Abstract:

Title: MODIFIED MRNAS ENCODING CELL-PENETRATING POLYPEPTIDES

This invention relates to modified nucleic acid compositions encoding cell-penetrating polypeptides to provoke an innate immune response in a cell and methods of delivering protein-binding partners to target cells.
MODIFIED MRNAS ENCODING CELL-PENETRATING POLYPEPTIDES

CROSS-REFERENCE TO RELATED APPLICATIONS
[0001] This application claims the benefit of U.S. Provisional Patent Application No. 61/581,322, filed December 29, 2011, entitled Modified mRNAs Encoding Cell-Penetrating Polypeptides, the contents of which are herein incorporated by reference in its entirety.

FIELD OF THE INVENTION
[0002] The invention relates to compositions and methods for the manufacture of modified mRNA.

BACKGROUND OF THE INVENTION

[0004] There are multiple problems with prior methodologies of effecting protein expression. For example, heterologous deoxyribonucleic acid (DNA) introduced into a cell can be inherited by daughter cells (whether or not the heterologous DNA has integrated into the chromosome) or by offspring. Introduced DNA can integrate into host cell genomic DNA at some frequency, resulting in alterations and/or damage to the host cell genomic DNA. In addition, multiple steps must occur before a protein is made. Once inside the cell, DNA must be transported into the nucleus where it is transcribed into RNA. The RNA transcribed from DNA must then enter the cytoplasm where it is translated into protein. This need for multiple processing steps creates lag times before the generation of a protein of interest. Further, it is difficult to obtain DNA expression in cells; frequently DNA enters cells but is not expressed or not expressed at reasonable rates or concentrations. This can be a particular problem when DNA is introduced into cells such as primary cells or modified cell lines. The role of nucleoside modifications on the immuno-stimulatory potential, stability, and on the translation efficiency of RNA, and the consequent benefits to this for enhancing protein expression and producing therapeutics however, is unclear.

[0005] There is a need in the art, therefore, for biological modalities to address the modulation of intracellular translation of nucleic acids. The present invention addresses this need by providing
modified nucleic acid compositions encoding cell-penetrating polypeptides to provoke a reduced innate immune response in a cell and methods of delivering protein-binding partners to target cells.

SUMMARY OF THE INVENTION

[0006] Described herein are compositions and methods for the manufacture and optimization of modified mRNA molecules via alteration of the terminal architecture of the molecules. Specifically disclosed are compositions and methods for provoking a reduced innate immune response in a cell.

[0007] The present invention provides a composition comprising a modified nucleic acid encoding a cell-penetrating polypeptide where the modified nucleic acid may provoke a reduced innate immune response when compared to an unmodified nucleic acid encoding the same cell-penetrating polypeptide when introduced into a first cell.

[0008] The cell-penetrating polypeptide may comprise a complex comprising a charged protein operably linked to a therapeutic protein, a fusion protein comprising a charged protein covalently linked to a therapeutic protein, a therapeutic protein comprising a surface charge, a detectable label, a supercharged polypeptide, or a signal sequence for secretion of the cell-penetrating polypeptide. The cell-penetrating polypeptide may comprise a ratio of surface charge to surface amino acids which may be approximately 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9.

[0009] In other embodiments, the cell-penetrating polypeptide comprises a first domain comprising a supercharged polypeptide and a second domain comprising a protein-binding partner. The cell-penetrating polypeptide may be capable of penetrating a first and/or a second cell. The cell-penetrating polypeptide penetrating the second cell may comprise an intracellular binding partner for the protein-binding partner selected from a supercharged antibody and a functional fragment thereof. The first and the second cells may be from the same tissue, or different tissue. Further, the first and second cells may be proximal or distal to each other.

[0010] The modified nucleic acid may be introduced into a cell where the cell-penetrating polypeptide is secreted from. The cell-penetrating polypeptide may be capable of penetrating a second cell comprising second cell comprises an intracellular binding partner for the supercharged antibody or functional fragment thereof.

[0011] In one embodiment, a formulation comprises a composition comprising a modified nucleic acid encoding a cell-penetrating polypeptide and a delivery agent. The delivery agent may comprise cationic material or a cell-penetrating polypeptide.
In one embodiment, a protein-binding partner may be delivered into a target cell comprising an intracellular binding partner for the protein-binding partner. A first cell is introduced to a composition comprising a modified nucleic acid which may encode a cell-penetrating polypeptide comprising a protein-binding partner. The introduction may occur under conditions where the cell-penetrating polypeptide may be produced and may be secreted from the first cell. The cell-penetrating polypeptide may further penetrate the target cell.

The delivery of the modified nucleic acid may provoke a reduced innate immune response as compared to an unmodified nucleic acid encoding the same cell-penetrating polypeptide introduced to the first cell. The first cell and the target cell may be the same cell, co-located in a tissue of a subject, proximate or distal to each other. The modified nucleic acid may also be retained in an extracellular depot which may be adjacent to the first cell prior to introduction of the modified nucleic acid to the first cell.

The details of various embodiments of the invention are set forth in the description below. Other features, objects, and advantages of the invention will be apparent from the description and the drawings, and from the claims.

**DETAILED DESCRIPTION**

Described herein are compositions and methods for the manufacture and optimization of modified mRNA molecules via alteration of the terminal architecture of the molecules. Specifically disclosed are modified nucleic acid compositions encoding cell-penetrating polypeptides to provoke a reduced innate immune response in a cell and methods of delivering protein-binding partners to target cells.

In general, exogenous nucleic acids, particularly viral nucleic acids, introduced into cells induce an innate immune response, resulting in interferon (IFN) production and cell death. However, it is of great interest for therapeutics, diagnostics, reagents and for biological assays to deliver a nucleic acid, e.g., a ribonucleic acid (RNA) inside a cell, either in vivo or ex vivo, such as to cause intracellular translation of the nucleic acid and production of the encoded protein. Of particular importance is the delivery and function of a non-integrative nucleic acid, as nucleic acids characterized by integration into a target cell are generally imprecise in their expression levels, deleteriously transferable to progeny and neighbor cells, and suffer from the substantial risk of mutation.
Provided herein in part are nucleic acid molecules encoding polypeptides capable of modulating a cell's status, function and/or activity, and methods of making and using these nucleic acids and polypeptides. As described herein and as in copending, co-owned applications International Application PCT/US201 1/046861 filed August 5, 2011 and PCT/US201 1/054636 filed October 3, 2011, the contents of which are incorporated by reference herein in their entirety, these modified nucleic acid molecules are capable of reducing the innate immune activity of a population of cells into which they are introduced, thus increasing the efficiency of protein production in that cell population.

In addition to utilization of non-natural nucleosides and nucleotides in the modified RNAs of the present invention, it has now been discovered that use of modified nucleic acids encoding cell-penetrating polypeptides may also serve to increase protein production from a cell population as well as reduce the innate cellular immune response.

**Modified nucleic acid molecules (modified RNAs)**

This invention provides nucleic acids, including RNAs such as mRNAs that contain one or more modified nucleosides (termed "modified nucleic acids" or "modified nucleic acid molecules"), which have useful properties including the lack of a substantial induction of the innate immune response of a cell into which the mRNA is introduced. Because these modified nucleic acids enhance the efficiency of protein production, intracellular retention of nucleic acids, and viability of contacted cells, as well as possess reduced immunogenicity, these nucleic acids having these properties are termed "enhanced" nucleic acids or modified RNAs herein.

The term "nucleic acid," in its broadest sense, includes any compound and/or substance that comprise a polymer of nucleotides linked via a phosphodiester bond. These polymers are often referred to as oligonucleotides.

Exemplary nucleic acids include ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs) or hybrids thereof. They may also include RNAi-inducing agents, RNAi agents, siRNAs, shRNAs, miRNAs, antisense RNAs, ribozymes, catalytic DNA, tRNA, RNAs that induce triple helix formation, aptamers, vectors, etc. In preferred embodiments, the modified nucleic acid molecule is one or more messenger RNAs (mRNAs). As described herein, the nucleic acids of the invention do not substantially induce an innate immune response of a cell into which the mRNA is introduced.
Provided are modified nucleic acids containing a translatable region and one, two, or more than two different nucleoside modifications. In some embodiments, the modified nucleic acid exhibits reduced degradation in a cell into which the nucleic acid is introduced, relative to a corresponding unmodified nucleic acid.

In another aspect, the present disclosure provides compounds comprising a nucleotide that can disrupts binding of a major groove interacting, e.g. binding, partner with a nucleic acid, wherein the nucleotide has decreased binding affinity to major groove interacting, e.g. binding, partners.

In some embodiments, the chemical modifications can be located on the major groove face of the nucleobase, and wherein the chemical modification can include replacing or substituting an atom of a pyrimidine nucleobase with an amine, an SH, a methyl or ethyl, or a chloro or fluoro.

In some embodiments, the chemical modifications can be located on the sugar moiety of the nucleotide.

In some embodiments, the chemical modifications can be located on the phosphate backbone of the nucleotide.

In some embodiments, the chemical modifications can alter the electrochemistry on the major groove face of the nucleotide.

In certain embodiments it is desirable to intracellularly degrade a modified nucleic acid introduced into the cell, for example if precise timing of protein production is desired. Thus, the invention provides a modified nucleic acid containing a degradation domain, which is capable of being acted on in a directed manner within a cell.

Synthesis of modified RNAs

Nucleic acids for use in accordance with the invention may be prepared according to any available technique including, but not limited to chemical synthesis, enzymatic synthesis, which is generally termed in vitro transcription, enzymatic or chemical cleavage of a longer precursor, etc. Methods of synthesizing RNAs are known in the art (see, e.g., Gait, M.J. (ed.) Oligonucleotide synthesis: a practical approach, Oxford [Oxfordshire], Washington, DC: IRL Press, 1984; and Heredewijn, P. (ed.) Oligonucleotide synthesis: methods and applications, Methods in Molecular Biology, v. 288 (Clifton, N.J.) Totowa, N.J.: Humana Press, 2005; both of which are incorporated herein by reference).
The modified nucleosides and nucleotides used in the synthesis of modified RNAs disclosed herein can be prepared from readily available starting materials using the following general methods and procedures. It is understood that where typical or preferred process conditions (i.e., reaction temperatures, times, mole ratios of reactants, solvents, pressures, etc.) are given; other process conditions can also be used unless otherwise stated. Optimum reaction conditions may vary with the particular reactants or solvent used, but such conditions can be determined by one skilled in the art by routine optimization procedures.

The processes described herein can be monitored according to any suitable method known in the art. For example, product formation can be monitored by spectroscopic means, such as nuclear magnetic resonance spectroscopy (e.g., H or 13C) infrared spectroscopy, spectrophotometry (e.g., UV-visible), or mass spectrometry, or by chromatography such as high performance liquid chromatography (HPLC) or thin layer chromatography.

Preparation of modified nucleosides and nucleotides used in the manufacture or synthesis of modified RNAs of the present invention can involve the protection and deprotection of various chemical groups. The need for protection and deprotection, and the selection of appropriate protecting groups can be readily determined by one skilled in the art.

The chemistry of protecting groups can be found, for example, in Greene, et al, Protective Groups in Organic Synthesis, 2d Ed., Wiley & Sons, 1991, which is incorporated herein by reference in its entirety.

The reactions of the processes described herein can be carried out in suitable solvents, which can be readily selected by one of skill in the art of organic synthesis. Suitable solvents can be substantially nonreactive with the starting materials (reactants), the intermediates, or products at the temperatures at which the reactions are carried out, i.e., temperatures which can range from the solvent's freezing temperature to the solvent's boiling temperature. A given reaction can be carried out in one solvent or a mixture of more than one solvent. Depending on the particular reaction step, suitable solvents for a particular reaction step can be selected.

Resolution of racemic mixtures of modified nucleosides and nucleotides can be carried out by any of numerous methods known in the art. An example method includes fractional recrystallization using a "chiral resolving acid" which is an optically active, salt-forming organic acid. Suitable resolving agents for fractional recrystallization methods are, for example, optically active acids, such as the D and L forms of tartaric acid, diacetyltartaric acid, dibenzoyltartaric acid, mandelic acid, malic acid, lactic acid or the various optically active camphorsulfonic acids.
Resolution of racemic mixtures can also be carried out by elution on a column packed with an optically active resolving agent (e.g., dinitrobenzoylphenylglycine). Suitable elution solvent composition can be determined by one skilled in the art.


[00037] Modified nucleic acids need not be uniformly modified along the entire length of the molecule. Different nucleotide modifications and/or backbone structures may exist at various positions in the nucleic acid. One of ordinary skill in the art will appreciate that the nucleotide analogs or other modification(s) may be located at any position(s) of a nucleic acid such that the function of the nucleic acid is not substantially decreased. A modification may also be a 5' or 3' terminal modification. The nucleic acids may contain at a minimum one and at maximum 100% modified nucleotides, or any intervening percentage, such as at least 50% modified nucleotides, at least 80% modified nucleotides, or at least 90% modified nucleotides.

[00038] For example, the nucleic acids may contain a modified pyrimidine such as uracil or cytosine. In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 80%, at least 90% or 100% of the uracil in the nucleic acid may be replaced with a modified uracil. The modified uracil can be replaced by a compound having a single unique structure, or can be replaced by a plurality of compounds having different structures (e.g., 2, 3, 4 or more unique structures). In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 80%, at least 90% or 100% of the cytosine in the nucleic acid may be replaced with a modified cytosine. The modified cytosine can be replaced by a compound having a single unique structure, or can be replaced by a plurality of compounds having different structures (e.g., 2, 3, 4 or more unique structures).

**Modified Nucleosides and Nucleotides**

[00039] The modified mRNAs may be synthesized chemically, enzymatically or recombinantly to include one or more modified or non-natural nucleosides.

[00040] The present disclosure provides for modified nucleosides and nucleotides. As described herein "nucleoside" is defined as a compound containing a five-carbon sugar molecule (a pentose or ribose) or derivative thereof, and an organic base, purine or pyrimidine, or a derivative thereof. As described herein, "nucleotide" is defined as a nucleoside consisting of a phosphate group. The
nucleosides and nucleotides described herein are generally chemically modified on the major groove face. In some embodiments, the major groove chemical modifications can include an amino group, a thiol group, an alkyl group, or a halo group.

**Modifications on the Sugar**

[00041] The modified nucleosides and nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, as described herein, can be modified on the sugar of the ribonucleic acid. For example, the 2' hydroxyl group (OH) can be modified or replaced with a number of different "oxy" or "deoxy" substituents. Examples of "oxy" -2' hydroxyl group modifications include, but are not limited to, alkoxy or aryloxy (-OR, e.g., R = H, alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar); polyethyleneglycols (PEG), -O(CH₂CH₂₂₀)ₙCH₂CH₂OR; "locked" nucleic acids (LNA) in which the 2' hydroxyl is connected, e.g., by a methylene bridge, to the 4' carbon of the same ribose sugar; and amino groups (-O-amino, wherein the amino group, e.g., NRR, can be alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroaryl amino, ethylene diamine, polyamino) or aminoalkoxy.

[00042] "Deoxy" modifications include hydrogen, amino (e.g. NH₂; alkylamino, dialkylamino, heterocyclyl, arylamino, diaryl amino, heteroaryl amino, diheteroaryl amino, or amino acid); or the amino group can be attached to the sugar through a linker, wherein the linker comprises one or more of the atoms C, N, and O.

[00043] The sugar group can also contain one or more carbons that possess the opposite stereochemical configuration than that of the corresponding carbon in ribose. Thus, a modified RNA can include nucleotides containing e.g., arabinose, as the sugar.

**Modifications on the Phosphate Backbone**

[00044] The modified nucleosides and nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, as described herein, can be modified on the phosphate backbone. The phosphate groups of the backbone can be modified by replacing one or more of the oxygen atoms with a different substituent. Further, the modified nucleosides and nucleotides can include the wholesale replacement of an unmodified phosphate moiety with a modified phosphate as described herein. Examples of modified phosphate groups include, but are not limited to, phosphorodithioate, phosphoroselenates, borano phosphates, borano phosphate esters, hydrogen phosphonates, phosphoroamidates, alkyl or aryl phosphonates and phosphotriesters. Phosphorodithioates have both non-linking oxygens replaced by sulfur. The phosphate linker can also be modified by the
replacement of a linking oxygen with nitrogen (bridged phosphoroamidates), sulfur (bridged phosphorothioates) and carbon (bridged methylene-phosphonates).

**Modifications on the Nucleobase**

The modified nucleosides and nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, as described herein, can be modified on the nucleobase. Examples of nucleobases found in RNA include, but are not limited to, adenine, guanine, cytosine and uracil. Examples of nucleobases found in DNA include, but are not limited to, adenine, guanine, cytosine and thymine. These bases can be modified or wholly replaced to provide nucleic acids having enhanced properties, e.g. resistance to nucleases through disruption of the binding of a major groove binding partner. For example, the nucleosides and nucleotides described herein can be chemically modified on the major groove face. In some embodiments, the major groove chemical modifications can include an amino group, a thiol group, an alkyl group, or a halo group.

Table 1 below identifies the chemical faces of each canonical nucleotide. Circles identify the atoms comprising the respective chemical regions.

**Table 1**

<table>
<thead>
<tr>
<th>Pyrimidines</th>
<th>Purines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytidine:</td>
<td>Adenosine:</td>
</tr>
<tr>
<td>Uridine:</td>
<td>Guanosine:</td>
</tr>
</tbody>
</table>

Circles identify the atoms comprising the respective chemical regions.
[00047] In some embodiments, modified nucleosides include pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine, dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, and 4-methoxy-2-thio-pseudouridine.

[00048] In some embodiments, modified nucleosides include 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, and 4-methoxy-1-methyl-pseudoisocytidine.

[00049] In other embodiments, modified nucleosides include 2-aminopurine, 2, 6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, and 2-methoxy-adenine.

[00050] In other embodiments, modified nucleosides include inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.
In some embodiments, the nucleotide can be modified on the major groove face and can include replacing hydrogen on C-5 of uracil with a methyl group or a halo group.

In specific embodiments, a modified nucleoside is 5'-0-(1-Thiophosphate)-Adenosine, 5'-0-(1-Thiophosphate)-Cytidine, 5'-0-(1-Thiophosphate)-Guanosine, 5'-0-(1-Thiophosphate)-Uridine or 5'-0-(1-Thiophosphate)-Pseudouridine.
The 3′-0-(1-Thiophosphate)-Uridine

5′-0-(1-Thiophosphate)-Uridine

The α-thio substituted phosphate moiety is provided to confer stability to RNA and DNA polymers through the unnatural phosphorothioate backbone linkages. Phosphorothioate DNA and RNA have increased nuclease resistance and subsequently a longer half-life in a cellular environment. Phosphorothioate linked nucleic acids are expected to also reduce the innate immune response through weaker binding/activation of cellular innate immune molecules.

In some embodiments, the nucleoside and nucleotide can be a compound of Formula I:

I

wherein:

Z is O or S;

each of Y^1 is independently selected from -OR^al, -NR^alR^bl, and -SR^al;

each of Y^2 is independently selected from O, NR^a, S or a linker comprising an atom selected from the group consisting of C, O, N, and S;

each of Y^3 is independently selected from O and S;

Y^4 is selected from H, -OR^a, -SR^a, and -NHR^a;

n is 0, 1, 2, or 3;
m is 0, 1, 2 or 3;

B is a nucleobase;

R is H, Ci<sub>20</sub> alkyl, C<sub>2-20</sub> alkenyl, C<sub>2-20</sub> alkynyl, or C<sub>6-20</sub> aryl;

R<sup>1</sup> and R<sup>2</sup> are each independently H or a counterion; and

Y<sup>3</sup>-R<sup>Cl</sup> is OH or SH at a pH of about 1 or -Y<sup>3</sup>-R<sup>Cl</sup> is O<sup>-</sup> or S<sup>-</sup> at physiological pH;

or -Y<sup>3</sup>-R<sup>Cl</sup> is Ci<sub>20</sub> alkoxy, C<sub>2-20</sub> -O-alkenyl, or Ci<sub>2-20</sub> -O-alkynyl;

wherein when B is an unmodified nucleobase selected from cytosine, guanine, uracil and adenine, then at least one of Z, Y<sup>1</sup> or Y<sup>2</sup> is not O or OH.

