.

22. The composition or medical preparation of any one of claims 8, and 11 to 21, wherein

- (i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, or an immunogenic fragment of the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 432 of SEQ ID NO: 48; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or a fragment of the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1349 of SEQ ID NO: 47.

23. The composition or medical preparation of any one of claims 8, and 11 to 22, wherein

- (i) the amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, or an immunogenic fragment of the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 121 to 273 of SEQ ID NO: 46; and/or
- (ii) the RNA sequence encoding the amino acid sequence comprising RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, or a fragment of the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 414 to 872 of SEQ ID NO: 45.

24. The composition or medical preparation of any one of claims 8, and 15 to 23, wherein

- (i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, or an immunogenic fragment of the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 748 of SEQ ID NO: 52; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, or a fragment of the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2297 of SEQ ID NO: 51.

25. The composition or medical preparation of any one of claims 8, and 11 to 24, wherein

(i) the amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, or an immunogenic fragment of the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 433 to 630 of SEQ ID NO: 48; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, or a fragment of the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 1350 to 1943 of SEQ ID NO: 47.

26. The composition or medical preparation of any one of claims 1 to 25, which comprises:

(i) an RNA molecule encoding an amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof;

(ii) an RNA molecule encoding an amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof;

(iii) an RNA molecule encoding an amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof; and

(iv) an RNA molecule encoding an amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof.

27. The composition or medical preparation of claim 26, wherein

(i) the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 465 of SEQ ID NO: 44; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising Ag85A, an immunogenic variant thereof, or an immunogenic fragment of the Ag85A or the immunogenic variant thereof and Hrp1, an immunogenic variant thereof, or an immunogenic fragment of the Hrp1 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1448 of SEQ ID NO: 43.

28. The composition or medical preparation of claim 26 or 27, wherein

(i) the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 273 of SEQ ID NO: 46; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising ESAT6, an immunogenic variant thereof, or an immunogenic fragment of the ESAT6 or the immunogenic variant thereof and RpfD, an immunogenic variant thereof, or an immunogenic fragment of the RpfD or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 872 of SEQ ID NO: 45.

29. The composition or medical preparation of any one of claims 26 to 28, wherein

(i) the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 630 of SEQ ID NO: 48; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising RpfA, an immunogenic variant thereof, or an immunogenic fragment of the RpfA or the immunogenic variant thereof and HbhA, an immunogenic variant thereof, or an immunogenic fragment of the HbhA or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 1943 of SEQ ID NO: 47.

30. The composition or medical preparation of any one of claims 26 to 29, wherein

(i) the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of positions 27 to 846 of SEQ ID NO: 52; and/or

(ii) the RNA sequence encoding the amino acid sequence comprising M72, an immunogenic variant thereof, or an immunogenic fragment of the M72 or the immunogenic variant thereof and VapB47, an immunogenic variant thereof, or an immunogenic fragment of the VapB47 or the immunogenic variant thereof comprises the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 132 to 2591 of SEQ ID NO: 51.

31. The composition or medical preparation of any one of claims 1 to 8, and 10 to 30, wherein the RNA encodes an amino acid sequence comprising a secretory signal peptide.

32. The composition or medical preparation of claim 31, wherein the secretory signal peptide is fused, preferably N-terminally, to the amino acid sequence.

33. The composition or medical preparation of claim 31 or 32, wherein

- (i) the secretory signal peptide comprises the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44, or a functional fragment of the amino acid sequence of positions 1 to 26 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 positions 1 to 26 of SEQ ID NO: 44; and/or
- (ii) the RNA sequence encoding the secretory signal peptide comprises the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or a fragment of the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of positions 54 to 131 of SEQ ID NO: 43.

34. The composition or medical preparation of any one of claims 1 to 33, which comprises:

- (i) RNA comprising the nucleotide sequence of SEQ ID NO: 43, 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: 43;
- (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 45, 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: 45;
- (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 47, 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: 47; and
- (iv) RNA comprising the nucleotide sequence of SEQ ID NO: 51, 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: 51,.