In some embodiments, B is a nucleobase of Formula II-a, II-b, or II-c:

![Chemical Structures](image)

wherein:

- \( \leftrightarrow \) denotes a single or double bond;
- X is O or S;
- U and W are each independently C or N;
- V is O, S, C or N;
- where when V is C then R<sup>1</sup> is H, Ci<sub>6</sub> alkyl, Ci<sub>6</sub> alkenyl, Ci<sub>6</sub> alkynyl, halo, or -OR<sup>c</sup>, wherein Ci<sub>2-6</sub> alkyl, C<sub>2-6</sub> alkenyl, C<sub>2-20</sub> alkynyl are each optionally substituted with -OH, -NR<sup>b</sup>R<sup>b</sup>, -SH, -C(0)R<sup>c</sup>, -C(0)OR<sup>c</sup>, -NHC(0)R<sup>c</sup>, or -NHC(0)OR<sup>c</sup>;
- and wherein when V is O, S, or N then R<sup>1</sup> is absent;
- R<sup>2</sup> is H, -OR<sup>c</sup>, -SR<sup>c</sup>, -NR<sup>b</sup>R<sup>b</sup>, or halo;
- or when V is C then R<sup>1</sup> and R<sup>2</sup> together with the carbon atoms to which they are attached can form a 5- or 6-membered ring optionally substituted with 1-4 substituents selected from halo, -OH, -SH, -NR<sup>b</sup>R<sup>b</sup>, Ci<sub>20</sub> alkyl, C<sub>2-20</sub> alkenyl, C<sub>2-20</sub> alkynyl, Ci<sub>2-20</sub> alkoxy, or Ci<sub>2-20</sub> thioalkyl;
- R<sup>3</sup> is H or Ci<sub>2-20</sub> alkyl;
- R<sup>4</sup> is H or Ci<sub>2-20</sub> alkyl; wherein when \( \leftrightarrow \) denotes a double bond then R<sup>4</sup> is absent, or N-R<sup>4</sup>, taken together, forms a positively charged N substituted with Ci<sub>2-20</sub> alkyl;
- R<sup>a</sup> and R<sup>b</sup> are each independently H, Ci<sub>2-6</sub> alkyl, C<sub>2-20</sub> alkenyl, C<sub>2-20</sub> alkynyl, or C<sub>6-20</sub> aryl; and
[00082] $R^c$ is H, $C_{1-2}O$ alkyl, $C_{2-2}O$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

[00083] In some embodiments, $B$ is a nucleobase of Formula II-al, II-a2, II-a3, II-a4, or II-a5:

$$
\begin{align*}
&\text{II-al} \\
&\text{II-a2} \\
&\text{II-a3} \\
&\text{II-a4} \\
&\text{II-a5}.
\end{align*}
$$

[00084] In some embodiments, $B$ is a nucleobase selected from the group consisting of cytosine, guanine, adenine, and uracil.

[00085] In some embodiments, $B$ is a pyrimidine or derivative thereof.

[00086] In some embodiments the nucleotide is a compound of Formula I-a:

$$
\begin{align*}
&\text{I-a}.
\end{align*}
$$

[00087] In some embodiments the nucleotide is a compound of Formula I-b:

$$
\begin{align*}
&\text{I-b}.
\end{align*}
$$

[00088] In some embodiments the nucleotide is a compound of Formula I-c:
In some embodiments, the nucleotide is selected from the group consisting of:

![Chemical Structures]

In some embodiments, the nucleotide is selected from the group consisting of:
For example, the modified nucleotide can be:

In some embodiments, the major groove chemical modification can include replacement of the C-H group at C-5 with an -NH- group or a -NH(CH₃)₂- group.

For example, the modified nucleotide can be:

In another embodiment, the major groove chemical modification can include replacement of the hydrogen at C-5 of cytosine with a halo group or a methyl group.
In yet a further embodiment, the major groove chemical modification can include a fused ring that is formed by the NH₂ at the C-4 position and the carbon atom at the C-5 position.

For example, the modified nucleotide can be:

Further examples of modified nucleotides and modified nucleotide combinations are provided below in Table 2.

<table>
<thead>
<tr>
<th>Modified Nucleotide</th>
<th>Modified Nucleotide Combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-aza-cytidine</td>
<td>α-thio-cytidine/5-iodo-uridine</td>
</tr>
<tr>
<td>2-thio-cytidine</td>
<td>α-thio-cytidine/N1-methyl-pseudo-uridine</td>
</tr>
<tr>
<td>α-thio-cytidine</td>
<td>α-thio-cytidine/α-thio-uridine</td>
</tr>
<tr>
<td>Pseudo-iso-cytidine</td>
<td>α-thio-cytidine/5-methyl-uridine</td>
</tr>
<tr>
<td>5-aminoallyl-uridine</td>
<td>α-thio-cytidine/pseudo-uridine</td>
</tr>
<tr>
<td>5-iodo-uridine</td>
<td>Pseudo-iso-cytidine/5-iodo-uridine</td>
</tr>
<tr>
<td>N1-methyl-pseudouridine</td>
<td>Pseudo-iso-cytidine/N1-methyl-pseudo-uridine</td>
</tr>
<tr>
<td>5,6-dihydrouridine</td>
<td>Pseudo-iso-cytidine/α-thio-uridine</td>
</tr>
<tr>
<td>α-thio-uridine</td>
<td>Pseudo-iso-cytidine/5-methyl-uridine</td>
</tr>
<tr>
<td>4-thio-uridine</td>
<td>Pseudo-iso-cytidine/Pseudo-uridine</td>
</tr>
<tr>
<td>6-aza-uridine</td>
<td>Pyrrolo-cytidine</td>
</tr>
<tr>
<td>5-hydroxy-uridine</td>
<td>Pyrrolo-cytidine/5-iodo-uridine</td>
</tr>
<tr>
<td>Deoxy-thymidine</td>
<td>Pyrrolo-cytidine/5-methyl-pseudo-uridine</td>
</tr>
<tr>
<td>Pseudo-uridine</td>
<td>Pyrrolo-cytidine/ a-thio-uridine</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Inosine</td>
<td>Pyrrolo-cytidine/5-methyl-uridine</td>
</tr>
<tr>
<td>a-thio-guanosine</td>
<td>Pyrrolo-cytidine/5-methyl-uridine</td>
</tr>
<tr>
<td>8-oxo-guanosine</td>
<td>5-methyl-cytidine/5-iodo-uridine</td>
</tr>
<tr>
<td>06-methyl-guanosine</td>
<td>5-methyl-cytidine/5-iodo-uridine</td>
</tr>
<tr>
<td>7-deaza-guanosine</td>
<td>5-methyl-cytidine/a-thio-uridine</td>
</tr>
<tr>
<td>No modification</td>
<td>5-methyl-cytidine/5-methyl-uridine</td>
</tr>
<tr>
<td>N 1-methyl-adenosine</td>
<td>5-methyl-cytidine/5-methyl-uridine</td>
</tr>
<tr>
<td>2-amino-6-Chloro-purine</td>
<td>5-methyl-cytidine</td>
</tr>
<tr>
<td>N6-methyl-2-amino-purine</td>
<td>25% Pseudo-iso-cytidine</td>
</tr>
<tr>
<td>6-Chloro-purine</td>
<td>25% N1-methyl-pseudo-uridine</td>
</tr>
<tr>
<td>N6-methyl-adenosine</td>
<td>25% N1-Methyl-pseudo-uridine/75%-pseudo-uridine</td>
</tr>
<tr>
<td>a-thio-adenosine</td>
<td>5-methyl-uridine</td>
</tr>
<tr>
<td>8-azido-adenosine</td>
<td>5-iodo-cytidine</td>
</tr>
<tr>
<td>7-deaza-adenosine</td>
<td>5-methyl-cytidine/a-thio-uridine</td>
</tr>
</tbody>
</table>

In some embodiments, at least 25% of the cytosines are replaced by a compound of Formula I-a (e.g., at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or about 100%).

In some embodiments, at least 25% of the uracils are replaced by a compound of Formula I-a (e.g., at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or about 100%).

In some embodiments, the nucleic acid sequences comprise a compound of Formula III:
wherein:

- **Z** is O or S;
- each of **Y**\(^1\) is independently selected from -OR\(^{al}\), -NR\(^{al}\)R\(^{b1}\), and -SR\(^{al}\);
- each of **Y**\(^2\) is independently selected from O, NR\(^a\), S or a linker comprising an atom selected from the group consisting of C, O, N, and S;
- each of **Y**\(^3\) is independently selected from O and S;
- **Y**\(^4\) is selected from H, -OR\(^a\), -SR\(^a\), and -NHR\(^a\);
- **B** is a nucleobase;
- **R**\(^a\) is H, C\(_{2-20}\) alkyl, C\(_{2-20}\) alkenyl, C\(_{2-20}\) alkynyl, or C\(_{6-20}\) aryl;
- **R**\(^{al}\) and **R**\(^{b1}\) are each independently H or a counterion; and
- **Y**\(^3\)-**R**\(^{cl}\) is OH or SH at a pH of about 1 or - **Y**\(^3\)-**R**\(^{cl}\) is O\(^-\) or S\(^-\) at physiological pH;
- or - **Y**\(^3\)-**R**\(^{cl}\) is C\(_{4-9}\) alkoxy, C\(_{2-20}\)-O-alkenyl, or C\(_{1-20}\)-O-alkynyl;
- wherein when **B** is an unmodified nucleobase selected from cytosine, guanine, thymidine, uracil and adenine, then at least one of **Z**, **Y**\(^1\) or **Y**\(^2\) is not O or OH.

In some embodiments, **B** is a nucleobase of Formula II-a, II-b, or II-c:

- **II-a**
- **II-b**
- **II-c**

wherein:

- denotes a single or double bond;
- **X** is O or S;
- **U** and **W** are each independently C or N;
V is O, S, C or N;

wherein when V is C then R^1 is H, CI\textsubscript{6} alkyl, CI\textsubscript{6} alkenyl, CI\textsubscript{6} alkynyl, halo, or -OR\textsuperscript{c}, wherein CI\textsubscript{2}o alkyl, C\textsubscript{2}≥\textsubscript{10} alkenyl, C\textsubscript{2}≥\textsubscript{10} alkynyl are each optionally substituted with -OH, -NR\textsubscript{a}R\textsubscript{b}, -SH, -C(0)R\textsubscript{c}, -C(0)OR\textsubscript{c}, -NHC(0)R\textsubscript{c}, or -NHC(0)OR\textsubscript{c};

and wherein when V is O, S, or N then R^1 is absent;

R^2 is H, -OR\textsuperscript{c}, -SR\textsuperscript{c}, -NR\textsubscript{a}R\textsubscript{b}, or halo;

or when V is C then R^1 and R^2 together with the carbon atoms to which they are attached can form a 5- or 6-membered ring optionally substituted with 1-4 substituents selected from halo, -OH, -SH, -NR\textsubscript{a}R\textsubscript{b}, CI\textsubscript{2}o alkyl, C\textsubscript{2}≥\textsubscript{10} alkenyl, C\textsubscript{2}≥\textsubscript{10} alkynyl, CI\textsubscript{2}o alkoxy, or CI\textsubscript{2}o thioalkyl;

R^3 is H or CI\textsubscript{2}o alkyl;

R^4 is H or CI\textsubscript{2}o alkyl; wherein when \simline denotes a double bond then R^4 is absent, or N-R^4, taken together, forms a positively charged N substituted with CI\textsubscript{2}o alkyl;

R\textsuperscript{a} and R\textsuperscript{b} are each independently H, CI\textsubscript{2}o alkyl, C\textsubscript{2}≥\textsubscript{10} alkenyl, C\textsubscript{2}≥\textsubscript{10} alkynyl, or C\textsubscript{6}≥\textsubscript{10} aryl; and

R\textsuperscript{c} is H, CI\textsubscript{2}o alkyl, C\textsubscript{2}≥\textsubscript{10} alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

In some embodiments, B is a nucleobase of Formula II-a1, II-a2, II-a3, II-a4, or II-a5:

\[
\begin{align*}
\text{II-a1} & \\
\text{II-a2} & \\
\text{II-a3} & \\
\text{II-a4} & \\
\text{II-a5} & 
\end{align*}
\]

In some embodiments, the nucleic acid is a compound of Formula IV:

\[
\begin{align*}
\text{IV} & 
\end{align*}
\]

wherein:
denotes an optional double bond;
— denotes an optional single bond;
U is O, S, -NRa-, or -CRaRb- when \( \equiv \equiv \) denotes a single bond, or U is -CRa- when \( \equiv \equiv \) denotes a double bond;
A is H, OH, phosphoryl, pyrophosphate, sulfate, -NH₂, -SH, an amino acid, a peptide comprising 2 to 12 amino acids;
X is O or S;
each of Y₁ is independently selected from -ORc₁, -NRc₁Rd₁, and -SRc₁;
each of Y² and Y³ are independently selected from O, -CRc₃Rc₂, NRc⁵, S or a linker comprising one or more atoms selected from the group consisting of C, O, N, and S;
R₁a and R₁b are each independently H, C₁₋₁₂ alkyl, C₂₋₁₂ alkenyl, C₂₋₁₂ alkynyl, or C₆₋₁₂ aryl;
R₁c is H, C₁₋₁₂ alkyl, C₂₋₁₂ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;
R₁al and R₁bl are each independently H or a counterion;
-ORc₁ is OH at a pH of about 1 or -ORc₁ is O⁻ at physiological pH; and
B is nucleobase;
provided that the ring encompassing the variables A, B, D, U, Z, Y² and Y³ cannot be ribose.
In some embodiments, B is a nucleobase of Formula II-a, II-b, or II-c:

\[
\begin{align*}
\text{II-a} & : \quad R^1_\equiv N \equiv R^4 \\
\text{II-b} & : \quad R^3 \equiv N \equiv NH \\
\text{II-c} & : \quad R^3 \equiv N \equiv O
\end{align*}
\]

wherein:
\( \equiv \equiv \) denotes a single or double bond;
X is O or S;
U and W are each independently C or N;
V is O, S, C or N;
wherein when V is C then \( R^1 \) is H, Ci₋₆ alkyl, Ci₋₆ alkenyl, Ci₋₆ alkynyl, halo, or -ORc, wherein Ci₋₂₀ alkyl, C₂₋₂₀ alkenyl, C₂₋₂₀ alkynyl are each optionally substituted with -OH, -NRcRb, -SH, -C(0)Rc, -C(0)ORc, -NHC(0)Rc, or -NHC(0)ORc;
and wherein when \( V \) is O, S, or N then \( R^1 \) is absent;

\[ R^2 \text{ is } H, \text{-OR}^c, \text{-SR}^c, \text{-NR}^a\text{R}^b, \text{ or halo; } \]

or when \( V \) is C then \( R^1 \) and \( R^2 \) together with the carbon atoms to which they are attached can form a 5- or 6-membered ring optionally substituted with 1-4 substituents selected from halo, \(-\text{OH}, \text{-SH}, \text{-NR}^a\text{R}^b, \text{Ci}_{20} \text{ alkyl, C}2-20 \text{ alkenyl, C}2-20 \text{ alkynyl, Ci}_{20} \text{ alkoxy, or Ci}_{20} \text{ thioalkyl; } \]

\[ R^3 \text{ is } H \text{ or Ci}_{20} \text{ alkyl; } \]

\[ R^4 \text{ is } H \text{ or Ci}_{20} \text{ alkyl; wherein when denotes a double bond then } R^4 \text{ is absent, or N-} \]

\[ R^4, \text{taken together, forms a positively charged N substituted with Ci}_{20} \text{ alkyl; } \]

\[ R^a \text{ and } R^b \text{ are each independently } H, \text{Ci}_{20} \text{ alkyl, C}2-20 \text{ alkenyl, C}2-20 \text{ alkynyl, or C}6-20 \text{ aryl; and } \]

\[ R^c \text{ is } H, \text{Ci}_{20} \text{ alkyl, C}2-20 \text{ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.} \]

In some embodiments, B is a nucleobase of Formula III-al, III-a2, III-a3, III-a4, or III-a5:

![Chemical structures](image)

II-al  II-a2  II-a3  II-a4  II-a5.

In some embodiments, the nucleobase is a pyrimidine or derivative thereof.

In some embodiments, the nucleic acid contains a plurality of structurally unique compounds of Formula II-a.

In a further aspect, the modified nucleoside and nucleotide, which may be incorporated into a nucleic acid, \( e.g. \), RNA or mRNA, as provided herein can be a compound of Formula V:

![Chemical structure](image)

\[ \text{wherein:} \]

\[ \text{denotes a single or a double bond;} \]
— denotes an optional single bond;

U is O, S, -NRₐ, or -CRₐRₐ when — denotes a single bond, or U is -CRₐ when —
denotes a double bond;

Z is H, C₁₋₁₂ alkyl, or C₆₋₁₀ aryl, or Z is absent when — denotes a double bond; and

Z can be -CRₐRₐ- and form a bond with A;

A is H, OH, NHR wherein R= alkyl or aryl or phorphoryl, sulfate, -NH₂, N₃, azido, -SH, N an amino acid, or a peptide comprising 1 to 12 amino acids;

D is H, OH, NHR wherein R= alkyl or aryl or phorphoryl, -NH₂, -SH, an amino acid, a peptide comprising 1 to 12 amino acids, or a group of Formula VII:

\[
\begin{array}{c}
Y_2 \\
\hline
P=X \\
\hline
OR^{\text{C}}_\text{m}
\end{array}
\]

VII

or A and D together with the carbon atoms to which they are attached form a 5-
membered ring;

X is O or S;

each of Y₁ is independently selected from -ORₐ, -NRₐRₐ, and -SRₐ;

each of Y₂ and Y₃ are independently selected from O, -CRₐRₐ, -NRₐ, S or a linker
comprising one or more atoms selected from the group consisting of C, O, N, and S;

n is 0, 1, 2, or 3;

m is 0, 1, 2 or 3;

B is nucleobase;

Rₐ and Rₐ are each independently H, C₁₋₁₂ alkyl, C₂₋₁₂ alkenyl, C₂₋₁₂ alkynyl, or C₆₋₁₀ aryl;

Rₜ is H, C₁₋₁₂ alkyl, C₂₋₁₂ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an
amino-polyethylene glycol group;

Rₐ and Rₐ are each independently H or a counterion; and

-ORₜ is OH at a pH of about 1 or -ORₜ is O⁻ at physiological pH;

provided that the ring encompassing the variables A, B, D, U, Z, Y² and Y³ cannot be
ribose.

In some embodiments, B may be a nucleobase of Formula II-a, II-b, or II-c:
wherein:
- \( \overline{\overset{\sim}{\sim}} \) denotes a single or double bond;
- \( X \) is O or S;
- U and W are each independently C or N;
- V is O, S, C, or N;
- when V is C then \( R^1 \) is H, \( \text{Ci}_{-6} \) alkyl, \( \text{Ci}_{-6} \) alkenyl, \( \text{Ci}_{-6} \) alkynyl, halo, or -OR;
- when \( \text{Ci}_{-2} \) alkyl, \( \text{Ci}_{-2} \) alkenyl, \( \text{Ci}_{-2} \) alkynyl are each optionally substituted with -OH, -NR, -SH, -C(O)R, -C(O)OR, -NHC(O)R, or -NHC(O)OR;
- and wherein when V is O, S, or N then \( R^1 \) is absent;
- \( R^2 \) is H, -OR, -SR, -NR, or halo;
- or when V is C then \( R^1 \) and \( R^2 \) together with the carbon atoms to which they are attached can form a 5- or 6-membered ring optionally substituted with 1-4 substituents selected from halo, -OH, -SH, -NR, -C(O)alkyl, \( \text{Ci}_{-2} \) alkyl, \( \text{Ci}_{-2} \) alkenyl, \( \text{Ci}_{-2} \) alkynyl, \( \text{Ci}_{-2} \) alkoxy, or \( \text{Ci}_{-2} \) thioalkyl;
- \( R^3 \) is H or \( \text{Ci}_{-2} \) alkyl;
- \( R^4 \) is H or \( \text{Ci}_{-2} \) alkyl;
- wherein \( \overline{\overset{\sim}{\sim}} \) denotes a double bond then \( R^4 \) is absent, or \( N-R^4 \), taken together, forms a positively charged N substituted with \( \text{Ci}_{-2} \) alkyl;
- \( R^a \) and \( R^b \) are each independently H, \( \text{Ci}_{-2} \) alkyl, \( \text{Ci}_{-2} \) alkenyl, \( \text{Ci}_{-2} \) alkynyl, or \( \text{C}_{6-2} \) aryl; and
- \( R^c \) is H, \( \text{Ci}_{-2} \) alkyl, \( \text{Ci}_{-2} \) alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

In some embodiments, B may be a nucleobase of Formula II-a1, II-a2, II-a3, II-a4, or II-a5:
In some embodiments, B may be a nucleobase selected from the group consisting of cytosine, guanine, adenine, and uracil.

In some embodiments, the nucleobase may be a pyrimidine or a derivative thereof.

In some embodiments, the modified nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, are a compound of Formula VII-a:

![Formula VII-a](image)

In some embodiments, the modified nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, are a compound of Formula VII-b:

![Formula VII-b](image)

In some embodiments, the modified nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, are a compound of Formula VII-cl, VII-c2, or VII-c3:

![Formulas VII-cl, VII-c2, VII-c3](image)

In some embodiments, the modified nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, are a compound of Formula XI:
[000205] wherein:
[000206] \( \equiv \) denotes a single or a double bond;
[000207] — denotes an optional single bond;
[000208] \( U \) is \( O, S, -NR^a, \) or \( -CR^b \) - when \( \equiv \) denotes a single bond, or \( U \) is \( -CR^a \) - when \( \equiv \) denotes a double bond;
[000209] \( Z \) is \( H, C_{1-12} \) alkyl, or \( C_{6-20} \) aryl, or \( Z \) is absent when \( \equiv \) denotes a double bond; and
[000210] \( Z \) can be \( -CR^b \) - and form a bond with \( A \);
[000211] \( A \) is \( H, OH, \) sulfate, \( -NH_2, -SH, \) an amino acid, or a peptide comprising 1 to 12 amino acids;
[000212] \( D \) is \( H, OH, -NH_2, -SH, \) an amino acid, a peptide comprising 1 to 12 amino acids, or a group of Formula VI:

[000213] or \( A \) and \( D \) together with the carbon atoms to which they are attached form a 5-membered ring;
[000214] \( X \) is \( O \) or \( S \);
[000215] each of \( Y^1 \) is independently selected from \( -OR^a, -NR^aR^b, \) and \( -SR^a \);
[000216] each of \( Y^2 \) and \( Y^3 \) is independently selected from \( O, -CR^b, -NR^c, \) \( S \) or a linker comprising one or more atoms selected from the group consisting of \( C, O, N, \) and \( S \);
[000217] \( n \) is 0, 1, 2, or 3;
[000218] \( m \) is 0, 1, 2 or 3;
[000219] \( B \) is a nucleobase of Formula IX:
wherein:

- **V** is N or positively charged NR;
- **R^3** is NR^d, -OR, or -SR;
- **R^4** is H or can optionally form a bond with Y^3;
- **R^5** is H, -NR^dR^d, or -OR;
- **R^a and R^b** are each independently H, C_{1-12} alkyl, C_{2-12} alkenyl, C_{2-12} alkynyl, or C_{6-26} aryl;
- **R^c** is H, C_{1-12} alkyl, C_{2-12} alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;
- **R^1 and R^b** are each independently H or a counterion; and
- **-OR^c** is OH at a pH of about 1 or -OR^c is O^- at physiological pH.

In some embodiments, **B** may be:

- **R^3** is -OH, -SH, or
- In some embodiments, **B** may be:
In some embodiments, the modified nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, may be a compound of Formula X:

\[
\text{Formula X}
\]

In some embodiments, the modified nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, may be a compound selected from the group consisting of:

\[
\text{Compound 1}
\]
\[
\text{Compound 2}
\]
\[
\text{Compound 3}
\]
\[
\text{Compound 4}
\]
In some embodiments, the modified nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, may be a compound selected from the group consisting of:
In one embodiment, the present invention provides methods of preparing a nucleic acid sequence comprising at least one nucleotide that disrupts binding of a major groove interacting partner with the nucleic acid sequence, wherein the nucleic acid sequence comprises a compound of Formula XI:

[000236]
wherein:

- $Z$ is $O$ or $S$;
- each of $Y^1$ is independently selected from $-OR^1$, $-NR^1R^b$, and $-SR^1$;
- each of $Y^2$ is independently selected from $O$, $NR^3$, or a linker comprising an atom selected from the group consisting of $C$, $O$, $N$, and $S$;
- each of $Y^3$ is independently selected from $O$ and $S$;
- $Y^4$ is selected from $H$, $-OR^3$, $-SR^3$, and $-NHR^3$;
- $B$ is a nucleobase;
- $R^a$ is $H$, $Ci_{-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl, or $C_{6-20}$ aryl;
- $R^b$ and $R^3$ are each independently $H$ or a counterion; and
- $-Y^3-R^3$ is $OH$ or $SH$ at a pH of about $1$ or $-Y^3-R^3$ is $O^-$ or $S^-$ at physiological pH;
- or $-Y^3-R^3$ is $Ci_{-20}$ alkoxy, $C_{2-20}$-O-alkenyl, or $Ci_{-20}$-O-alkynyl;
- wherein when $B$ is an unmodified nucleobase selected from cytosine, guanine, uracil and adenine, then at least one of $Z$, $Y^1$ or $Y^2$ is not $O$ or $OH$;
- the method comprising reacting a compound of Formula XII:

\[
\begin{align*}
\frac{1}{2}Z &- Y^1 & & \frac{1}{2}Z \\
Y^3 & & & Y^2 \\
R^c & & & R^c \\
\frac{1}{2}Z &- P & & \frac{1}{2}Z \\
Y^3 & & & Y^2 \\
R^c & & & R^c \\
\end{align*}
\]

XII

- with an RNA polymerase, and a cDNA template.
- In some embodiments, the reaction may be repeated from 1 to about 7,000 times.
- In some embodiments, $B$ may be a nucleobase of Formula II-a, II-b, or II-c:
[000253] wherein:
[000254] \( \rightleftharpoons \) denotes a single or double bond;
[000255] X is O or S;
[000256] U and W are each independently C or N;
[000257] V is O, S, C or N;
[000258] wherein when V is C then \( R^1 \) is H, \( C_1-6 \) alkyl, \( C_{1-6} \) alkenyl, \( C_{1-6} \) alkynyl, halo, or -OR; wherein \( C_{1-20} \) alkyl, \( C_{2-20} \) alkenyl, \( C_{2-20} \) alkynyl are each optionally substituted with -OH, -NR,Rb, -SH, -C(0)Rc, -C(0)OR, -NHC(0)Rc, or -NHC(0)ORc;
[000259] and wherein when V is O, S, or N then \( R^1 \) is absent;
[000260] \( R^2 \) is H, -OR, -SR, -NR,Rb, or halo;
[000261] or when V is C then \( R^1 \) and \( R^2 \) together with the carbon atoms to which they are attached can form a 5- or 6-membered ring optionally substituted with 1-4 substituents selected from halo, -OH, -SH, -NR,Rb, \( C_{1-20} \) alkyl, \( C_{2-20} \) alkenyl, \( C_{2-20} \) alkynyl, \( C_{1-20} \) alkoxy, or \( C_{1-20} \) thioalkyl;
[000262] \( R^3 \) is H or \( C_{1-20} \) alkyl;
[000263] \( R^4 \) is H or \( C_{1-20} \) alkyl;
[000264] wherein when \( \rightleftharpoons \) denotes a double bond then \( R^4 \) is absent, or N-R,R taken together, forms a positively charged N substituted with \( C_{1-20} \) alkyl;
[000265] \( R^a \) and \( R^b \) are each independently H, \( C_{1-20} \) alkyl, \( C_{2-20} \) alkenyl, \( C_{2-20} \) alkynyl, or \( C_{6-20} \) aryl; and
[000266] \( R^c \) is H, \( C_{1-20} \) alkyl, \( C_{2-20} \) alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.
[000267] In some embodiments, B may be a nucleobase of Formula II-al, II-a2, II-a3, II-a4, or II-a5:
In some embodiments, the methods further comprise a nucleotide selected from the group consisting of adenosine, cytosine, guanosine, and uracil.

In some embodiments, the nucleobase may be a pyrimidine or derivative thereof.

In a further embodiment, the present invention provides methods of amplifying a nucleic acid sequence comprising at least one nucleotide that disrupts binding of a major groove binding partner with the nucleic acid sequence, the method comprising:

reacting a compound of Formula XII:

\[
\text{II-XII}
\]

Z is O or S;

each of Y^1 is independently selected from -OR^a, -NR^aR^b, and -SR^a;

each of Y^2 is independently selected from O, NR^a, S or a linker comprising an atom selected from the group consisting of C, O, N, and S;

each of Y^3 is independently selected from O and S;

Y^4 is selected from H, -OR^a, -SR^a, and -NHR^a;

B is a nucleobase;

R^a is H, Cl_2-20 alkyl, C_2-20 alkenyl, C_2-20 alkynyl, or C_6-20 aryl;

R^a and R^b are each independently H or a counterion; and

-Y^3-R^c is OH or SH at a pH of about 1 or -Y^3-R^c is O^- or S^- at physiological pH;

or -Y^3-R^c is Cl_2-20 alkoxy, C_2-20-O-alkenyl, or Cl_2-20-O-alkynyl;

wherein when B is an unmodified nucleobase selected from cytosine, guanine, uracil and adenine, then at least one of Z, Y^1 or Y^2 is not O or OH;

with a primer, a cDNA template, and an RNA polymerase.

In some embodiments, B may be a nucleobase of Formula II-a, II-b, or II-c:
wherein:

- denotes a single or double bond;
- \( X \) is \( O \) or \( S \);  
- \( U \) and \( W \) are each independently \( C \) or \( N \);  
- \( V \) is \( O \), \( S \), \( C \), or \( N \);  
- when \( V \) is \( C \) then \( R^1 \) is \( H \), \( \text{Ci-}2\text{o} \) alkyl, \( \text{Ci-}2\text{o} \) alkenyl, \( \text{Ci-}2\text{o} \) alkynyl, halo, or -OR;  
- when \( V \) is \( C \) then \( R^1 \) and \( R^2 \) together with the carbon atoms to which they are attached can form a 5- or 6-membered ring optionally substituted with 1-4 substituents selected from halo, -OH, -SH, -NR\( a \)R, -C(\( 0 \)R, or -NHC(\( 0 \)OR;  
- and when \( V \) is \( O \), \( S \), or \( N \) then \( R^1 \) is absent;  
- \( R^2 \) is \( H \), -OR, -SR, -NR\( a \)R, or halo;  
- or when \( V \) is \( C \) then \( R^1 \) and \( R^2 \) together with the carbon atoms to which they are attached can form a 5- or 6-membered ring optionally substituted with 1-4 substituents selected from halo, -OH, -SH, -NR\( a \)R, -C(\( 0 \)alkyl, \( C_2\text{-}2\text{o} \) alkynyl, \( C_2\text{-}2\text{o} \) alkynyl, \( C_2\text{-}2\text{o} \) alkyl, or \( C_2\text{-}2\text{o} \) thioalkyl;  
- \( R^3 \) is \( H \) or \( \text{Ci-}2\text{o} \) alkyl;  
- \( R^4 \) is \( H \) or \( \text{Ci-}2\text{o} \) alkyl;  
- wherein \( \Rightarrow \) denotes a double bond then \( R^4 \) is absent, or \( N\text{-}R^4 \), taken together, forms a positively charged \( N \) substituted with \( \text{Ci-}2\text{o} \) alkyl;  
- \( R^a \) and \( R^b \) are each independently \( H \), \( \text{Ci-}2\text{o} \) alkyl, \( C_2\text{-}2\text{o} \) alkenyl, \( C_2\text{-}2\text{o} \) alkynyl, or \( C_6\text{-}2\text{o} \) aryl;  
- and \( R^c \) is \( H \), \( \text{Ci-}2\text{o} \) alkyl, \( C_2\text{-}2\text{o} \) alkenyl, \( \text{phenyl} \), \( \text{benzyl} \), a polyethylene glycol group, or an amino-polyethylene glycol group.  

In some embodiments, \( B \) may be a nucleobase of Formula II-a1, II-a2, II-a3, II-a4, or II-a5:
In some embodiments, the methods further comprise a nucleotide selected from the group consisting of adenosine, cytosine, guanosine, and uracil. In some embodiments, the nucleobase may be a pyrimidine or derivative thereof.

In some embodiments, the nucleic acid is translatable.

Other components of nucleic acid are optional, and are beneficial in some embodiments. For example, a 5' untranslated region (UTR) and/or a 3'UTR are provided, wherein either or both may independently contain one or more different nucleoside modifications. In such embodiments, nucleoside modifications may also be present in the translatable region. Also provided are nucleic acids containing a Kozak sequence.

Additionally, provided are nucleic acids containing one or more intronic nucleotide sequences capable of being excised from the nucleic acid.

Further, provided are nucleic acids containing an internal ribosome entry site (IRES). An IRES may act as the sole ribosome binding site, or may serve as one of multiple ribosome binding sites of an mRNA. An mRNA containing more than one functional ribosome binding site may encode several peptides or polypeptides that are translated independently by the ribosomes ("multicistrionic mRNA"). When nucleic acids are provided with an IRES, further optionally provided is a second translatable region. Examples of IRES sequences that can be used according to the invention include without limitation, those from picornaviruses (e.g. FMDV), pest viruses (CFFV), polio viruses (PV), encephalomyocarditis viruses (ECMV), foot-and-mouth disease viruses (FMDV), hepatitis C viruses (HCV), classical swine fever viruses (CSFV), murine leukemia virus (MLV), simian immune deficiency viruses (SIV) or cricket paralysis viruses (CrPV).

Modified Nucleotide with a Linker and a Payload

The nucleobase of the nucleotide, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, can be covalently linked at any chemically appropriate position to a payload, e.g., detectable agent or therapeutic agent. For example, the nucleobase can be deaza-adenosine or deaza-guanosine and the linker can be attached at the C-7 or C-8 positions of the deaza-adenosine or deaza-guanosine. In other embodiments, the nucleobase can be cytosine or uracil and the linker can be attached to the N-3 or C-5 positions of cytosine or uracil. Scheme 1, below, depicts a modified nucleotide wherein the nucleobase, adenine, is attached to a linker at the C-7 carbon of 7-deaza adenine. In addition, Scheme 1 depicts the modified nucleotide with the linker and payload, e.g., a
detectable agent, incorporated onto the 3’ end of the mRNA. Disulfide cleavage and a 1,2-addition of the thiol group onto the propargyl ester releases the detectable agent. The remaining structure (depicted, for example, as pApC5Parg in Scheme 1) is the inhibitor. The structure of the modified nucleotide is important as the tethered inhibitor sterically interferes with the ability of the polymerase to incorporate a second base. Thus, it is critical that the tether be long enough to affect the incorporation of a second base and that the inhibitor be in a stereochemical orientation to inhibit or prohibits second and follow on nucleotides into the growing polynucleotide strand.
Scheme 1

A Capless pCp5 Parg

 incorporation

Cleavage of S-S bond

Linker
The term "linker" as used herein refers to a group of atoms, e.g., 10-1,000 atoms, and can be comprised of the atoms or groups such as, but not limited to, carbon, amino, alkyamino, oxygen, sulfur, sulfoxide, sulfanyl, carbonyl, and imine. The linker can be attached to a modified nucleoside or nucleotide on the nucleobase or sugar moiety at a first end, and to a payload, e.g., detectable or therapeutic agent, at a second end. The linker may be of sufficient length as to not interfere with incorporation into a nucleic acid sequence.

Examples of chemical groups that can be incorporated into the linker include, but are not limited to, an alkyl, an alkene, an alkyne, an amido, an ether, a thioether or an ester group. The linker chain can also comprise part of a saturated, unsaturated or aromatic ring, including polycyclic and heteroaromatic rings wherein the heteroaromatic ring may be an aryl group containing one to four heteroatoms, N, O or S. Specific examples of linkers include, but are not limited to, unsaturated alkanes, polyethylene glycols, and dextran polymers.

For example, the linker can include, but is not limited to, ethylene or propylene glycol monomeric units, e.g., diethylene glycol, dipropylene glycol, triethylene glycol, tripropylene glycol, tetraethylene glycol, or tetraethylene glycol. In some embodiments, the linker can include, but is not limited to, a divalent alkyl, alkenyl, and/or alkynyl moiety. The linker can include an ester, amide, or ether moiety.

Other examples include, but are not limited to, cleavable moieties within the linker, such as, for example, a disulfide bond (-S-S-) or an azo bond (-N=N-), which can be cleaved using a reducing agent or photolysis. When a cleavable bond which has been incorporated into the linker and attached to a modified nucleotide, is cleaved, a short "scar" or chemical modification on the nucleotide may result. For example, after cleaving, the resulting scar on a nucleotide base, which formed part of the modified nucleotide, and is incorporated into a polynucleotide strand, is unreactive and does not need to be chemically neutralized. This increases the ease with which a subsequent nucleotide can be incorporated during sequencing of a nucleic acid polymer template. For example, conditions include the use of tris(2-carboxyethyl)phosphine (TCEP), dithiothreitol (DTT) and/or other reducing agents for cleavage of a disulfide bond. A selectively severable bond that includes an amido bond can be cleaved for example by the use of TCEP or other reducing agents, and/or photolysis. A selectively severable bond that includes an ester bond can be cleaved for example by acidic or basic hydrolysis.

Payload
The methods and compositions described herein are useful for delivering a payload to a biological target. The payload can be used, e.g., for labeling (e.g., a detectable agent such as a fluorophore), or for therapeutic purposes (e.g., a cytotoxin or other therapeutic agent).

**Therapeutic Agents**

In some embodiments the payload may be a therapeutic agent such as a cytotoxin, radioactive ion, chemotherapeutic, or other therapeutic agent. A cytotoxin or cytotoxic agent includes any agent that may be detrimental to cells. Examples include, but are not limited to, taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracing dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, puromycin, maytansinoids, e.g., maytansinol (see U.S. Pat. No. 5,208,020 incorporated herein in its entirety), CC-1065 (see U.S. Pat. Nos. 5,475,092, 5,585,499, 5,846,545; all of which are incorporated herein by reference) and analogs or homologs thereof. Radioactive ions include, but are not limited to iodine (e.g., iodine 125 or iodine 131), strontium 89, phosphorous, palladium, cesium, iridium, phosphate, cobalt, yttrium 90, Samarium 153 and praseodymium. Other therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil carabazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, CC-1065, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine, vinblastine, taxol and maytansinoids).

**Detectable Agents**

Examples of detectable substances include, but are not limited to, various organic small molecules, inorganic compounds, nanoparticles, enzymes or enzyme substrates, fluorescent materials, luminescent materials, bioluminescent materials, chemiluminescent materials, radioactive materials, and contrast agents. Such optically-detectable labels include for example, without limitation, 4-acetamido-4'-isothiocyanoatostilbene-2,2'disulfonic acid; acridine and derivatives: acridine, acridine isothiocyanate; 5-(2'-aminoethyl)aminonaphthalene-l-sulfonic acid (EDANS); 4-amino-N-[3-vinylsulfonyl]phenyl]napthalimide-3,5 disulfonate; N-(4-anilino-1-
naphthyl)maleimide; anthranilamide; BODIPY; Brilliant Yellow; coumarin and derivatives; coumarin, 7-amino-4-methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcoumarin (Coumaran 151); cyanine dyes; cyanosine; 4',6-diaminindino-2-phenylindole (DAPI); 5'5''-dibromopyrogallol-sulfonaphthalein (Bromopyrogallol Red); 7-diethylamino-3-(4' isothiocyanatophenyl)-4-methylcoumarin; diethylenetriamine pentaacetate; 4,4'-diisothiocyanatodihydro-stilbene-2,2 '-disulfonic acid; 4,4'-diisothiocyanatostilbene-2,2 '-disulfonic acid; 5-[dimethylamino]-naphthalene-1-sulfonate (DNS, dansylchloride); 4- dimethylaminophenylazophenyl-4'-isothiocyanate (DABITC); eosin and derivatives; eosin, eosin isothiocyanate, erythrosin and derivatives; erythrosin B, erythrosin, isothiocyanate; ethidium; fluorescein and derivatives; 5-carboxyfluorescein (FAM), 5-(4,6-dichlorotriazin-2-yl)aminofluorescein (DTAF), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein, fluorescein, fluorescein isothiocyanate, QFITC, (XRITC); fluorescamine; IR144; IR1446; Malachite Green isothiocyanate; 4-methylumbelliferoneortho cresphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-phthalaldehyde; pyrene and derivatives: pyrene, pyrene butyrate, succinimidyl 1-pyrene; butyrate quantum dots; Reactive Red 4 (CibacronTM Brilliant Red 3B-A) rhodamine and derivatives: 6-carboxy-X-rhodamine (ROX), 6-carboxyrhodamine (R6G), lissamine rhodamine B sulfonyl chloride rhodamine (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, sulforhodamine B, sulforhodamine 101, sulfonyl chloride derivative of sulforhodamine 101 (Texas Red); N,N,N',N'tetramethyl-6-carboxyrhodamine (TAMRA) tetramethyl rhodamine; tetramethyl rhodamine isothiocyanate (TRITC); riboflavin; rosolic acid; terbium chelate derivatives; Cyanine-3 (Cy3); Cyanine-5 (Cy5); Cyanine-5.5 (Cy5.5), Cyanine-7 (Cy7); IRD 700; IRD 800; Alexa 647; La Jolta Blue; phthalo cyanine; and naphthalo cyanine. In some embodiments, the detectable label may be a fluorescent dye, such as Cy5 and Cy3.

[000314] An example of luminescent material includes, but is not limited to, luminol; examples of bioluminescent materials includes, but is not limited to, luciferase, luciferin, and aeruorin.

[000315] Examples of suitable radioactive material include, but is not limited to, $^{18}$F, $^{67}$Ga, $^{81m}$Kr, $^{82}$Rb, $^{111}$In, $^{123}$I, $^{133}$Xe, $^{204}$Tl, $^{125}$I, $^{35}$S, $^{14}$C, or $^3$H, $^{99m}$Tc (e.g., as pertechnetate (technetate(VII)), Tc(V) either directly or indirectly, or other radioisotope detectable by direct counting of radioemmission or by scintillation counting.

[000316] In addition, contrast agents, e.g., contrast agents for MRI or NMR, for X-ray CT, Raman imaging, optical coherence tomgraphy, absorption imaging, ultrasound imaging, or thermal imaging can be used. Exemplary contrast agents include, but are not limited to, gold (e.g., gold
nanoparticles), gadolinium (e.g., chelated Gd), iron oxides (e.g., superparamagnetic iron oxide (SPIO), monocrystalline iron oxide nanoparticles (MIONs), and ultrasmall superparamagnetic iron oxide (USPIO)), manganese chelates (e.g., Mn-DPDP), barium sulfate, iodinated contrast media (iohexol), microbubbles, or perfluorocarbons.

In some embodiments, the detectable agent may be a non-detectable pre-cursor that becomes detectable upon activation. Examples include, but are not limited to, fluorogenic tetrazine-fluorophore constructs (e.g., tetrazine-BODIPY FL, tetrazine-Oregon Green 488, or tetrazine-BODIPY TMR-X) or enzyme activatable fluorogenic agents (e.g., PROSENSE (VisEn Medical)).

When the compounds are enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, the enzymatic label may be detected by the determination of the conversion of an appropriate substrate to a product. In vitro assays in which the enzyme labeled compositions can be used include, but are not limited to, enzyme linked immunosorbent assays (ELISAs), immunoprecipitations, immunofluorescence, enzyme immunoassay (EIA), radioimmunoassay (RIA), and Western blot analysis.

Labels, other than those described herein, are contemplated by the present disclosure, including, but not limited to, other optically-detectable labels. Labels can be attached to the modified nucleotide of the present disclosure at any position using standard chemistries such that the label can be removed from the incorporated base upon cleavage of the cleavable linker.