35. The composition or medical preparation of any one of claims 1 to 34, which comprises:

- (i) RNA comprising the nucleotide sequence of SEQ ID NO: 43;
- (ii) RNA comprising the nucleotide sequence of SEQ ID NO: 45;
- (iii) RNA comprising the nucleotide sequence of SEQ ID NO: 47; and
- (iv) RNA comprising the nucleotide sequence of SEQ ID NO: 51.

36. The composition or medical preparation of any one of claims 1 to 35, wherein the RNA is formulated in lipid nanoparticles.

37. The composition or medical preparation of claim 36, wherein the lipid nanoparticles comprise each of:

- (i) a cationically ionizable lipid;
- (ii) a steroid;
- (iii) a neutral lipid; and
- (iv) a polymer-conjugated lipid.

38. The composition or medical preparation of claim 37, wherein the cationically ionizable lipid is present in a concentration ranging from about 40 to about 60 mol percent of the total lipids.

39. The composition or medical preparation of claim 37 or 38, wherein the steroid is present in a concentration ranging from about 30 to about 50 mol percent of the total lipids.
40. The composition or medical preparation of any one of claims 37 to 39, wherein the neutral lipid is present in a concentration ranging from about 5 to about 15 mol percent of the total lipids.
41. The composition or medical preparation of any one of claims 37 to 40, wherein the polymer-conjugated lipid is present in a concentration ranging from about 1 to about 10 mol percent of the total lipids.
42. The composition or medical preparation of any one of claims 37 to 41, wherein the cationically ionizable lipid is within a range of about 40 to about 60 mole percent, the steroid is within a range of about 30 to about 50 mole percent, the neutral lipid is within a range of about 5 to about 15 mole percent, and the polymer-conjugated lipid is within a range of about 1 to about 10 mole percent.
43. The composition or medical preparation of any one of claims 37 to 42, wherein the cationically ionizable lipid is or comprises ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate).
44. The composition or medical preparation of any one of claims 37 to 43, wherein the steroid is or comprises cholesterol.
45. The composition or medical preparation of any one of claims 37 to 44, wherein the neutral lipid is or comprises a phospholipid.
46. The composition or medical preparation of claim 45, wherein the phospholipid is or comprises distearoylphosphatidylcholine (DSPC).
47. The composition or medical preparation of any one of claims 37 to 46, wherein the polymer-conjugated lipid is or comprises a polyethylene glycol (PEG)-lipid.
48. The composition or medical preparation of claim 47, wherein the PEG-lipid is or comprises 2-[(polyethylene glycol)-2000]-*N,N*-ditetradecylacetamide.
49. The composition or medical preparation of any one of claims 36 to 48, wherein the lipid nanoparticles comprise:
- (a) ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate);
 - (b) cholesterol;
 - (c) distearoylphosphatidylcholine (DSPC); and
 - (d) 2-[(polyethylene glycol)-2000]-*N,N*-ditetradecylacetamide.
50. The composition or medical preparation of claim 49, wherein ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate) is within a range of about 40 to about 60 mole percent, cholesterol is within a range of about 30 to about 50 mole percent, distearoylphosphatidylcholine (DSPC) is within a range of about 5 to about 15 mole percent, and 2-[(polyethylene glycol)-2000]-*N,N*-ditetradecylacetamide is within a range of about 1 to about 10 mole percent.

51. The composition or medical preparation of any one of claims 1 to 50, wherein the RNA comprises a 5'-cap.
52. The composition or medical preparation of claim 51, wherein the 5' cap is or comprises a cap1 structure.
53. The composition or medical preparation of claim 51 or 52, wherein the 5'-cap is or comprises $m_2^{7,3'}\text{-OGppp}(m_1^{2'}\text{-})\text{ApG}$.
54. The composition or medical preparation of any one of claims 1 to 53, wherein the RNA comprises a 5'-UTR.
55. The composition or medical preparation of claim 54, wherein the 5'-UTR is or comprises a modified human alpha-globin 5'-UTR.
56. The composition or medical preparation of claim 54 or 55, wherein the 5'-UTR is or comprises the nucleotide sequence of SEQ ID NO: 56, 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: 56.
57. The composition or medical preparation of any one of claims 1 to 56, wherein the RNA comprises a 3'-UTR.
58. The composition or medical preparation of claim 57, wherein the 3'-UTR is or comprises a first sequence from the amino terminal enhancer of split (AES) messenger RNA and a second sequence from the mitochondrial encoded 12S ribosomal RNA.
59. The composition or medical preparation of claim 56 or 57, wherein the 3'-UTR is or comprises the nucleotide sequence of SEQ ID NO: 58, 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: 58.
60. The composition or medical preparation of any one of claims 1 to 59, wherein the RNA comprises a polyA sequence.
61. The composition or medical preparation of claim 60, wherein the polyA sequence is an interrupted sequence of A nucleotides.
62. The composition or medical preparation of claim 60 or 61, wherein the polyA sequence comprises 30 adenine nucleotides followed by 70 adenine nucleotides, wherein the 30 adenine nucleotides and 70 adenine nucleotides are separated by a linker sequence of 10 nucleotides.
63. The composition or medical preparation of any one of claims 60 to 62, wherein the polyA sequence is or comprises the nucleotide sequence of SEQ ID NO: 59, 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: 59.
64. The composition or medical preparation of any one of claims 1 to 63, wherein the RNA comprises a 5'-cap, a 5'-UTR, a 3'-UTR, and a polyA sequence.

65. The composition or medical preparation of any one of claims 1 to 64, wherein the RNA comprises modified uridines.
66. The composition or medical preparation of any one of claims 1 to 65, wherein the RNA comprises modified uridines in place of all uridines.
67. The composition or medical preparation of claim 65 or 66, wherein the modified uridines are N1-methyl-pseudouridine.
68. The composition or medical preparation of any one of claims 1 to 67, wherein the coding sequence of the RNA is codon-optimized and/or is characterized in that its G/C content is increased compared to the parental sequence.
69. The composition or medical preparation of any one of claims 1 to 68, wherein the RNA is in a liquid formulation.
70. The composition or medical preparation of any one of claims 1 to 68, wherein the RNA is in a frozen formulation.
71. The composition or medical preparation of any one of claims 1 to 68, wherein the RNA is in a lyophilized formulation.
72. The composition or medical preparation of any one of claims 1 to 71, wherein the RNA is formulated for injection.
73. The composition or medical preparation of any one of claims 1 to 72, wherein the RNA is formulated for intramuscular administration.
74. The composition or medical preparation of any one of claims 1 to 73, which is a pharmaceutical composition.
75. The composition or medical preparation of claim 74, wherein the pharmaceutical composition further comprises one or more pharmaceutically acceptable carriers, diluents and/or excipients.
76. The composition or medical preparation of any one of claims 1 to 75, which is a vaccine.
77. The composition or medical preparation of any one of claims 1 to 73, which is a kit.
78. The composition or medical preparation of claim 77, wherein different RNA molecules are in separate vials.
79. The composition or medical preparation of claim 77 or 78, further comprising instructions for use of the composition or medical preparation for treating or preventing tuberculosis.
80. The composition or medical preparation of any one of claims 1 to 79 for pharmaceutical use.
81. The composition or medical preparation of claim 80, wherein the pharmaceutical use comprises a therapeutic or prophylactic treatment of a disease or disorder.

82. The composition or medical preparation of claim 81, wherein the therapeutic or prophylactic treatment of a disease or disorder comprises treating or preventing tuberculosis.
83. The composition or medical preparation of any one of claims 1 to 82, which is for administration to a human.
84. A method of vaccinating a subject comprising administering the composition of any one of claims 1 to 83 to the subject.
85. The method of claim 84, wherein the vaccination is for preventing tuberculosis.
86. The method of claim 84 or 85, wherein administration is by intramuscular administration.
87. The method of any one of claims 84 to 86, comprising administering to the subject at least one dose of the composition.
88. The method of any one of claims 84 to 87, comprising administering to the subject at least two doses of the composition.
89. The method of any one of claims 84 to 88, wherein an amount of the RNA of at least 10 µg per dose is administered.
90. The method of any one of claims 84 to 89, wherein the subject is a human.