Cell Penetrating Payloads

In some embodiments, the modified nucleotides and modified nucleic acid molecules, which are incorporated into a nucleic acid, e.g., RNA or mRNA, can also include a payload that can be a cell penetrating moiety or agent that enhances intracellular delivery of the compositions. For example, the compositions can include, but are not limited to, a cell-penetrating peptide sequence that facilitates delivery to the intracellular space, e.g., HIV-derived TAT peptide, penetratins, transportans, or hCT derived cell-penetrating peptides, see, e.g., Caron et al, (2001) Mol Ther. 3(3):3 10-8; Langel, Cell-Penetrating Peptides: Processes and Applications (CRC Press, Boca Raton FL 2002); El-Andaloussi et al, (2005) Curr Pharm Des. 11(28):3597-61; and Deshayes et al, (2005) Cell Mol Life Sci. 62(16): 1839-49; all of which are incorporated herein by reference. The compositions can also be formulated to include a cell penetrating agent, e.g., liposomes, which enhance delivery of the compositions to the intracellular space.


**Biological Targets**

[000322] The modified nucleotides and modified nucleic acid molecules described herein, which are incorporated into a nucleic acid, *e.g.*, RNA or mRNA, can be used to deliver a payload to any biological target for which a specific ligand exists or can be generated. The ligand can bind to the biological target either covalently or non-covalently.

[000323] Examples of biological targets include, but are not limited to, biopolymers, *e.g.*, antibodies, nucleic acids such as RNA and DNA, proteins, enzymes; examples of proteins include, but are not limited to, enzymes, receptors, and ion channels. In some embodiments the target may be a tissue- or a cell-type specific marker, *e.g.*, a protein that is expressed specifically on a selected tissue or cell type. In some embodiments, the target may be a receptor, such as, but not limited to, plasma membrane receptors and nuclear receptors; more specific examples include, but are not limited to, G-protein-coupled receptors, cell pore proteins, transporter proteins, surface-expressed antibodies, HLA proteins, MHC proteins and growth factor receptors.

**Terminal Architecture Modifications: 5′-Capping**

[000324] Endogenous eukaryotic cellular messenger RNA (mRNA) molecules contain a 5′-cap structure on the 5′-end of a mature mRNA molecule. The 5′-cap contains a 5′-5′-triphosphate linkage between the 5′-most nucleotide and guanine nucleotide. The conjugated guanine nucleotide is methylated at the N7 position.

[000325] Additional modifications include methylation of the ultimate and penultimate most 5′-nucleotides on the 2′-hydroxyl group. The 5′-cap structure is responsible for binding the mRNA Cap Binding Protein (CBP), which is responsibility for mRNA stability in the cell and translation competency.

[000326] Multiple distinct 5′-cap structures can be used to generate the 5′-cap of a synthetic mRNA molecule. Many chemical cap analogs are used to co-transcriptionally cap a synthetic mRNA molecule. For example, the Anti-Reverse Cap Analog (ARCA) cap contains a 5′-5′-triphosphate guanine-guanine linkage where one guanine contains an N7 methyl group as well as a 3′-0-methyl group. While chemical cap analogs allow for the concomitant capping of an RNA molecule, up 20% of transcripts remain uncapped and the synthetic cap analog is not identical to an endogenous 5′-cap structure of an authentic cellular mRNA. This may lead to reduced translationally-competency and reduced cellular stability.
[000327] Synthetic mRNA molecules may also be capped post-transcriptionally using enzymes responsible for generating a more authentic 5\'-cap structure. As used herein the phrase "more authentic" refers to a feature that closely mirrors or mimics, either structurally or functionally an endogenous or wild type feature. More authentic 5\' cap structures of the present invention are those which, among other things, have enhanced binding of cap binding proteins, increased half life, reduced susceptibility to 5\' endonucleases and/or reduced 5\' decapping. For example, recombinant Vaccinia Virus Capping Enzyme and recombinant 2'-0-methyltransferase enzyme can create a canonical 5\'-5\' triphosphate linkage between the 5\'-most nucleotide of an mRNA and a guanine nucleotide where the guanine contains an N7 methylation and the ultimate 5\'-nucleotide contains a 2'-0-methyl generating the Capl structure. This results in a cap with higher translational-competency and cellular stability and reduced activation of cellular pro-inflammatory cytokines. Because the synthetic mRNA is capped post-transcriptionally, nearly 100% of the mRNA molecules are capped in contrast to -80% of synthetic mRNAs containing a chemical cap analog.

Cell-Penetrating Polypeptides

[000328] The modified nucleic acids disclosed herein may encode a cell-penetrating polypeptide. As used herein, "cell-penetrating polypeptide" refers to a polypeptide which may facilitate the cellular uptake of molecules. "Cell-penetrating peptide" refers to a peptide which may facilitate the cellular uptake of molecules. It is known in the art that "CPP" refers to cell-penetration polypeptides and cell-penetrating peptides. When used herein, it will be clarified as to which of either cell-penetrating polypeptides or cell-penetrating peptides the abbreviation CPP refers to.

[000329] A cell-penetrating polypeptide of the present invention may contain one or more detectable labels. The polypeptides may be partially labeled or completely labeled throughout. The modified nucleic acid may encode the detectable label completely, partially or not at all. The cell-penetrating peptide may also include a signal sequence. As used herein, a "signal sequence" refers to a sequence of amino acid residues bound at the amino terminus of a nascent protein during protein translation. The signal sequence may be used to signal the secretion of the cell-penetrating polypeptide.

[000330] In one embodiment, a modified nucleic acid may encode a cell-penetrating polypeptide. When the modified nucleic acid is introduced into a cell, the modified nucleic acid may provoke an innate immune response. The response may result in an increase or a decrease in the response as compared to the response when an unmodified nucleic acid is introduced into the cell. In a preferred
embodiment the modified nucleic acid is introduced into an animal cell such as, but not limited to, a mammalian cell.

[000331] The modified nucleic acid may encode a cell-penetrating polypeptide that may comprise a fusion protein. The fusion protein may be created by operably linking a charged protein to a therapeutic protein. As used herein, "operably linked" refers to the therapeutic protein and the charged protein being connected in such a way to permit the expression of the complex when introduced into the cell. As used herein, "charged protein" refers to a protein that carries a positive, negative or overall neutral electrical charge. Preferably, the therapeutic protein may be covalently linked to the charged protein in the formation of the fusion protein. The ratio of surface charge to amino acids may be approximately 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9.

[000332] The cell-penetrating polypeptide encoded by the modified nucleic acids may form a complex. The complex may comprise a charged protein linked, e.g. covalently linked, to the cell-penetrating polypeptide. The complex may further include a therapeutic protein operably linked to the charged protein. "Therapeutic protein" refers to a protein that, when administered to a cell has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect. The therapeutic protein may further include a surface charge.

[000333] The modified nucleic acid may encode a cell-penetrating polypeptide that may also comprise a therapeutic protein. The therapeutic protein may further comprise a surface charge such as a positive or negative charge. As used herein, "surface charge" refers to the electric charge present at an interface. Additionally, the cell-penetrating peptide may comprise a supercharged polypeptide. A "supercharged polypeptide" refers to a polypeptide which can result in an increase or decrease in the overall surface charge on the polypeptide when introduced into the cell. The supercharged polypeptide may be positively or negatively charged or have an overall neutral charge. The supercharged polypeptide may be charged in order to aid in the delivery of the modified nucleic acid into the cell.

[000334] In one embodiment, the cell-penetrating polypeptide may comprise a first domain and a second domain. The first domain may comprise a supercharged polypeptide. The second domain may comprise a protein-binding partner. As used herein, "protein-binding partner" includes, but are not limited to, antibodies and functional fragments thereof, scaffold proteins, or peptides. The cell-penetrating polypeptide may further comprise an intracellular binding partner for the protein-binding partner. The cell-penetrating polypeptide may be capable of being secreted from a cell where the
modified nucleic acid may be introduced. The cell-penetrating polypeptide may also be capable of penetrating the first cell.

[000335] In a further embodiment, the cell-penetrating polypeptide is capable of penetrating a second cell. The second cell may from the same area as the first cell, or it may be from a different area. The area may include, but is not limited to, tissues and organs. The second cell may also be proximal or distal to the first cell.

[000336] In one embodiment, the modified nucleic acid may encode a cell-penetrating polypeptide which may comprise a protein-binding partner. The protein binding partner may include, but is not limited to, an antibody, a supercharged antibody or a functional fragment. The modified nucleic acid may be introduced into the cell where a cell-penetrating polypeptide comprising the protein-binding partner is introduced. In a further embodiment, the cell-penetrating polypeptide may penetrate the first cell. The cell-penetrating polypeptide comprising the protein-binding partner may be capable of penetrating a second cell. The second cell may comprise an intracellular binding partner for the protein-binding partner. In a preferred embodiment, the second cell comprises an intracellular binding partner for a supercharged antibody. In a preferred embodiment, the second cell comprises an intracellular binding partner for a functional fragment.

[000337] In one embodiment, a formulation comprises the modified nucleic acid encoding a cell-penetrating polypeptide and a delivery agent. The delivery agent may comprise a cationic material or it may comprise a cell-penetrating polypeptide.

[000338] In one embodiment, a protein-binding partner may be delivered into a target cell which may comprise an intracellular binding partner. A modified nucleic acid which encodes a cell-penetrating polypeptide comprising a protein-binding partner may be introduced into a first cell when the surrounding conditions would facilitate the production of the cell-penetrating polypeptide. The resulting conditions may cause secretion from the first cell of the cell-penetrating polypeptide and resulting in the penetration of the target cell. The introduction of the modified nucleic acid encoding a cell-penetrating polypeptide to the first cell may provoke a change in the innate immune response to the first cell. This change may be a reduction in the innate response as compared to an introduction with an unmodified nucleic acid encoding the same or similar cell-penetrating polypeptide. Preferably, the first cell and the target cell are the same cell, or the first and second cells are co-located in a tissue of a subject. The first cell may also be located proximally or distally to the target cell.
In a further embodiment, the modified nucleic acid is retained in an extracellular depot before the introduction of the modified nucleic acid into the first cell. The extracellular depot may be located adjacent to the first cell. As used herein, "extracellular depot" refers to a housing which may contain the modified nucleic acid prior to administration in the interstitial space. This housing may be selected from, but not limited to, rigid, semi-rigid, flexible, or gel-like materials.

Synthesis of Modified Nucleotides

The modified nucleosides and nucleotides, which may be incorporated into a nucleic acid, e.g., RNA or mRNA, disclosed herein can be prepared from readily available starting materials using the following general methods and procedures. It is understood that where typical or preferred process conditions (i.e., reaction temperatures, times, mole ratios of reactants, solvents, pressures, etc.) are given; other process conditions can also be used unless otherwise stated. Optimum reaction conditions may vary with the particular reactants or solvent used, but such conditions can be determined by one skilled in the art by routine optimization procedures.

The processes described herein can be monitored according to any suitable method known in the art. For example, product formation can be monitored by spectroscopic means, such as nuclear magnetic resonance spectroscopy (e.g., H or 13C) infrared spectroscopy, spectrophotometry (e.g., UV-visible), mass spectrometry, or by chromatography such as high performance liquid chromatography (HPLC) or thin layer chromatography.

Preparation of modified nucleosides and nucleotides can involve the protection and deprotection of various chemical groups. The need for protection and deprotection, and the selection of appropriate protecting groups can be readily determined by one skilled in the art. The chemistry of protecting groups can be found, for example, in Greene, et al, Protective Groups in Organic Synthesis, 2d. Ed., Wiley & Sons, 1991, which is incorporated herein by reference in its entirety.

The reactions of the processes described herein can be carried out in suitable solvents, which can be readily selected by one of skill in the art of organic synthesis. Suitable solvents can be substantially nonreactive with the starting materials (reactants), the intermediates, or products at the temperatures at which the reactions are carried out, i.e., temperatures which can range from the solvent's freezing temperature to the solvent's boiling temperature. A given reaction can be carried out in one solvent or a mixture of more than one solvent. Depending on the particular reaction step, suitable solvents for a particular reaction step can be selected.
Resolution of racemic mixtures of modified nucleosides and nucleotides can be carried out by any of numerous methods known in the art. An example method includes, but is not limited to, fractional recrystallization using a "chiral resolving acid" which is an optically active, salt-forming organic acid. Suitable resolving agents for fractional recrystallization methods are, for example, optically active acids, such as the D and L forms of tartaric acid, diacetyltartaric acid, dibenzoyltartaric acid, mandelic acid, malic acid, lactic acid or the various optically active camphorsulfonic acids. Resolution of racemic mixtures can also be carried out by elution on a column packed with an optically active resolving agent (e.g., dinitrobenzoylphenylglycine). Suitable elution solvent composition can be determined by one skilled in the art.

Exemplary syntheses of modified nucleotides, which are incorporated into a nucleic acid, e.g., RNA or mRNA, are provided below in Schemes 2 through 12.
Scheme 3

\[
\text{RBr/Heat} \quad \text{R = alkyl, alkenyl, allyl, and benzyl}
\]

Scheme 4

\[
1) \text{POCl}_3 \\
2) \text{Pyrophosphate}
\]
Scheme 5
Scheme 6

Scheme 7
Scheme 8

1) POCl₃
2) Pyrophosphate

RBr/Heat
R = alkyl, alkenyl, allyl, and benzyl

1) POCl₃
2) Pyrophosphate
Scheme 9

RBr/Heat
R = alkyl, alkenyl, allyl, and benzyl

1) POCl₃
2) Pyrophosphate
Scheme 10

CH$_3$NH$_2$/Heat

1) POCl$_3$
2) Pyrophosphate
Scheme 11

1. DCC, Morpholine

2. Enzymatic Hydrolysis

3. Uracil

4. (1) OsO₄, (2) Acetone, TsOH

5. \((\text{EtO})_2\text{POCH}_2\text{OTs}\)

6. TMSiI

7. DCC, Morpholine

8. Pyrophosphate
Terminal Architecture Modifications: Poly-A tails

[000346] During RNA processing, a long chain of adenine nucleotides (poly-A tail) is normally added to a messenger RNA (mRNA) molecules to increase the stability of the molecule. Immediately after transcription, the 3' end of the transcript is cleaved to free a 3' hydroxyl. Then poly-A polymerase adds a chain of adenine nucleotides to the RNA. The process, called polyadenylation, adds a poly-A tail that is between 100 and 250 residues long.
It has been discovered that unique poly-A tail lengths provide certain advantages to the modified RNAs of the present invention.

Generally, the length of a poly-A tail of the present invention is greater than 30 nucleotides in length. In another embodiment, the poly-A tail is greater than 35 nucleotides in length. In another embodiment, the length is at least 40 nucleotides. In another embodiment, the length is at least 45 nucleotides. In another embodiment, the length is at least 55 nucleotides. In another embodiment, the length is at least 60 nucleotides. In another embodiment, the length is at least 60 nucleotides. In another embodiment, the length is at least 80 nucleotides. In another embodiment, the length is at least 90 nucleotides. In another embodiment, the length is at least 100 nucleotides. In another embodiment, the length is at least 120 nucleotides. In another embodiment, the length is at least 140 nucleotides. In another embodiment, the length is at least 160 nucleotides. In another embodiment, the length is at least 180 nucleotides. In another embodiment, the length is at least 200 nucleotides. In another embodiment, the length is at least 250 nucleotides. In another embodiment, the length is at least 300 nucleotides. In another embodiment, the length is at least 350 nucleotides. In another embodiment, the length is at least 400 nucleotides. In another embodiment, the length is at least 450 nucleotides. In another embodiment, the length is at least 500 nucleotides. In another embodiment, the length is at least 600 nucleotides. In another embodiment, the length is at least 700 nucleotides. In another embodiment, the length is at least 800 nucleotides. In another embodiment, the length is at least 900 nucleotides. In another embodiment, the length is at least 1000 nucleotides.

In one embodiment, the poly-A tail is designed relative to the length of the overall modified RNA molecule. This design may be based on the length of the coding region of the modified RNA, the length of a particular feature or region of the modified RNA (such as the mRNA), or based on the length of the ultimate product expressed from the modified RNA.

In this context the poly-A tail may be 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100% greater in length than the modified RNA or feature thereof. The poly-A tail may also be designed as a fraction of the modified RNA to which it belongs. In this context, the poly-A tail may be 10, 20, 30, 40, 50, 60, 70, 80, or 90% or more of the total length of the construct or the total length of the construct minus the poly-A tail.

Length

Generally, the shortest length of a modified mRNA, herein "mmRNA," of the present disclosure can be the length of an mRNA sequence that may be sufficient to encode for a dipeptide.
In another embodiment, the length of the mRNA sequence may be sufficient to encode for a tripeptide. In another embodiment, the length of an mRNA sequence may be sufficient to encode for a tetrapeptide. In another embodiment, the length of an mRNA sequence may be sufficient to encode for a pentapeptide. In another embodiment, the length of an mRNA sequence may be sufficient to encode for a hexapeptide. In another embodiment, the length of an mRNA sequence may be sufficient to encode for a heptapeptide. In another embodiment, the length of an mRNA sequence may be sufficient to encode for an octapeptide. In another embodiment, the length of an mRNA sequence may be sufficient to encode for a nonapeptide. In another embodiment, the length of an mRNA sequence may be sufficient to encode for a decapptide. Examples of dipeptides that the modified nucleic acid molecule sequences can encode for include, but are not limited to, carnosine and anserine.

[000352] Generally, the length of a modified mRNA of the present invention is greater than 30 nucleotides in length. In another embodiment, the RNA molecule is greater than 35 nucleotides in length. In another embodiment, the length is at least 40 nucleotides. In another embodiment, the length is at least 45 nucleotides. In another embodiment, the length is at least 55 nucleotides. In another embodiment, the length is at least 60 nucleotides. In another embodiment, the length is at least 80 nucleotides. In another embodiment, the length is at least 90 nucleotides. In another embodiment, the length is at least 100 nucleotides. In another embodiment, the length is at least 120 nucleotides. In another embodiment, the length is at least 140 nucleotides. In another embodiment, the length is at least 160 nucleotides. In another embodiment, the length is at least 180 nucleotides. In another embodiment, the length is at least 200 nucleotides. In another embodiment, the length is at least 250 nucleotides. In another embodiment, the length is at least 300 nucleotides. In another embodiment, the length is at least 350 nucleotides. In another embodiment, the length is at least 400 nucleotides. In another embodiment, the length is at least 450 nucleotides. In another embodiment, the length is at least 500 nucleotides. In another embodiment, the length is at least 600 nucleotides. In another embodiment, the length is at least 700 nucleotides. In another embodiment, the length is at least 800 nucleotides. In another embodiment, the length is at least 900 nucleotides. In another embodiment, the length is at least 1000 nucleotides. In another embodiment, the length is at least 1100 nucleotides. In another embodiment, the length is at least 1200 nucleotides. In another embodiment, the length is at least 1300 nucleotides. In another embodiment, the length is at least 1400 nucleotides. In another embodiment, the length is at least 1500 nucleotides. In another embodiment, the length is at least
1600 nucleotides. In another embodiment, the length is at least 1800 nucleotides. In another embodiment, the length is at least 2000 nucleotides. In another embodiment, the length is at least 2500 nucleotides. In another embodiment, the length is at least 3000 nucleotides. In another embodiment, the length is at least 4000 nucleotides. In another embodiment, the length is at least 5000 nucleotides, or greater than 5000 nucleotides. In another embodiment, the length is at least 5000 nucleotides, or greater than 6000 nucleotides. In another embodiment, the length is at least 7000 nucleotides, or greater than 7000 nucleotides. In another embodiment, the length is at least 8000 nucleotides, or greater than 8000 nucleotides. In another embodiment, the length is at least 9000 nucleotides, or greater than 9000 nucleotides. In another embodiment, the length is at least 10,000 nucleotides, or greater than 10,000 nucleotides.

**Use of modified RNAs**

Prevention or reduction of innate cellular immune response activation

[000353] The term "innate immune response" includes a cellular response to exogenous single stranded nucleic acids, generally of viral or bacterial origin, which involves the induction of cytokine expression and release, particularly the interferons, and cell death. Protein synthesis is also reduced during the innate cellular immune response. While it is advantageous to eliminate the innate immune response in a cell, the invention provides modified mRNAs that substantially reduce the immune response, including interferon signaling, without entirely eliminating such a response. In some embodiments, the immune response is reduced by 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 99.9%, or greater than 99.9% as compared to the immune response induced by a corresponding unmodified nucleic acid. Such a reduction can be measured by expression or activity level of Type 1 interferons or the expression of interferon-regulated genes such as the toll-like receptors (e.g., TLR7 and TLR8). Reduction of innate immune response can also be measured by decreased cell death following one or more administrations of modified RNAs to a cell population; e.g., cell death is 10%, 25%, 50%, 75%, 85%, 90%, 95%, or over 95% less than the cell death frequency observed with a corresponding unmodified nucleic acid. Moreover, cell death may affect fewer than 50%, 40%, 30%, 20%, 10%, 5%, 1%, 0.1%, 0.01% or fewer than 0.01% of cells contacted with the modified nucleic acids.

[000354] The invention provides for the repeated introduction (e.g., transfection) of modified nucleic acids into a target cell population, e.g., *in vitro*, *ex vivo*, or *in vivo*. The step of contacting the cell population may be repeated one or more times (such as two, three, four, five or more than five
times). In some embodiments, the step of contacting the cell population with the modified nucleic acids is repeated a number of times sufficient such that a predetermined efficiency of protein translation in the cell population is achieved. Given the reduced cytotoxicity of the target cell population provided by the nucleic acid modifications, such repeated transfections are achievable in a diverse array of cell types.

**Major Groove Interacting Partners**

[000355] As described herein, the phrase "major groove interacting partner" refers to RNA recognition receptors that detect and respond to RNA ligands through interactions, e.g. binding, with the major groove face of a nucleotide or nucleic acid. As such, RNA ligands comprising modified nucleotides or nucleic acids such as the modified RNAs as described herein decrease interactions with major groove binding partners, and therefore decrease an innate immune response.

[000356] Example major groove interacting, e.g. binding, partners include, but are not limited to the following nuclease and helicases. Within membranes, TLRs (Toll-like Receptors) 3, 7, and 8 can respond to single- and double-stranded RNAs. Within the cytoplasm, members of the superfamily 2 class of DEX(D/H) helicases and ATPases can sense RNAs to initiate antiviral responses. These helicases include the RIG-I (retinoic acid-inducible gene 1) and MDA5 (melanoma differentiation-associated gene 5). Other examples include laboratory of genetics and physiology 2 (LGP2), HIN-200 domain containing proteins, or Helicase-domain containing proteins.

**Polypeptide variants**

[000357] Provided are nucleic acids that encode variant polypeptides, which have a certain identity with a reference polypeptide sequence. The term "identity" as known in the art, refers to a relationship between the sequences of two or more peptides, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between peptides, as determined by the number of matches between strings of two or more amino acid residues.

"Identity" measures the percent of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program (i.e., "algorithms"). Identity of related peptides can be readily calculated by known methods. Such methods include, but are not limited to, those described in Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1. Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994;

[000358] In some embodiments, the polypeptide variant has the same or a similar activity as the reference polypeptide. Alternatively, the variant has an altered activity (e.g., increased or decreased) relative to a reference polypeptide. Generally, variants of a particular polynucleotide or polypeptide of the invention will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to that particular reference polynucleotide or polypeptide as determined by sequence alignment programs and parameters described herein and known to those skilled in the art.

[000359] As recognized by those skilled in the art, protein fragments, functional protein domains, and homologous proteins are also considered to be within the scope of this invention. For example, provided herein is any protein fragment of a reference protein (meaning a polypeptide sequence at least one amino acid residue shorter than a reference polypeptide sequence but otherwise identical) 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or greater than 100 amino acids in length In another example, any protein that includes a stretch of about 20, about 30, about 40, about 50, or about 100 amino acids which are about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, or about 100% identical to any of the sequences described herein can be utilized in accordance with the invention. In certain embodiments, a protein sequence to be utilized in accordance with the invention includes 2, 3, 4, 5, 6, 7, 8, 9, 10, or more mutations as shown in any of the sequences provided or referenced herein.

**Polypeptide libraries**

[000360] Also provided are polynucleotide libraries containing nucleoside modifications, wherein the polynucleotides individually contain a first nucleic acid sequence encoding a polypeptide, such as an antibody, protein binding partner, scaffold protein, and other polypeptides known in the art. Preferably, the polynucleotides are mRNA in a form suitable for direct introduction into a target cell host, which in turn synthesizes the encoded polypeptide.

[000361] In certain embodiments, multiple variants of a protein, each with different amino acid modification(s), are produced and tested to determine the best variant in terms of pharmacokinetics, stability, biocompatibility, and/or biological activity, or a biophysical property such as expression level. Such a library may contain 10, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, or over 10^9 possible
variants (including substitutions, deletions of one or more residues, and insertion of one or more residues).

**Polypeptide-nucleic acid complexes**

[000362] Proper protein translation involves the physical aggregation of a number of polypeptides and nucleic acids associated with the mRNA. Provided by the invention are complexes containing conjugates of protein and nucleic acids, containing a translatable mRNA having one or more nucleoside modifications (e.g., at least two different nucleoside modifications) and one or more polypeptides bound to the mRNA. Generally, the proteins are provided in an amount effective to prevent or reduce an innate immune response of a cell into which the complex is introduced.

**Targeting Moieties**

[000363] In embodiments of the invention, modified nucleic acids are provided to express a protein-binding partner or a receptor on the surface of the cell, which functions to target the cell to a specific tissue space or to interact with a specific moiety, either in vivo or in vitro. Suitable protein-binding partners include antibodies and functional fragments thereof, scaffold proteins, or peptides. Additionally, modified nucleic acids can be employed to direct the synthesis and extracellular localization of lipids, carbohydrates, or other biological moieties.

[000364] As described herein, a useful feature of the modified nucleic acids of the invention is the capacity to reduce the innate immune response of a cell to an exogenous nucleic acid. Provided are methods for performing the titration, reduction or elimination of the immune response in a cell or a population of cells. In some embodiments, the cell is contacted with a first composition that contains a first dose of a first exogenous nucleic acid including a translatable region and at least one nucleoside modification, and the level of the innate immune response of the cell to the first exogenous nucleic acid is determined. Subsequently, the cell is contacted with a second composition, which includes a second dose of the first exogenous nucleic acid, the second dose containing a lesser amount of the first exogenous nucleic acid as compared to the first dose.

Alternatively, the cell is contacted with a first dose of a second exogenous nucleic acid. The second exogenous nucleic acid may contain one or more modified nucleosides, which may be the same or different from the first exogenous nucleic acid or, alternatively, the second exogenous nucleic acid may not contain modified nucleosides. The steps of contacting the cell with the first composition and/or the second composition may be repeated one or more times.
Additionally, efficiency of protein production (e.g., protein translation) in the cell is optionally determined, and the cell may be re-transfected with the first and/or second composition repeatedly until a target protein production efficiency is achieved.

Vaccines

As described herein, provided are mRNAs having sequences that are substantially not translatable. Such mRNA is effective as a vaccine when administered to a mammalian subject.

Also provided are modified nucleic acids that contain one or more noncoding regions. Such modified nucleic acids are generally not translated, but are capable of binding to and sequestering one or more translational machinery component such as a ribosomal protein or a transfer RNA (tRNA), thereby effectively reducing protein expression in the cell. The modified nucleic acid may contain a small nucleolar RNA (sno-RNA), micro RNA (miRNA), small interfering RNA (siRNA) or Piwi-interacting RNA (piRNA).

Additionally, certain modified nucleosides, or combinations thereof, when introduced into modified nucleic acids activate the innate immune response. Such activating modified nucleic acids, e.g., modified RNAs, are useful as adjuvants when combined with polypeptide or other vaccines. In certain embodiments, the activated modified mRNAs contain a translatable region which encodes for a polypeptide sequence useful as a vaccine, thus providing the ability to be a self-adjuvant.

Therapeutic Agents

The modified nucleic acids (modified RNAs) and the proteins translated from the modified nucleic acids described herein can be used as therapeutic agents. For example, a modified nucleic acid described herein can be administered to a subject, wherein the modified nucleic acid is translated in vivo to produce a therapeutic peptide in the subject. Provided are compositions, methods, kits, and reagents for treatment or prevention of disease or conditions in humans and other mammals. The active therapeutic agents of the invention include modified nucleic acids, cells containing modified nucleic acids or polypeptides translated from the modified nucleic acids, polypeptides translated from modified nucleic acids, and cells contacted with cells containing modified nucleic acids or polypeptides translated from the modified nucleic acids.

In certain embodiments, provided are combination therapeutics containing one or more modified nucleic acids containing translatable regions that encode for a protein or proteins that boost a mammalian subject's immunity along with a protein that induces antibody-dependent cellular
toxicity. For example, provided are therapeutics containing one or more nucleic acids that encode trastuzumab and granulocyte-colony stimulating factor (G-CSF). In particular, such combination therapeutics are useful in Her2+ breast cancer patients who develop induced resistance to trastuzumab. (See, e.g., Albrecht, Immunotherapy. 2(6):795-8 (2010)).

[000371] Provided are methods of inducing translation of a recombinant polypeptide in a cell population using the modified nucleic acids described herein. Such translation can be in vivo, ex vivo, in culture, or in vitro. The cell population is contacted with an effective amount of a composition containing a nucleic acid that has at least one nucleoside modification, and a translatable region encoding the recombinant polypeptide. The population is contacted under conditions such that the nucleic acid is localized into one or more cells of the cell population and the recombinant polypeptide is translated in the cell from the nucleic acid.

[000372] An effective amount of the composition is provided based, at least in part, on the target tissue, target cell type, means of administration, physical characteristics of the nucleic acid (e.g., size, and extent of modified nucleosides), and other determinants. In general, an effective amount of the composition provides efficient protein production in the cell, preferably more efficient than a composition containing a corresponding unmodified nucleic acid. Increased efficiency may be demonstrated by increased cell transfection (i.e., the percentage of cells transfected with the nucleic acid), increased protein translation from the nucleic acid, decreased nucleic acid degradation (as demonstrated, e.g., by increased duration of protein translation from a modified nucleic acid), or reduced innate immune response of the host cell.

[000373] Aspects of the invention are directed to methods of inducing in vivo translation of a recombinant polypeptide in a mammalian subject in need thereof. Therein, an effective amount of a composition containing a nucleic acid that has at least one nucleoside modification and a translatable region encoding the recombinant polypeptide is administered to the subject using the delivery methods described herein. The nucleic acid is provided in an amount and under other conditions such that the nucleic acid is localized into a cell of the subject and the recombinant polypeptide is translated in the cell from the nucleic acid. The cell in which the nucleic acid is localized, or the tissue in which the cell is present, may be targeted with one or more than one rounds of nucleic acid administration.

[000374] Other aspects of the invention relate to transplantation of cells containing modified nucleic acids to a mammalian subject. Administration of cells to mammalian subjects is known to those of ordinary skill in the art, such as local implantation (e.g., topical or subcutaneous
administration), organ delivery or systemic injection (e.g., intravenous injection or inhalation), as is
the formulation of cells in pharmaceutically acceptable carrier. Compositions containing modified
nucleic acids are formulated for administration intramuscularly, transarterially, intraocularly,
vaginally, rectally, intraperitoneally, intravenously, intranasally, subcutaneously, endoscopically,
transdermally, intramuscularly, intraventricularly, intradermally, intrathecially, topically (e.g. by
powders, ointments, creams, gels, lotions, and/or drops), mucosally, nasal, enterally, intratumorally,
by intratracheal instillation, bronchial instillation, and/or inhalation; nasal spray and/or aerosol,
and/or through a portal vein catheter. In some embodiments, the composition is formulated for
extended release. In specific embodiments, modified nucleic acid molecules or complexes, and/or
pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof, may be administered in a
way which allows the modified nucleic acid molecules or complex to cross the blood-brain barrier,
vascular barrier, or other epithelial barrier.

[000375] However, the present disclosure encompasses the delivery of modified nucleic acid
molecules or complexes, and/or pharmaceutical, prophylactic, diagnostic, or imaging compositions
thereof, by any appropriate route taking into consideration likely advances in the sciences of drug
delivery.

[000376] The subject to whom the therapeutic agent is administered suffers from or is at risk of
developing a disease, disorder, or deleterious condition. Provided are methods of identifying,
diagnosing, and classifying subjects on these bases, which may include clinical diagnosis, biomarker
levels, genome-wide association studies (GWAS), and other methods known in the art.

[000377] In certain embodiments, the administered modified nucleic acid directs production of one
or more recombinant polypeptides that provide a functional activity which is substantially absent in
the cell in which the recombinant polypeptide is translated. For example, the missing functional
activity may be enzymatic, structural, or gene regulatory in nature. In related embodiments, the
administered modified nucleic acid directs production of one or more recombinant polypeptides that
increases (e.g., synergistically) a functional activity which is present but substantially deficient in the
cell in which the recombinant polypeptide is translated.

[000378] In other embodiments, the administered modified nucleic acid directs production of one
or more recombinant polypeptides that replace a polypeptide (or multiple polypeptides) that is
substantially absent in the cell in which the recombinant polypeptide is translated. Such absence may
be due to genetic mutation of the encoding gene or regulatory pathway thereof. In some
embodiments, the recombinant polypeptide increases the level of an endogenous protein in the cell
to a desirable level; such an increase may bring the level of the endogenous protein from a subnormal level to a normal level, or from a normal level to a super-normal level.

[000379] Alternatively, the recombinant polypeptide functions to antagonize the activity of an endogenous protein present in, on the surface of, or secreted from the cell. Usually, the activity of the endogenous protein is deleterious to the subject, for example, do to mutation of the endogenous protein resulting in altered activity or localization. Additionally, the recombinant polypeptide antagonizes, directly or indirectly, the activity of a biological moiety present in, on the surface of, or secreted from the cell. Examples of antagonized biological moieties include lipids (e.g., cholesterol), a lipoprotein (e.g., low density lipoprotein), a nucleic acid, a carbohydrate, a protein toxin such as shiga and tetanus toxins, or a small molecule toxin such as botulinum, cholera, and diphtheria toxins. Additionally, the antagonized biological molecule may be an endogenous protein that exhibits an undesirable activity, such as a cytotoxic or cytostatic activity.

[000380] The recombinant proteins described herein are engineered for localization within the cell, potentially within a specific compartment such as the nucleus, or are engineered for secretion from the cell or translocation to the plasma membrane of the cell.

**Therapeutics**

[000381] Provided are methods for treating or preventing a symptom of diseases characterized by missing or aberrant protein activity, by replacing the missing protein activity or overcoming the aberrant protein activity. Because of the rapid initiation of protein production following introduction of modified mRNAs, as compared to viral DNA vectors, the compounds of the present invention are particularly advantageous in treating acute diseases such as sepsis, stroke, and myocardial infarction. Moreover, the lack of transcriptional regulation of the modified mRNAs of the invention is advantageous in that accurate titration of protein production is achievable.

[000382] In some embodiments, modified mRNAs may be derived from cDNA.

[000383] In some embodiments, modified mRNAs and their encoded polypeptides in accordance with the present invention may be used for therapeutic purposes. In some embodiments, modified mRNAs and their encoded polypeptides in accordance with the present invention may be used for treatment of any of a variety of diseases, disorders, and/or conditions, including but not limited to one or more of the following: autoimmune disorders (e.g., diabetes, lupus, multiple sclerosis, psoriasis, rheumatoid arthritis); inflammatory disorders (e.g., arthritis, pelvic inflammatory disease); infectious diseases (e.g., viral infections (e.g., HIV, HCV, RSV), bacterial infections, fungal
infections, sepsis); neurological disorders (e.g. Alzheimer's disease, Huntington's disease; autism; Duchenne muscular dystrophy); cardiovascular disorders (e.g. atherosclerosis, hypercholesterolemia, thrombosis, clotting disorders, angiogenic disorders such as macular degeneration); proliferative disorders (e.g. cancer, benign neoplasms); respiratory disorders (e.g. chronic obstructive pulmonary disease); digestive disorders (e.g. inflammatory bowel disease, ulcers); musculoskeletal disorders (e.g. fibromyalgia, arthritis); endocrine, metabolic, and nutritional disorders (e.g. diabetes, osteoporosis); urological disorders (e.g. renal disease); psychological disorders (e.g. depression, schizophrenia); skin disorders (e.g. wounds, eczema); blood and lymphatic disorders (e.g. anemia, hemophilia); etc.

Diseases characterized by dysfunctional or aberrant protein activity include cystic fibrosis, sickle cell anemia, epidermolysis bullosa, amyotrophic lateral sclerosis, and glucose-phosphate dehydrogenase deficiency. The present invention provides a method for treating such conditions or diseases in a subject by introducing nucleic acid or cell-based therapeutics containing the modified nucleic acids provided herein, wherein the modified nucleic acids encode for a protein that antagonizes or otherwise overcomes the aberrant protein activity present in the cell of the subject. Specific examples of a dysfunctional protein are the missense mutation variants of the cystic fibrosis transmembrane conductance regulator (CFTR) gene, which produce a dysfunctional protein variant of CFTR protein, which causes cystic fibrosis.

Diseases characterized by missing (or substantially diminished such that proper protein function does not occur) protein activity include cystic fibrosis, Niemann-Pick type C, β thalassemia major, Duchenne muscular dystrophy, Hurler Syndrome, Hunter Syndrome, and Hemophilia A. Such proteins may not be present, or are essentially non-functional. The present invention provides a method for treating such conditions or diseases in a subject by introducing nucleic acid or cell-based therapeutics containing the modified nucleic acids provided herein, wherein the modified nucleic acids encode for a protein that replaces the protein activity missing from the target cells of the subject. Specific examples of a dysfunctional protein are the nonsense mutation variants of the cystic fibrosis transmembrane conductance regulator (CFTR) gene, which produce a nonfunctional protein variant of CFTR protein, which causes cystic fibrosis.

Thus, provided are methods of treating cystic fibrosis in a mammalian subject by contacting a cell of the subject with a modified nucleic acid having a translatable region that encodes a functional CFTR polypeptide, under conditions such that an effective amount of the CTFR polypeptide is present in the cell. Preferred target cells are epithelial, endothelial and mesothelial
cells, such as the lung, and methods of administration are determined in view of the target tissue; i.e.,
for lung delivery, the RNA molecules are formulated for administration by inhalation.

In another embodiment, the present invention provides a method for treating hyperlipidemia in a subject, by introducing into a cell population of the subject with a modified mRNA molecule encoding Sortilin, a protein recently characterized by genomic studies, thereby ameliorating the hyperlipidemia in a subject. The SORT1 gene encodes a trans-Golgi network (TGN) transmembrane protein called Sortilin. Genetic studies have shown that one of five individuals has a single nucleotide polymorphism, rs12740374, in the lpl3 locus of the SORT1 gene that predisposes them to having low levels of low-density lipoprotein (LDL) and very-low-density lipoprotein (VLDL). Each copy of the minor allele, present in about 30% of people, alters LDL cholesterol by 8 mg/dL, while two copies of the minor allele, present in about 5% of the population, lowers LDL cholesterol 16 mg/dL. Carriers of the minor allele have also been shown to have a 40% decreased risk of myocardial infarction. Functional in vivo studies in mice describes that overexpression of SORT1 in mouse liver tissue led to significantly lower LDL-cholesterol levels, as much as 80% lower, and that silencing SORT1 increased LDL cholesterol approximately 200% (Musunuru K et al. From noncoding variant to phenotype via SORT1 at the lpl3 cholesterol locus. Nature 2010; 466: 714-721).

Modulation of cell fate

Provided are methods of inducing an alteration in cell fate in a target mammalian cell. The target mammalian cell may be a precursor cell and the alteration may involve driving differentiation into a lineage, or blocking such differentiation. Alternatively, the target mammalian cell may be a differentiated cell, and the cell fate alteration includes driving de-differentiation into a pluripotent precursor cell, or blocking such de-differentiation, such as the dedifferentiation of cancer cells into cancer stem cells. In situations where a change in cell fate is desired, effective amounts of mRNAs encoding a cell fate inductive polypeptide is introduced into a target cell under conditions such that an alteration in cell fate is induced. In some embodiments, the modified mRNAs are useful to reprogram a subpopulation of cells from a first phenotype to a second phenotype. Such a reprogramming may be temporary or permanent. Optionally, the reprogramming induces a target cell to adopt an intermediate phenotype.

Additionally, the methods of the present invention are particularly useful to generate induced pluripotent stem cells (iPS cells) because of the high efficiency of transfection, the ability to
re-transfect cells, and the tenability of the amount of recombinant polypeptides produced in the target cells. Further, the use of iPS cells generated using the methods described herein is expected to have a reduced incidence of teratoma formation.

[000390] Also provided are methods of reducing cellular differentiation in a target cell population. For example, a target cell population containing one or more precursor cell types is contacted with a composition having an effective amount of a modified mRNA encoding a polypeptide, under conditions such that the polypeptide is translated and reduces the differentiation of the precursor cell. In non-limiting embodiments, the target cell population contains injured tissue in a mammalian subject or tissue affected by a surgical procedure. The precursor cell is, e.g., a stromal precursor cell, a neural precursor cell, or a mesenchymal precursor cell.

[000391] In a specific embodiment, provided are modified nucleic acids that encode one or more differentiation factors Gata4, Mef2c and Tbx4. These mRNA-generated factors are introduced into fibroblasts and drive the reprogramming into cardiomyocytes. Such a reprogramming can be performed in vivo, by contacting an mRNA-containing patch or other material to damaged cardiac tissue to facilitate cardiac regeneration. Such a process promotes cardiomyocyte genesis as opposed to fibrosis.

**Targeting of pathogenic organisms: purification of biological materials**

[000392] Provided herein are methods for targeting pathogenic microorganisms, such as bacteria, yeast, protozoa, helminthes and the like, using modified mRNAs that encode cytostatic or cytotoxic polypeptides. Preferably the mRNA introduced into the target pathogenic organism contains modified nucleosides or other nucleic acid sequence modifications that the mRNA is translated exclusively, or preferentially, in the target pathogenic organism, to reduce possible off-target effects of the therapeutic. Such methods are useful for removing pathogenic organisms from biological material, including blood, semen, eggs, and transplant materials including embryos, tissues, and organs.

**Targeting diseased cells**

[000393] Provided herein are methods for targeting pathogenic or diseased cells, particularly cancer cells, using modified mRNAs that encode cytostatic or cytotoxic polypeptides. Preferably the mRNA introduced into the target pathogenic cell contains modified nucleosides or other nucleic acid sequence modifications that the mRNA is translated exclusively, or preferentially, in the target pathogenic cell, to reduce possible off-target effects of the therapeutic. Alternatively, the invention
provides targeting moieties that are capable of targeting the modified mRNAs to preferentially bind to and enter the target pathogenic cell.

**Protein production**

[000394] The methods provided herein are useful for enhancing protein product yield in a cell culture process. In a cell culture containing a plurality of host cells, introduction of the modified mRNAs described herein results in increased protein production efficiency relative to a corresponding unmodified nucleic acid. Such increased protein production efficiency can be demonstrated, e.g., by showing increased cell transfection, increased protein translation from the nucleic acid, decreased nucleic acid degradation, and/or reduced innate immune response of the host cell. Protein production can be measured by ELISA, and protein activity can be measured by various functional assays known in the art. The protein production may be generated in a continuous or a fed-batch mammalian process.

[000395] Additionally, it is useful to optimize the expression of a specific polypeptide in a cell line or collection of cell lines of potential interest, particularly an engineered protein such as a protein variant of a reference protein having a known activity. In one embodiment, provided is a method of optimizing expression of an engineered protein in a target cell, by providing a plurality of target cell types, and independently contacting with each of the plurality of target cell types a modified mRNA encoding an engineered polypeptide. Additionally, culture conditions may be altered to increase protein production efficiency. Subsequently, the presence and/or level of the engineered polypeptide in the plurality of target cell types is detected and/or quantitated, allowing for the optimization of an engineered polypeptide's expression by selection of an efficient target cell and cell culture conditions relating thereto. Such methods are particularly useful when the engineered polypeptide contains one or more post-translational modifications or has substantial tertiary structure, situations which often complicate efficient protein production.

**Gene silencing**

[000396] The modified mRNAs described herein are useful to silence (i.e., prevent or substantially reduce) expression of one or more target genes in a cell population. A modified mRNA encoding a polypeptide capable of directing sequence-specific histone H3 methylation is introduced into the cells in the population under conditions such that the polypeptide is translated and reduces gene transcription of a target gene via histone H3 methylation and subsequent heterochromatin formation. In some embodiments, the silencing mechanism is performed on a cell population present in a
mammalian subject. By way of non-limiting example, a useful target gene is a mutated Janus Kinase-2 family member, wherein the mammalian subject expresses the mutant target gene suffers from a myeloproliferative disease resulting from aberrant kinase activity.

[000397] Co-administration of modified mRNAs and siRNAs are also provided herein. As demonstrated in yeast, sequence-specific trans silencing is an effective mechanism for altering cell function. Fission yeast require two RNAi complexes for siRNA-mediated heterochromatin assembly: the RNA-induced transcriptional silencing (RITS) complex and the RNA-directed RNA polymerase complex (RDRC) (Motamedi et al. Cell 2004, 119, 789-802). In fission yeast, the RITS complex contains the siRNA binding Argonaute family protein Agol, a chromodomain protein Chpl, and Tas3. The fission yeast RDRC complex is composed of an RNA-dependent RNA Polymerase Rdpl, a putative RNA helicase Hrl, and a polyA polymerase family protein Cidl2. These two complexes require the Dicer ribonuclease and Clr4 histone H3 methyltransferase for activity. Together, Agol binds siRNA molecules generated through Dicer-mediated cleavage of Rdpl co-transcriptionally generated dsRNA transcripts and allows for the sequence-specific direct association of Chpl, Tas3, Hrl, and Clr4 to regions of DNA destined for methylation and histone modification and subsequent compaction into transcriptionally silenced heterochromatin. While this mechanism functions in cis- with centromeric regions of DNA, sequence-specific trans silencing is possible through co-transfection with double-stranded siRNAs for specific regions of DNA and concomitant RNAi-directed silencing of the siRNA ribonuclease Eril (Buhler et al. Cell 2006, 125, 873-886).

Modulation of biological pathways

[000398] The rapid translation of modified mRNAs introduced into cells provides a desirable mechanism of modulating target biological pathways. Such modulation includes antagonism or agonism of a given pathway. In one embodiment, a method is provided for antagonizing a biological pathway in a cell by contacting the cell with an effective amount of a composition comprising a modified nucleic acid encoding a recombinant polypeptide, under conditions such that the nucleic acid is localized into the cell and the recombinant polypeptide is capable of being translated in the cell from the nucleic acid, wherein the recombinant polypeptide inhibits the activity of a polypeptide functional in the biological pathway. Exemplary biological pathways are those defective in an autoimmune or inflammatory disorder such as multiple sclerosis, rheumatoid arthritis, psoriasis, lupus erythematosus, ankylosing spondylitis colitis, or Crohn's disease; in particular, antagonism of

Further, provided are modified nucleic acids encoding an antagonist for chemokine receptors; chemokine receptors CXCR-4 and CCR-5 are required for, e.g., HIV entry into host cells (Arenzana-Seisdedos F et al, (1996) Nature. Oct 3;383(6599):400).

Alternatively, provided are methods of agonizing a biological pathway in a cell by contacting the cell with an effective amount of a modified nucleic acid encoding a recombinant polypeptide under conditions such that the nucleic acid is localized into the cell and the recombinant polypeptide is capable of being translated in the cell from the nucleic acid, and the recombinant polypeptide induces the activity of a polypeptide functional in the biological pathway. Exemplary agonized biological pathways include pathways that modulate cell fate determination. Such agonization is reversible or, alternatively, irreversible.

Cellular nucleic acid delivery

Methods of the present invention enhance nucleic acid delivery into a cell population, in vivo, ex vivo, or in culture. For example, a cell culture containing a plurality of host cells (e.g., eukaryotic cells such as yeast or mammalian cells) is contacted with a composition that contains an enhanced nucleic acid having at least one nucleoside modification and, optionally, a translatable region. The composition also generally contains a transfection reagent or other compound that increases the efficiency of enhanced nucleic acid uptake into the host cells. The enhanced nucleic acid exhibits enhanced retention in the cell population, relative to a corresponding unmodified nucleic acid. The retention of the enhanced nucleic acid is greater than the retention of the unmodified nucleic acid. In some embodiments, it is at least about 50%, 75%, 90%, 95%, 100%, 150%, 200% or more than 200% greater than the retention of the unmodified nucleic acid. Such retention advantage may be achieved by one round of transfection with the enhanced nucleic acid, or may be obtained following repeated rounds of transfection.

In some embodiments, the enhanced nucleic acid is delivered to a target cell population with one or more additional nucleic acids. Such delivery may be at the same time, or the enhanced nucleic acid is delivered prior to delivery of the one or more additional nucleic acids. The additional one or more nucleic acids may be modified nucleic acids or unmodified nucleic acids. It is understood that the initial presence of the enhanced nucleic acids does not substantially induce an innate immune response of the cell population and, moreover, that the innate immune response will
not be activated by the later presence of the unmodified nucleic acids. In this regard, the enhanced nucleic acid may not itself contain a translatable region, if the protein desired to be present in the target cell population is translated from the unmodified nucleic acids.

Expression of Ligand or Receptor on Cell Surface

[000403] In some aspects and embodiments of the aspects described herein, the modified RNAs can be used to express a ligand or ligand receptor on the surface of a cell (e.g., a homing moiety). A ligand or ligand receptor moiety attached to a cell surface can permit the cell to have a desired biological interaction with a tissue or an agent in vivo. A ligand can be an antibody, an antibody fragment, an aptamer, a peptide, a vitamin, a carbohydrate, a protein or polypeptide, a receptor, e.g., cell-surface receptor, an adhesion molecule, a glycoprotein, a sugar residue, a therapeutic agent, a drug, a glycosaminoglycan, or any combination thereof. For example, a ligand can be an antibody that recognizes a cancer-cell specific antigen, rendering the cell capable of preferentially interacting with tumor cells to permit tumor-specific localization of a modified cell. A ligand can confer the ability of a cell composition to accumulate in a tissue to be treated, since a preferred ligand may be capable of interacting with a target molecule on the external face of a tissue to be treated. Ligands having limited cross-reactivity to other tissues are generally preferred.

[000404] In some cases, a ligand can act as a homing moiety which permits the cell to target to a specific tissue or interact with a specific ligand. Such homing moieties can include, but are not limited to, any member of a specific binding pair, antibodies, monoclonal antibodies, or derivatives or analogs thereof, including without limitation: Fv fragments, single chain Fv (scFv) fragments, Fab' fragments, F(\(\text{ab}'\))2 fragments, single domain antibodies, camelized antibodies and antibody fragments, humanized antibodies and antibody fragments, and multivalent versions of the foregoing; multivalent binding reagents including without limitation: monospecific or bispecific antibodies, such as disulfide stabilized Fv fragments, scFv tandems ((SCFV)2 fragments), diabodies, tribodies or tetrabodies, which typically are covalently linked or otherwise stabilized (i.e., leucine zipper or helix stabilized) scFv fragments; and other homing moieties include for example, aptamers, receptors, and fusion proteins.

[000405] In some embodiments, the homing moiety may be a surface-bound antibody, which can permit tuning of cell targeting specificity. This is especially useful since highly specific antibodies can be raised against an epitope of interest for the desired targeting site. In one embodiment, multiple antibodies are expressed on the surface of a cell, and each antibody can have a different
specificity for a desired target. Such approaches can increase the avidity and specificity of homing interactions.

A skilled artisan can select any homing moiety based on the desired localization or function of the cell, for example an estrogen receptor ligand, such as tamoxifen, can target cells to estrogen-dependent breast cancer cells that have an increased number of estrogen receptors on the cell surface. Other non-limiting examples of ligand/receptor interactions include CCRI (e.g., for treatment of inflamed joint tissues or brain in rheumatoid arthritis, and/or multiple sclerosis), CCR7, CCR8 (e.g., targeting to lymph node tissue), CCR6, CCR9,CCR10 (e.g., to target to intestinal tissue), CCR4, CCR10 (e.g., for targeting to skin), CXCR4 (e.g., for general enhanced transmigration), HCELL (e.g., for treatment of inflammation and inflammatory disorders, bone marrow), Alpha4beta7 (e.g., for intestinal mucosa targeting), VLA-4/VCAM-1 (e.g., targeting to endothelium). In general, any receptor involved in targeting (e.g., cancer metastasis) can be harnessed for use in the methods and compositions described herein.

Mediators of Cell Death

In one embodiment, a modified nucleic acid molecule composition can be used to induce apoptosis in a cell (e.g., a cancer cell) by increasing the expression of a death receptor, a death receptor ligand or a combination thereof. This method can be used to induce cell death in any desired cell and has particular usefulness in the treatment of cancer where cells escape natural apoptotic signals.

Apoptosis can be induced by multiple independent signaling pathways that converge upon a final effector mechanism consisting of multiple interactions between several "death receptors" and their ligands, which belong to the tumor necrosis factor (TNF) receptor/ligand superfamily. The best-characterized death receptors are CD95 ("Fas"), TNFRI (p55), death receptor 3 (DR3 or Apo3/TRAMO), DR4 and DR5 (apo2-TRAIL-R2). The final effector mechanism of apoptosis may be the activation of a series of proteinases designated as caspases. The activation of these caspases results in the cleavage of a series of vital cellular proteins and cell death. The molecular mechanism of death receptors/ligands-induced apoptosis is well known in the art. For example, Fas/FasL-mediated apoptosis is induced by binding of three FasL molecules which induces trimerization of Fas receptor via C-terminus death domains (DDs), which in turn recruits an adapter protein FADD (Fas-associated protein with death domain) and Caspase-8. The oligomerization of this trimolecular complex, Fas/FAIDD/caspase-8, results in proteolytic cleavage of proenzyme
caspase-8 into active caspase-8 that, in turn, initiates the apoptosis process by activating other downstream caspases through proteolysis, including caspase-3. Death ligands in general are apoptotic when formed into trimers or higher order of structures. As monomers, they may serve as antiapoptotic agents by competing with the trimers for binding to the death receptors.

[000409] In one embodiment, the modified nucleic acid molecule composition encodes for a death receptor (e.g., Fas, TRAIL, TRAMO, TNFR, TLR etc). Cells made to express a death receptor by transfection of modified RNA become susceptible to death induced by the ligand that activates that receptor. Similarly, cells made to express a death ligand, e.g., on their surface, will induce death of cells with the receptor when the transfected cell contacts the target cell. In another embodiment, the modified RNA composition encodes for a death receptor ligand (e.g., FasL, TNF, etc). In another embodiment, the modified RNA composition encodes a caspase (e.g., caspase 3, caspase 8, caspase 9 etc). Where cancer cells often exhibit a failure to properly differentiate to a non-proliferative or controlled proliferative form, in another embodiment, the synthetic, modified RNA composition encodes for both a death receptor and its appropriate activating ligand. In another embodiment, the synthetic, modified RNA composition encodes for a differentiation factor that when expressed in the cancer cell, such as a cancer stem cell, will induce the cell to differentiate to a non-pathogenic or nonself-renewing phenotype (e.g., reduced cell growth rate, reduced cell division etc) or to induce the cell to enter a dormant cell phase (e.g., G0/resting phase).

[000410] One of skill in the art will appreciate that the use of apoptosis-inducing techniques may require that the modified nucleic acid molecules are appropriately targeted to e.g., tumor cells to prevent unwanted wide-spread cell death. Thus, one can use a delivery mechanism (e.g., attached ligand or antibody, targeted liposome etc) that recognizes a cancer antigen such that the modified nucleic acid molecules are expressed only in cancer cells.

Devices

[000411] The present invention provides for devices, in particular portable devices, which incorporate modified nucleosides and nucleotides into nucleic acids such as ribonucleic acids (RNA) that encode proteins of interest. These devices contain in a stable formulation the reagents to synthesize a modified RNA in a formulation available to be immediately delivered to a subject in need thereof, such as a human patient. Non-limiting examples of such a protein of interest include a growth factor and/or angiogenesis stimulator for wound healing, a peptide antibiotic to facilitate infection control, and an antigen to rapidly stimulate an immune response to a newly identified virus.
[000412] In some embodiments the device is self-contained, and is optionally capable of wireless remote access to obtain instructions for synthesis and/or analysis of the generated nucleic acid. The device is capable of mobile synthesis of at least one nucleic acid, and preferably an unlimited number of different nucleic acid sequences. In certain embodiments, the device is capable of being transported by one or a small number of individuals. In other embodiments, the device is scaled to fit on a benchtop or desk. In other embodiments, the device is scaled to fit into a suitcase, backpack or similarly sized object. In further embodiments, the device is scaled to fit into a vehicle, such as a car, truck or ambulance, or a military vehicle such as a tank or personnel carrier. The information necessary to generate a modified mRNA encoding protein of interest is present within a computer readable medium present in the device.

[000413] In some embodiments, the device is capable of communication (e.g., wireless communication) with a database of nucleic acid and polypeptide sequences. The device contains at least one sample block for insertion of one or more sample vessels. Such sample vessels are capable of accepting in liquid or other form any number of materials such as template DNA, nucleotides, enzymes, buffers, and other reagents. The sample vessels are also capable of being heated and cooled by contact with the sample block. The sample block is generally in communication with a device base with one or more electronic control units for the at least one sample block. The sample block preferably contains a heating module, such heating molecule capable of heating and/or cooling the sample vessels and contents thereof to temperatures between about -20C and above +100C. The device base is in communication with a voltage supply such as a battery or external voltage supply. The device also contains means for storing and distributing the materials for RNA synthesis.

[000414] Optionally, the sample block contains a module for separating the synthesized nucleic acids. Alternatively, the device contains a separation module operably linked to the sample block. Preferably the device contains a means for analysis of the synthesized nucleic acid. Such analysis includes sequence identity (demonstrated such as by hybridization), absence of non-desired sequences, measurement of integrity of synthesized mRNA (such has by microfluidic viscometry combined with spectrophotometry), and concentration and/or potency of modified RNA (such as by spectrophotometry).

[000415] In certain embodiments, the device is combined with a means for detection of pathogens present in a biological material obtained from a subject, e.g., the IBIS PLEX-ID system (Abbott) for microbial identification.
**Compositions and formulations containing modified RNAs**

[000416] Provided are formulations containing an effective amount of a ribonucleic acid (e.g., an mRNA or a nucleic acid containing an mRNA) engineered to avoid an innate immune response of a cell into which the ribonucleic acid enters. The ribonucleic acid generally includes a nucleotide sequence encoding a polypeptide of interest.

[000417] In certain embodiments, the formulations include one or more cell penetration agents, e.g., transfection agents. In one specific embodiment, a ribonucleic acid is mixed or admixed with a transfection agent (or mixture thereof) and the resulting mixture is employed to transfect cells. Preferred transfection agents are cationic lipid compositions, particularly monovalent and polyvalent cationic lipid compositions, more particularly "LIPOFECTIN," "LIPOFECTACE," "LIPOFECTAMINE," "CELLFECTIN," DMRIE-C, DMRIE, DOTAP, DOSPA, and DOSPER, and dendrimer compositions, particularly G5-G10 dendrimers, including dense star dendrimers, PAMAM dendrimers, grafted dendrimers, and dendrimers known as dendrigrafts and "SUPERFECT." In a second specific transfection method, a ribonucleic acid is conjugated to a nucleic acid-binding group, for example a polyamine and more particularly a spermine, which is then introduced into the cell or admixed with a transfection agent (or mixture thereof) and the resulting mixture is employed to transfect cells. In a third specific embodiment, a mixture of one or more transfection-enhancing peptides, proteins, or protein fragments, including fusogenic peptides or proteins, transport or trafficking peptides or proteins, receptor-ligand peptides or proteins, or nuclear localization peptides or proteins and/or their modified analogs (e.g., spermine modified peptides or proteins) or combinations thereof are mixed with and complexed with a ribonucleic acid to be introduced into a cell, optionally being admixed with transfection agent and the resulting mixture is employed to transfect cells. Further, a component of a transfection agent (e.g., lipids, cationic lipids or dendrimers) is covalently conjugated to selected peptides, proteins, or protein fragments directly or via a linking or spacer group. Of particular interest in this embodiment are peptides or proteins that are fusogenic, membrane-permeabilizing, transport or trafficking, or which function for cell-targeting. The peptide- or protein-transfection agent complex is combined with a ribonucleic acid and employed for transfection.

[000418] In certain embodiments, the formulations include a pharmaceutically acceptable carrier that causes the effective amount of ribonucleic acid to be substantially retained in a target tissue containing the cell.
In certain embodiments, the formulation may include at least a modified nucleic acid and a delivery agent. In some embodiments, the delivery agent may comprise lipidoid-based formulations allowed for localized and systemic delivery of mmRNA.

Also provided are compositions for generation of an in vivo depot containing an engineered ribonucleotide. For example, the composition contains a bioerodible, biocompatible polymer, a solvent present in an amount effective to plasticize the polymer and form a gel therewith, and an engineered ribonucleic acid. In certain embodiments the composition also includes a cell penetration agent as described herein. In other embodiments, the composition also contains a thixotropic amount of a thixotropic agent mixable with the polymer so as to be effective to form a thixotropic composition. Further compositions include a stabilizing agent, a bulking agent, a chelating agent, or a buffering agent.

In other embodiments, provided are sustained-release delivery depots, such as for administration of an engineered ribonucleic acid an environment (meaning an organ or tissue site) in a patient. Such depots generally contain an engineered ribonucleic acid and a flexible chain polymer where both the engineered ribonucleic acid and the flexible chain polymer are entrapped within a porous matrix of a crosslinked matrix protein. Usually, the pore size is less than 1 nm, such as 900 nm, 800 nm, 700 nm, 600 nm, 500 nm, 400 nm, 300 nm, 200 nm, 100 nm, or less than 100 nm.

Usually the flexible chain polymer is hydrophilic. Usually the flexible chain polymer has a molecular weight of at least 50 kDa, such as 75 kDa, 100 kDa, 150 kDa, 200 kDa, 250 kDa, 300 kDa, 400 kDa, 500 kDa, or greater than 500 kDa. Usually the flexible chain polymer has a persistence length of less than 10%, such as 9, 8, 7, 6, 5, 4, 3, 2, 1 or less than 1% of the persistence length of the matrix protein. Usually the flexible chain polymer has a charge similar to that of the matrix protein. In some embodiments, the flexible chain polymer alters the effective pore size of a matrix of crosslinked matrix protein to a size capable of sustaining the diffusion of the engineered ribonucleic acid from the matrix into a surrounding tissue comprising a cell into which the engineered ribonucleic acid is capable of entering.

**In vivo delivery of ribonucleic acids**

As described herein, compositions containing the nucleic acids of the invention are formulated for administration intramuscularly, transarterially, intraperitoneally, intravenously, intranasally, subcutaneously, endoscopically, transdermally, or intrathecally. As described herein, in
some embodiments, the composition is formulated in depots for extended release. Generally, a
specific organ or tissue (a "target tissue") is targeted for administration.

[000423] In some aspects of the invention, the nucleic acids (particularly ribonucleic acids
encoding polypeptides) are spatially retained within or proximal to a target tissue. Provided are
method of providing a composition to a target tissue of a mammalian subject by contacting the target
tissue (which contains one or more target cells) with the composition under conditions such that the
composition, in particular the nucleic acid component(s) of the composition, is substantially retained
in the target tissue, meaning that at least 10, 20, 30, 40, 50, 60, 70, 80, 85, 90, 95, 96, 97, 98, 99,
99.9, 99.99 or greater than 99.99% of the composition is retained in the target tissue.

Advantageously, retention is determined by measuring the amount of the nucleic acid present in the
composition that enters one or more target cells. For example, at least 1, 5, 10, 20, 30, 40, 50, 60,
70, 80, 85, 90, 95, 96, 97, 98, 99, 99.9, 99.99 or greater than 99.99% of the nucleic acids
administered to the subject are present intracellularly at a period of time following administration.
For example, intramuscular injection to a mammalian subject is performed using an aqueous
composition containing a ribonucleic acid and a transfection reagent, and retention of the
composition is determined by measuring the amount of the ribonucleic acid present in the muscle
cells.

[000424] Aspects of the invention are directed to methods of providing a composition to a target
tissue of a mammalian subject, by contacting the target tissue (containing one or more target cells)
with the composition under conditions such that the composition is substantially retained in the
target tissue. The composition contains an effective amount of a ribonucleic acid engineered to
avoid an innate immune response of a cell into which the ribonucleic acid enters, where the
ribonucleic acid contains a nucleotide sequence encoding a polypeptide of interest, under conditions
such that the polypeptide of interest is produced in at least one target cell. The compositions
generally contain a cell penetration agent, although "naked" nucleic acid (such as nucleic acids
without a cell penetration agent or other agent) is also contemplated, and a pharmaceutically
acceptable carrier.

[000425] In some circumstances, the amount of a protein produced by cells in a tissue is desirably
increased. Preferably, this increase in protein production is spatially restricted to cells within the
target tissue. Thus, provided are methods of increasing production of a protein of interest in a tissue
of a mammalian subject. A composition is provided that contains a ribonucleic acid that is
engineered to avoid an innate immune response of a cell into which the ribonucleic acid enters and
encodes the polypeptide of interest and the composition is characterized in that a unit quantity of composition has been determined to produce the polypeptide of interest in a substantial percentage of cells contained within a predetermined volume of the target tissue. In some embodiments, the composition includes a plurality of different ribonucleic acids, where one or more than one of the ribonucleic acids is engineered to avoid an innate immune response of a cell into which the ribonucleic acid enters, and where one or more than one of the ribonucleic acids encodes a polypeptide of interest. Optionally, the composition also contains a cell penetration agent to assist in the intracellular delivery of the ribonucleic acid. A determination is made of the dose of the composition required to produce the polypeptide of interest in a substantial percentage of cells contained within the predetermined volume of the target tissue (generally, without inducing significant production of the polypeptide of interest in tissue adjacent to the predetermined volume, or distally to the target tissue). Subsequent to this determination, the determined dose is introduced directly into the tissue of the mammalian subject.

**Pharmaceutical Compositions**

[000426] The present invention provides enhanced nucleic acids, and complexes containing enhanced nucleic acids associated with other deliverable moieties. Thus, the present invention provides pharmaceutical compositions comprising one or more enhanced nucleic acids, or one or more such complexes, and one or more pharmaceutically acceptable excipients. Pharmaceutical compositions may optionally comprise one or more additional therapeutically active substances. In some embodiments, compositions are administered to humans. For the purposes of the present disclosure, the phrase "active ingredient" generally refers to an enhanced nucleic acid to be delivered as described herein.

[000427] Although the descriptions of pharmaceutical compositions provided herein are principally directed to pharmaceutical compositions which are suitable for administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to animals of all sorts. Modification of pharmaceutical compositions suitable for administration to humans in order to render the compositions suitable for administration to various animals is well understood, and the ordinarily skilled veterinary pharmacologist can design and/or perform such modification with merely ordinary, if any, experimentation. Subjects to which administration of the pharmaceutical compositions is contemplated include, but are not limited to, humans and/or other primates; mammals, including commercially relevant mammals such as cattle, pigs, horses, sheep,
cats, dogs, mice, and/or rats; and/or birds, including commercially relevant birds such as chickens, ducks, geese, and/or turkeys.

**000428** Formulations of the pharmaceutical compositions described herein may be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of bringing the active ingredient into association with an excipient and/or one or more other accessory ingredients, and then, if necessary and/or desirable, shaping and/or packaging the product into a desired single- or multi-dose unit.

**000429** A pharmaceutical composition in accordance with the invention may be prepared, packaged, and/or sold in bulk, as a single unit dose, and/or as a plurality of single unit doses. As used herein, a "unit dose" is discrete amount of the pharmaceutical composition comprising a predetermined amount of the active ingredient. The amount of the active ingredient is generally equal to the dosage of the active ingredient which would be administered to a subject and/or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage.

**000430** Relative amounts of the active ingredient, the pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the invention will vary, depending upon the identity, size, and/or condition of the subject treated and further depending upon the route by which the composition is to be administered. By way of example, the composition may comprise between 0.1% and 100% (w/w) active ingredient.

**000431** Pharmaceutical formulations may additionally comprise a pharmaceutically acceptable excipient, which, as used herein, includes any and all solvents, dispersion media, diluents, or other liquid vehicles, dispersion or suspension aids, surface active agents, isotonic agents, thickening or emulsifying agents, preservatives, solid binders, lubricants and the like, as suited to the particular dosage form desired. Remington's *The Science and Practice of Pharmacy, 21st Edition, A.R.* Gennaro (Lippincott, Williams & Wilkins, Baltimore, MD, 2006; incorporated herein by reference) discloses various excipients used in formulating pharmaceutical compositions and known techniques for the preparation thereof. Except insofar as any conventional excipient medium is incompatible with a substance or its derivatives, such as by producing any undesirable biological effect or otherwise interacting in a deleterious manner with any other component(s) of the pharmaceutical composition, its use is contemplated to be within the scope of this invention.

**000432** In some embodiments, a pharmaceutically acceptable excipient is at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% pure. In some embodiments, an excipient is approved for use in humans and for veterinary use. In some embodiments, an excipient is approved
by United States Food and Drug Administration. In some embodiments, an excipient is pharmaceutical grade. In some embodiments, an excipient meets the standards of the United States Pharmacopoeia (USP), the European Pharmacopoeia (EP), the British Pharmacopoeia, and/or the International Pharmacopoeia.

[000433] Pharmaceutically acceptable excipients used in the manufacture of pharmaceutical compositions include, but are not limited to, inert diluents, dispersing and/or granulating agents, surface active agents and/or emulsifiers, disintegrating agents, binding agents, preservatives, buffering agents, lubricating agents, and/or oils. Such excipients may optionally be included in pharmaceutical formulations. Excipients such as cocoa butter and suppository waxes, coloring agents, coating agents, sweetening, flavoring, and/or perfuming agents can be present in the composition, according to the judgment of the formulator.

[000434] Exemplary diluents include, but are not limited to, calcium carbonate, sodium carbonate, calcium phosphate, dicalcium phosphate, calcium sulfate, calcium hydrogen phosphate, sodium phosphate lactose, sucrose, cellulose, microcrystalline cellulose, kaolin, mannitol, sorbitol, inositol, sodium chloride, dry starch, cornstarch, powdered sugar, etc., and/or combinations thereof.

[000435] Exemplary granulating and/or dispersing agents include, but are not limited to, potato starch, corn starch, tapioca starch, sodium starch glycolate, clays, alginic acid, guar gum, citrus pulp, agar, bentonite, cellulose and wood products, natural sponge, cation-exchange resins, calcium carbonate, silicates, sodium carbonate, cross-linked poly(vinyl-pyrrolidone) (crospovidone), sodium carboxymethyl starch (sodium starch glycolate), carboxymethyl cellulose, cross-linked sodium carboxymethyl cellulose (croskarmellose), methylcellulose, pregelatinized starch (starch 1500), microcrystalline starch, water insoluble starch, calcium carboxymethyl cellulose, magnesium aluminum silicate (Veegum), sodium lauryl sulfate, quaternary ammonium compounds, etc., and/or combinations thereof.

[000436] Exemplary surface active agents and/or emulsifiers include, but are not limited to, natural emulsifiers (e.g. acacia, agar, alginic acid, sodium alginate, tragacanth, chondrux, cholesterol, xanthan, pectin, gelatin, egg yolk, casein, wool fat, cholesterol, wax, and lecithin), colloidal clays (e.g. bentonite [aluminum silicate] and Veegum® [magnesium aluminum silicate]), long chain amino acid derivatives, high molecular weight alcohols (e.g. stearyl alcohol, cetyl alcohol, oleyl alcohol, triacetin monostearate, ethylene glycol distearate, glycercyl monostearate, and propylene glycol monostearate, polyvinyl alcohol), carbomers (e.g. carboxy polymethylene, polyacrylic acid, acrylic acid polymer, and carboxyvinyl polymer), carrageenan, cellulosic derivatives (e.g. 82
carboxymethylcellulose sodium, powdered cellulose, hydroxymethyl cellulose, hydroxypropyl cellulose, hydroxypropyl methylcellulose, methylcellulose), sorbitan fatty acid esters (e.g. polyoxyethylene sorbitan monolaurate [Tween®20], polyoxyethylene sorbitan [Tween®60], polyoxyethylene sorbitan monooleate [Tween®80], sorbitan monopalmitate [Span®40], sorbitan monostearate [Span®60], sorbitan tristearate [Span®65], glycercy monooleate, sorbitan monooleate [Span®80]), polyoxyethylene esters (e.g. polyoxyethylene monostearate [Myrj®45], polyoxyethylene hydrogenated castor oil, polyethoxylated castor oil, polyoxyethylene stearate, and Solutol®), sucrose fatty acid esters, polyethylene glycol fatty acid esters (e.g. Cremophor®), polyoxyethylene ethers, (e.g. polyoxyethylene lauryl ether [Brij®30]), poly(vinyl-pyrrolidone), diethylene glycol monolaurate, triethanolamine oleate, sodium oleate, potassium oleate, ethyl oleate, oleic acid, ethyl laurate, sodium lauryl sulfate, Pluronic®F 68, Poloxamer®188, cetrimonium bromide, cetylpyridinium chloride, benzalkonium chloride, docusate sodium, etc. and/or combinations thereof.

[000437] Exemplary binding agents include, but are not limited to, starch (e.g. cornstarch and starch paste); gelatin; sugars (e.g. sucrose, glucose, dextrose, dextrin, molasses, lactose, lactitol, mannitol.); natural and synthetic gums (e.g. acacia, sodium alginate, extract of Irish moss, panwar gum, ghatti gum, mucilage of isapol husks, carboxymethylcellulose, methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropyl cellulose, hydroxypropyl methylcellulose, microcrystalline cellulose, cellulose acetate, poly(vinyl-pyrrolidone), magnesium aluminum silicate (Veegum®), and larch arabogalactan); alginites; polyethylene oxide; polyethylene glycol; inorganic calcium salts; silicic acid; polymethacrylates; waxes; water; alcohol; etc.; and combinations thereof.

[000438] Exemplary preservatives may include, but are not limited to, antioxidants, chelating agents, antimicrobial preservatives, antifungal preservatives, alcohol preservatives, acidic preservatives, and/or other preservatives. Exemplary antioxidants include, but are not limited to, alpha tocopherol, ascorbic acid, acorbaryl palmitate, butylated hydroxyanisole, butylated hydroxytoluene, monothioglycerol, potassium metabisulfite, propionic acid, propyl gallate, sodium ascorbate, sodium bisulfite, sodium metabisulfite, and/or sodium sulfite. Exemplary chelating agents include ethylenediaminetetraacetic acid (EDTA), citric acid monohydrate, disodium edetate, dipotassium edetate, edetic acid, fumaric acid, malic acid, phosphoric acid, sodium edetate, tartaric acid, and/or trisodium edetate. Exemplary antimicrobial preservatives include, but are not limited to, benzalkonium chloride, benzethonium chloride, benzyl alcohol, bronopol, cetrimide, cetylpyridinium chloride, chlorhexidine, chlorobutanol, chlorocresol, chloroxylenol, cresol, ethyl alcohol, glycerin, hexetidine, imidurea, phenol, phenoxyethanol, phenylethyl alcohol,
phenylmercuric nitrate, propylene glycol, and/or thimerosal. Exemplary antifungal preservatives
include, but are not limited to, butyl paraben, methyl paraben, ethyl paraben, propyl paraben,
benzoic acid, hydroxybenzoic acid, potassium benzoate, potassium sorbate, sodium benzoate,
sodium propionate, and/or sorbic acid. Exemplary alcohol preservatives include, but are not limited
to, ethanol, polyethylene glycol, phenol, phenolic compounds, bisphenol, chlorobutanol,
hydroxybenzoate, and/or phenylethyl alcohol. Exemplary acidic preservatives include, but are not
limited to, vitamin A, vitamin C, vitamin E, beta-carotene, citric acid, acetic acid, dehydroacetic
acid, ascorbic acid, sorbic acid, and/or phytic acid. Other preservatives include, but are not limited
to, tocopherol, tocopherol acetate, dextrose x mesylate, cetrimide, butylated hydroxyanisol (BHA),
butylated hydroxytoluene (BHT), ethylenediamine, sodium lauryl sulfate (SLS), sodium lauryl
ether sulfate (SLES), sodium bisulfite, sodium metabisulfite, potassium sulfite, potassium
metabisulfite, Glydant Plus®, Phenonip®, methylparaben, Germall®115, Germaben®II, Neolone™,
Kathon™, and/or Euxyl®.

Exemplary buffering agents include, but are not limited to, citrate buffer solutions,
acetate buffer solutions, phosphate buffer solutions, ammonium chloride, calcium carbonate, calcium
chloride, calcium citrate, calcium gluconate, calcium lactate, propanoic acid, calcium levulinate, pentanoic acid,
dibasic calcium phosphate, phosphoric acid, tribasic calcium phosphate, calcium hydroxide
phosphate, potassium acetate, potassium chloride, potassium gluconate, potassium mixtures, dibasic
potassium phosphate, monobasic potassium phosphate, potassium phosphate mixtures, sodium
acetate, sodium bicarbonate, sodium chloride, sodium citrate, sodium lactate, dibasic sodium
phosphate, monobasic sodium phosphate, sodium phosphate mixtures, tromethamine, magnesium
hydroxide, aluminum hydroxide, alginic acid, pyrogen-free water, isotonic saline, Ringer's solution,
ethyl alcohol, etc., and/or combinations thereof.

Exemplary lubricating agents include, but are not limited to, magnesium stearate, calcium
stearate, stearic acid, silica, talc, malt, glyceryl behenate, hydrogenated vegetable oils, polyethylene
glycol, sodium benzoate, sodium acetate, sodium chloride, leucine, magnesium lauryl sulfate,
sodium lauryl sulfate, etc., and combinations thereof.

Exemplary oils include, but are not limited to, almond, apricot kernel, avocado, babassu,
bergamot, black current seed, borage, cade, camomile, canola, caraway, carnauba, castor, cinnamon,
cocoa butter, coconut, cod liver, coffee, corn, cotton seed, emu, eucalyptus, evening primrose, fish,
flaxseed, geraniol, gourd, grape seed, hazel nut, hyssop, isopropyl myristate, jojoba, kukui nut,
lavandin, lavender, lemon, litsea cubeba, macadamia nut, mallow, mango seed, meadowfoam seed,
mink, nutmeg, olive, orange, orange roughy, palm, palm kernel, peach kernel, peanut, poppy seed,
pumpkin seed, rapeseed, rice bran, rosemary, safflower, sandalwood, sasquana, savoury, sea
buckthorn, sesame, shea butter, silicone, soybean, sunflower, tea tree, thistle, tsubaki, vetiver,
walnut, and wheat germ oils. Exemplary oils include, but are not limited to, butyl stearate, caprylic
triglyceride, capric triglyceride, cyclomethicone, diethicone 360, isopropyl myristate, mineral oil, octyldecanol, oleyl alcohol, silicone oil, and/or combinations thereof.

[000442] Liquid dosage forms for oral and parenteral administration include, but are not limited to,
pharmaceutically acceptable emulsions, microemulsions, solutions, suspensions, syrups, and/or
elixirs. In addition to active ingredients, liquid dosage forms may comprise inert diluents commonly
used in the art such as, for example, water or other solvents, solubilizing agents and emulsifiers such
as ethyl alcohol, isopropyl alcohol, ethyl carbonate, ethyl acetate, benzyl alcohol, benzyl benzoate,
propylene glycol, 1,3-butylen glycol, dimethylformamide, oils (in particular, cottonseed, groundnut,
corn, germ, olive, castor, and sesame oils), glycerol, tetrahydrofurfuryl alcohol, polyethylene glycols
and fatty acid esters of sorbitan, and mixtures thereof. Besides inert diluents, oral compositions can
include adjuvants such as wetting agents, emulsifying and suspending agents, sweetening, flavoring,
and/or perfuming agents. In certain embodiments for parenteral administration, compositions are
mixed with solubilizing agents such as Cremophor®, alcohols, oils, modified oils, glycols,
polysorbates, cyclodextrins, polymers, and/or combinations thereof.

[000443] Injectable preparations, for example, sterile injectable aqueous or oleaginous suspensions
may be formulated according to the known art using suitable dispersing agents, wetting agents,
and/or suspending agents. Sterile injectable preparations may be sterile injectable solutions,
suspensions, and/or emulsions in nontoxic parenterally acceptable diluents and/or solvents, for
example, as a solution in 1,3-butane diol. Among the acceptable vehicles and solvents that may be
employed are water, Ringer's solution, U.S.P., and isotonic sodium chloride solution. Sterile, fixed
oils are conventionally employed as a solvent or suspending medium. For this purpose any bland
fixed oil can be employed including synthetic mono- or diglycerides. Fatty acids such as oleic acid
can be used in the preparation of injectables.

[000444] Injectable formulations can be sterilized, for example, by filtration through a bacterial-
retaining filter, and/or by incorporating sterilizing agents in the form of sterile solid compositions
which can be dissolved or dispersed in sterile water or other sterile injectable medium prior to use.

85
In order to prolong the effect of an active ingredient, it is often desirable to slow the absorption of the active ingredient from subcutaneous or intramuscular injection. This may be accomplished by the use of a liquid suspension of crystalline or amorphous material with poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution which, in turn, may depend upon crystal size and crystalline form. Alternatively, delayed absorption of a parenterally administered drug form is accomplished by dissolving or suspending the drug in an oil vehicle. Injectable depot forms are made by forming microencapsule matrices of the drug in biodegradable polymers such as polylactide-polyglycolide. Depending upon the ratio of drug to polymer and the nature of the particular polymer employed, the rate of drug release can be controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot injectable formulations are prepared by entrapping the drug in liposomes or microemulsions which are compatible with body tissues.

Compositions for rectal or vaginal administration are typically suppositories which can be prepared by mixing compositions with suitable non-irritating excipients such as cocoa butter, polyethylene glycol or a suppository wax which are solid at ambient temperature but liquid at body temperature and therefore melt in the rectum or vaginal cavity and release the active ingredient. Solid dosage forms for oral administration include capsules, tablets, pills, powders, and granules. In such solid dosage forms, an active ingredient is mixed with at least one inert, pharmaceutically acceptable excipient such as sodium citrate or dicalcium phosphate and/or fillers or extenders (e.g. starches, lactose, sucrose, glucose, mannitol, and silicic acid), binders (e.g. carboxymethylcellulose, alginates, gelatin, polyvinylpyrrolidinone, sucrose, and acacia), humectants (e.g. glycerol), disintegrating agents (e.g. agar, calcium carbonate, potato or tapioca starch, alginic acid, certain silicates, and sodium carbonate), solution retarding agents (e.g. paraffin), absorption accelerators (e.g. quaternary ammonium compounds), wetting agents (e.g. cetyl alcohol and glycerol monostearate), absorbents (e.g. kaolin and bentonite clay), and lubricants (e.g. talc, calcium stearate, magnesium stearate, solid polyethylene glycols, sodium lauryl sulfate), and mixtures thereof. In the case of capsules, tablets and pills, the dosage form may comprise buffering agents.

Solid compositions of a similar type may be employed as fillers in soft and hard-filled gelatin capsules using such excipients as lactose or milk sugar as well as high molecular weight polyethylene glycols and the like. Solid dosage forms of tablets, dragees, capsules, pills, and granules can be prepared with coatings and shells such as enteric coatings and other coatings well known in the pharmaceutical formulating art. They may optionally comprise opacifying agents and
can be of a composition that they release the active ingredient(s) only, or preferentially, in a certain part of the intestinal tract, optionally, in a delayed manner. Examples of embedding compositions which can be used include polymeric substances and waxes. Solid compositions of a similar type may be employed as fillers in soft and hard-filled gelatin capsules using such excipients as lactose or milk sugar as well as high molecular weight polyethylene glycols and the like.

**[000448]** Dosage forms for topical and/or transdermal administration of a composition may include ointments, pastes, creams, lotions, gels, powders, solutions, sprays, inhalants and/or patches. Generally, an active ingredient is admixed under sterile conditions with a pharmaceutically acceptable excipient and/or any needed preservatives and/or buffers as may be required.

Additionally, the present invention contemplates the use of transdermal patches, which often have the added advantage of providing controlled delivery of a compound to the body. Such dosage forms may be prepared, for example, by dissolving and/or dispensing the compound in the proper medium. Alternatively or additionally, rate may be controlled by either providing a rate controlling membrane and/or by dispersing the compound in a polymer matrix and/or gel.

**[000449]** Suitable devices for use in delivering intradermal pharmaceutical compositions described herein include short needle devices such as those described in U.S. Patents 4,886,499; 5,190,521; 5,328,483; 5,527,288; 4,270,537; 5,015,235; 5,141,496; and 5,417,662. Intradermal compositions may be administered by devices which limit the effective penetration length of a needle into the skin, such as those described in PCT publication WO 99/34850 and functional equivalents thereof. Jet injection devices which deliver liquid compositions to the dermis via a liquid jet injector and/or via a needle which pierces the stratum corneum and produces a jet which reaches the dermis are suitable. Jet injection devices are described, for example, in U.S. Patents 5,480,381; 5,599,302; 5,334,144; 5,993,412; 5,649,912; 5,569,189; 5,704,912; 5,338,351; 5,893,397; 5,466,220; 5,339,163; 5,312,335; 5,503,627; 5,064,413; 5,520,639; 4,596,556; 4,790,824; 4,941,880; 4,940,460; and PCT publications WO 97/37705 and WO 97/13537. Ballistic powder/particle delivery devices which use compressed gas to accelerate vaccine in powder form through the outer layers of the skin to the dermis are suitable. Alternatively or additionally, conventional syringes may be used in the classical mantoux method of intradermal administration.

**[000450]** Formulations suitable for topical administration include, but are not limited to, liquid and/or semi liquid preparations such as liniments, lotions, oil in water and/or water in oil emulsions such as creams, ointments and/or pastes, and/or solutions and/or suspensions. Topically-administrable formulations may, for example, comprise from about 1% to about 10% (w/w) active
ingredient, although the concentration of active ingredient may be as high as the solubility limit of
the active ingredient in the solvent. Formulations for topical administration may further comprise
one or more of the additional ingredients described herein.

[000451] A pharmaceutical composition may be prepared, packaged, and/or sold in a formulation
suitable for pulmonary administration via the buccal cavity. Such a formulation may comprise dry
particles which comprise the active ingredient and which have a diameter in the range from about
0.5 nm to about 7 nm or from about 1 nm to about 6 nm. Such compositions are suitably in the form
of dry powders for administration using a device comprising a dry powder reservoir to which a
stream of propellant may be directed to disperse the powder and/or using a self propelling
solvent/powder dispensing container such as a device comprising the active ingredient dissolved
and/or suspended in a low-boiling propellant in a sealed container. Such powders comprise particles
wherein at least 98% of the particles by weight have a diameter greater than 0.5 nm and at least 95%
of the particles by number have a diameter less than 7 nm. Alternatively, at least 95% of the
particles by weight have a diameter greater than 1 nm and at least 90% of the particles by number
have a diameter less than 6 nm. Dry powder compositions may include a solid fine powder diluent
such as sugar and are conveniently provided in a unit dose form.

[000452] Low boiling propellants generally include liquid propellants having a boiling point of
below 65 °F at atmospheric pressure. Generally the propellant may constitute 50% to 99.9% (w/w)
of the composition, and active ingredient may constitute 0.1% to 20% (w/w) of the composition. A
propellant may further comprise additional ingredients such as a liquid non-ionic and/or solid
anionic surfactant and/or a solid diluent (which may have a particle size of the same order as
particles comprising the active ingredient).

[000453] Pharmaceutical compositions formulated for pulmonary delivery may provide an active
ingredient in the form of droplets of a solution and/or suspension. Such formulations may be
prepared, packaged, and/or sold as aqueous and/or dilute alcoholic solutions and/or suspensions,
optionally sterile, comprising active ingredient, and may conveniently be administered using any
nebulization and/or atomization device. Such formulations may further comprise one or more
additional ingredients including, but not limited to, a flavoring agent such as saccharin sodium, a
volatile oil, a buffering agent, a surface active agent, and/or a preservative such as
methylhydroxybenzoate. Droplets provided by this route of administration may have an average
diameter in the range from about 0.1 nm to about 200 nm.
Formulations described herein as being useful for pulmonary delivery are useful for intranasal delivery of a pharmaceutical composition. Another formulation suitable for intranasal administration is a coarse powder comprising the active ingredient and having an average particle from about 0.2 µm to 500 µm. Such a formulation is administered in the manner in which snuff is taken, i.e. by rapid inhalation through the nasal passage from a container of the powder held close to the nose.

Formulations suitable for nasal administration may, for example, comprise from about as little as 0.1% (w/w) and as much as 100% (w/w) of active ingredient, and may comprise one or more of the additional ingredients described herein. A pharmaceutical composition may be prepared, packaged, and/or sold in a formulation suitable for buccal administration. Such formulations may, for example, be in the form of tablets and/or lozenges made using conventional methods, and may, for example, 0.1% to 20% (w/w) active ingredient, the balance comprising an orally dissolvable and/or degradable composition and, optionally, one or more of the additional ingredients described herein. Alternately, formulations suitable for buccal administration may comprise a powder and/or an aerosolized and/or atomized solution and/or suspension comprising active ingredient. Such powdered, aerosolized, and/or aerosolized formulations, when dispersed, may have an average particle and/or droplet size in the range from about 0.1 nm to about 200 nm, and may further comprise one or more of any additional ingredients described herein.

A pharmaceutical composition may be prepared, packaged, and/or sold in a formulation suitable for ophthalmic administration. Such formulations may, for example, be in the form of eye drops including, for example, a 0.1/1.0% (w/w) solution and/or suspension of the active ingredient in an aqueous or oily liquid excipient. Such drops may further comprise buffering agents, salts, and/or one or more other of any additional ingredients described herein. Other ophthalmically-administrable formulations which are useful include those which comprise the active ingredient in microcrystalline form and/or in a liposomal preparation. Ear drops and/or eye drops are contemplated as being within the scope of this invention.

General considerations in the formulation and/or manufacture of pharmaceutical agents may be found, for example, in Remington: The Science and Practice of Pharmacy 21st ed., Lippincott Williams & Wilkins, 2005 (incorporated herein by reference).

The present invention provides methods comprising administering modified mRNAs and their encoded proteins or complexes in accordance with the invention to a subject in need thereof. Nucleic acids, proteins or complexes, or pharmaceutical, imaging, diagnostic, or prophylactic
compositions thereof, may be administered to a subject using any amount and any route of administration effective for preventing, treating, diagnosing, or imaging a disease, disorder, and/or condition (e.g., a disease, disorder, and/or condition relating to working memory deficits). The exact amount required will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the disease, the particular composition, its mode of administration, its mode of activity, and the like. Compositions in accordance with the invention are typically formulated in dosage unit form for ease of administration and uniformity of dosage. It will be understood, however, that the total daily usage of the compositions of the present invention will be decided by the attending physician within the scope of sound medical judgment. The specific therapeutically effective, prophylactically effective, or appropriate imaging dose level for any particular patient will depend upon a variety of factors including the disorder being treated and the severity of the disorder; the activity of the specific compound employed; the specific composition employed; the age, body weight, general health, sex and diet of the patient; the time of administration, route of administration, and rate of excretion of the specific compound employed; the duration of the treatment; drugs used in combination or coincidental with the specific compound employed; and like factors well known in the medical arts.

[000459] In certain embodiments, compositions in accordance with the present invention may be administered at dosage levels sufficient to deliver from about 0.0001 mg/kg to about 100 mg/kg, from about 0.01 mg/kg to about 50 mg/kg, from about 0.1 mg/kg to about 40 mg/kg, from about 0.5 mg/kg to about 30 mg/kg, from about 0.01 mg/kg to about 10 mg/kg, from about 0.1 mg/kg to about 10 mg/kg, or from about 1 mg/kg to about 25 mg/kg, of subject body weight per day, one or more times a day, to obtain the desired therapeutic, diagnostic, prophylactic, or imaging effect. The desired dosage may be delivered three times a day, two times a day, once a day, every other day, every third day, every week, every two weeks, every three weeks, or every four weeks. In certain embodiments, the desired dosage may be delivered using multiple administrations (e.g., two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, or more administrations).

[000460] Modified nucleic acid molecules or complexes may be used in combination with one or more other therapeutic, prophylactic, diagnostic, or imaging agents. By "in combination with," it is not intended to imply that the agents must be administered at the same time and/or formulated for delivery together, although these methods of delivery are within the scope of the present disclosure. Compositions can be administered concurrently with, prior to, or subsequent to, one or more other desired therapeutics or medical procedures. In general, each agent will be administered at a dose
and/or on a time schedule determined for that agent. In some embodiments, the present disclosure encompasses the delivery of pharmaceutical, prophylactic, diagnostic, or imaging compositions in combination with agents that may improve their bioavailability, reduce and/or modify their metabolism, inhibit their excretion, and/or modify their distribution within the body.

[000461] The pharmaceutical compositions described herein include lipidoid-based formulations allowing for the localized and systemic delivery of mmRNA. The synthesis of lipidoids has been extensively described and formulations containing these compounds are particularly suited for delivery of polynucleotides (see Mahon et al, Bioconjug Chem. 2010 21:1448-1454; Schroeder et al, J Intern Med. 2010 267:9-21; Akinc et al, Nat Biotechnol. 2008 26:561-569; Love et al, Proc Natl Acad Sci U S A. 2010 107:1864-1869; Siegwart et al, Proc Natl Acad Sci U S A. 2011 108:12996-3001; all of which are incorporated herein in their entireties). While these lipidoids have been used to effectively deliver double stranded small interfering RNA molecules in rodents and non-human primates (see Akinc et al., Nat Biotechnol. 2008 26:561-569; Frank-Kamenetsky et al., Proc Natl Acad Sci U S A. 2008 105:1 1915-1 1920; Akinc et al, Mol Ther. 2009 17:872-879; Love et al, Proc Natl Acad Sci U S A. 2010 107: 1864-1869; Leuschner et al, Nat Biotechnol 2011 29:1005-1010; all of which is incorporated herein in their entirety), the present disclosure describes their formulation and use in delivering single stranded mmRNA. Complexes, micelles, liposomes or particles can be prepared containing these lipidoids and therefore, result in an effective delivery of mmRNA, as judged by the production of an encoded protein, following the injection of an mmRNA-formulated lipidoids via localized and systemic routes of administration. Modified mRNA-lipidoid complexes can be administered by various means including, but not limited to, intravenous, intramuscular, or subcutaneous routes.

[000462] In vivo delivery of nucleic acids may be affected by many parameters, including, but not limited to, the formulation composition, nature of particle PEGylation, degree of drug loading, oligonucleotide to lipid ratio, and biophysical parameters such as particle size (Akinc et al., Mol Ther. 2009 17:872-879; herein incorporated by reference in its entirety). As an example, small changes in the anchor chain length of poly(ethylene glycol) (PEG) lipids may result in significant effects on in vivo efficacy. Formulations with the different lipidoids, including, but not limited to, 98N12-5, C12-200, and MD1, can be tested for in vivo activity. The lipidoid referred to herein as "98N12-5" is disclosed by Akinc et al, Mol Ther. 2009 17:872-879; is incorporated by reference in its entirety (see Figure 1). The lipidoid referred to herein as "C12-200" is disclosed by Love et al, Proc Natl Acad Sci U S A. 2010 107:1864-1869 (see Figure 1) and Liu and Huang, Molecular
Therapy. 2010 669-670 (see Figure 1); both of which are herein incorporated by reference in their entirety. Initial mmRNA-lipidoid formulations can include particles comprising either 3 or 4 or more components in addition to mmRNA. As an example, mmRNA formulations with certain lipidoids, include, but are not limited to, 98N12-5 and may contain 42% lipidoid, 48% cholesterol and 10% PEG (Cl4 alkyl chain length). As another example, mmRNA formulations with certain lipidoids, include, but are not limited to, C12-200 and may contain 50% lipidoid, 10%> disteroylphosphatidyl choline, 38.5% cholesterol, and 1.5% PEG-DMG.

In one embodiment of the present invention, systemic intravenous administration of the formulation composition of the present invention may target the liver. For example, a final optimized intravenous formulation using siRNA, and comprising a lipid molar composition of 42% 98N12-5, 48% cholesterol, and 10% PEG-lipid with a final weight ratio of about 7.5 to 1 total lipid to siRNA, and a C14 alkyl chain length on the PEG lipid, with a mean particle size of roughly 50-60 nm, may result in the distribution of the formulation to be greater than 90% to the liver.(see, Akinc et al., Mol Ther. 2009 17:872-879; herein incorporated in its entirety). In another example, an intravenous formulation using a C12-200 (see US provisional application 61/175,770 and published international application WO2010129709 herein incorporated by reference it their entirety) lipidoid may have a molar ratio of 50/10/38.5/1.5 of C12-200/disteroylphosphatidyl choline/cholesterol/PEG-DMG, with a weight ratio of 7 to 1 total lipid to siRNA = 7:1, and a mean particle size of 80nm may be effective to deliver siRNA to hepatocytes and may result in gene silencing (see, Love et al, Proc Natl Acad Sci U S A. 2010 107:1864-1869 herein incorporated by reference). In another embodiment, an MD1 lipidoid-containing formulation may be used to effectively deliver siRNA to hepatocytes in vivo (see Oligonucleotide Therapeutics Society Meeting, Copenhagen, Denmark September 10-12, 2011 and Alnylam Press Release Sept 12, 2011; both of which are incorporated herein by reference in their entirety)

The characteristics of optimized lipidoid formulations for intramuscular or subcutaneous routes may vary significantly depending on the target cell type and the ability of formulations to diffuse through the extracellular matrix into the blood stream. While a particle size of less than 150nm may be desired for effective hepatocyte delivery due to the size of the endothelial fenestrae (see, Akinc et al, Mol Ther. 2009 17:872-879 herein incorporated by reference), use of lipidoid oligonucleotides to deliver the formulation to other cells types including, but not limited to, endothelial cells, myeloid cells, and muscle cells may not be similarly size-limited. Use of lipidoid formulations to deliver siRNA in vivo to other non-hepatocyte cells such as myeloid cells and
endothelium has also been reported (see Akinc et al., Nat Biotechnol. 2008 26:561-569; Leuschner et al., Nat Biotechnol. 201 1 29:1005-1010; 8th International Judah Folkman Conference, Cambridge, MA October 8-9, 2010 and Alnylam Press Release Oct 8, 2010; herein incorporated by reference in its entirety). In one aspect, effective delivery to myeloid cells, such as monocytes, lipidoid formulations may have a similar component molar ratio. Different ratios of lipidoids and other components including, but not limited to, disteroylphosphatidyl choline, cholesterol and PEG-DMG, may be used to optimize the formulation of the modified nucleic acid molecule for delivery to different cell types including, but not limited to, hepatocytes, myeloid cells, muscle cells, etc. For example, the component molar ratio may include, but is not limited to, 50% C12-200, 10% disteroylphosphatidyl choline, 38.5% cholesterol, and %1.5 PEG-DMG (see Leuschner et al, Nat Biotechnol 201 1 29:1005-1010; herein incorporated by reference in its entirety). The use of lipidoid formulations for the localized delivery of nucleic acids to cells (such as, but not limited to, adipose cells and muscle cells) via either subcutaneous or intramuscular delivery, may also not require all of the formulation components which may be required for systemic delivery, and as such may comprise the lipidoid and the mmRNA.

[000465] In a further embodiment, combinations of different lipidoids may be used to improve the efficacy of mmRNA-directed protein production as the lipidoids may be able to synergize for improved siRNA delivery and silencing (Whitehead et al., Mol Ther. 201 1, 19:1688-1694, herein incorporated by reference in its entirety).

[000466] In one embodiment, the pharmaceutical compositions described herein can include one or more pharmaceutically acceptable carriers.

[000467] When administered to a subject the pharmaceutical compositions described herein may provide proteins which have been generated from modified mRNAs. Pharmaceutical compositions may optionally comprise one or more additional therapeutically active substances. In accordance with some embodiments, a method of administering pharmaceutical compositions comprising one or more proteins to be delivered to a subject in need thereof is provided. In some embodiments, compositions are administered to human subjects. In a further embodiment, the compositions are administered to a subject who is a patient.

Properties of the Pharmaceutical Compositions

[000468] The pharmaceutical compositions described herein can be characterized by one or more of the following properties:
Bioavailability

[000469] The modified nucleic acid molecules, when formulated into a composition with a delivery agent as described herein, can exhibit an increase in bioavailability as compared to a composition lacking a delivery agent as described herein. As used herein, the term "bioavailability" refers to the systemic availability of a given amount of a modified nucleic acid molecule administered to a mammal. Bioavailability can be assessed by measuring the area under the curve (AUC) or the maximum serum or plasma concentration \( C_{\text{max}} \) of the unchanged form of a compound following administration of the compound to a mammal. AUC is a determination of the area under the curve plotting the serum or plasma concentration of a compound along the ordinate (Y-axis) against time along the abscissa (X-axis). Generally, the AUC for a particular compound can be calculated using methods known to those of ordinary skill in the art and as described in G. S. Banker, Modern Pharmaceutics, Drugs and the Pharmaceutical Sciences, v. 72, Marcel Dekker, New York, Inc., 1996, herein incorporated by reference.

[000470] The \( C_{\text{max}} \) value is the maximum concentration of the compound achieved in the serum or plasma of a mammal following administration of the compound to the mammal. The \( C_{\text{max}} \) value of a particular compound can be measured using methods known to those of ordinary skill in the art. The phrases "increasing bioavailability" or "improving the pharmacokinetics," as used herein mean that the systemic availability of a first modified nucleic acid molecule, measured as AUC, \( C_{\text{max}} \), or \( C_{\text{m}} \), in a mammal is greater, when co-administered with a delivery agent as described herein, than when such co-administration does not take place. In some embodiments, the bioavailability of the modified nucleic acid molecule can increase by at least about 2%, at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or about 100%.

Therapeutic Window

[000471] The modified nucleic acid molecules, when formulated into a composition with a delivery agent as described herein, can exhibit an increase in the therapeutic window of the administered modified nucleic acid molecule composition as compared to the therapeutic window of the administered modified nucleic acid molecule composition lacking a delivery agent as described herein. As used herein "therapeutic window" refers to the range of plasma concentrations, or the
range of levels of therapeutically active substance at the site of action, with a high probability of eliciting a therapeutic effect. In some embodiments, the therapeutic window of the modified nucleic acid molecule when co-administered with a delivery agent as described herein can increase by at least about 2%, at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or about 100%.

Volume of Distribution

[000472] The modified nucleic acid molecules, when formulated into a composition with a delivery agent as described herein, can exhibit an improved volume of distribution (V<sub>dist</sub>), e.g., reduced or targeted, relative to a modified nucleic acid molecule composition lacking a delivery agent as described herein. The volume of distribution (V<sub>dist</sub>) relates the amount of the drug in the body to the concentration of the drug in the blood or plasma. As used herein, the term "volume of distribution" refers to the fluid volume that would be required to contain the total amount of the drug in the body at the same concentration as in the blood or plasma: V<sub>dist</sub> equals the amount of drug in the body/concentration of drug in blood or plasma. For example, for a 10 mg dose and a plasma concentration of 10 mg/L, the volume of distribution would be 1 liter. The volume of distribution reflects the extent to which the drug is present in the extravascular tissue. A large volume of distribution reflects the tendency of a compound to bind to the tissue components compared with plasma protein binding. In a clinical setting, V<sub>dist</sub> can be used to determine a loading dose to achieve a steady state concentration. In some embodiments, the volume of distribution of the modified nucleic acid molecule when co-administered with a delivery agent as described herein can decrease at least about 2%, at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%.

Kits

[000473] The invention provides a variety of kits for conveniently and/or effectively carrying out methods of the present invention. Typically kits will comprise sufficient amounts and/or numbers of components to allow a user to perform multiple treatments of a subject(s) and/or to perform multiple experiments.
In one aspect, the present invention provides kits for protein production, comprising a first isolated nucleic acid comprising a translatable region and a nucleic acid modification, wherein the nucleic acid may be capable of evading an innate immune response of a cell into which the first isolated nucleic acid may be introduced, and packaging and instructions. The kit may further comprise a delivery agent to form a formulation composition. The delivery composition may comprise a lipidoid. The lipidoid may be selected from the group consisting of CI2-200, 98N12-5 and MD1.

In one aspect, the present invention provides kits for protein production, comprising: a first isolated nucleic acid comprising a translatable region, provided in an amount effective to produce a desired amount of a protein encoded by the translatable region when introduced into a target cell; a second nucleic acid comprising an inhibitory nucleic acid, provided in an amount effective to substantially inhibit the innate immune response of the cell; and packaging and instructions.

In one aspect, the present invention provides kits for protein production, comprising a first isolated nucleic acid comprising a translatable region and a nucleoside modification, wherein the nucleic acid exhibits reduced degradation by a cellular nuclease, and packaging and instructions.

In one aspect, the present invention provides kits for protein production, comprising a first isolated nucleic acid comprising a translatable region and at least two different nucleoside modifications, wherein the nucleic acid exhibits reduced degradation by a cellular nuclease, and packaging and instructions.

In one aspect, the present invention provides kits for protein production, comprising a first isolated nucleic acid comprising a translatable region and at least one nucleoside modification, wherein the nucleic acid exhibits reduced degradation by a cellular nuclease; a second nucleic acid comprising an inhibitory nucleic acid; and packaging and instructions.

In some embodiments, the first isolated nucleic acid comprises messenger RNA (mRNA). In some embodiments the mRNA comprises at least one nucleoside selected from the group consisting of pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 3-methyluridine, 5-carboxymethyluridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-
deaza-pseudouridine, dihydouridine, dihydropseudouridine, 2-thio-dihydouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, and 4-methoxy-2-thio-pseudouridine.

[000480] In some embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolocytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, and 4-methoxy-1-methyl-pseudoisocytidine.

[000481] In some embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of 2-aminopurine, 2, 6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, and 2-methoxy-adenine.

[000482] In some embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

[000483] In another aspect, the disclosure provides compositions for protein production, comprising a first isolated nucleic acid comprising a translatable region and a nucleoside modification, wherein the nucleic acid exhibits reduced degradation by a cellular nuclease, and a mammalian cell suitable for translation of the translatable region of the first nucleic acid.
Definitions

[000484] At various places in the present specification, substituents of compounds of the present disclosure are disclosed in groups or in ranges. It is specifically intended that the present disclosure include each and every individual subcombination of the members of such groups and ranges. For example, the term "Ci_6 alkyl" is specifically intended to individually disclose methyl, ethyl, C_3 alkyl, C_4 alkyl, C_5 alkyl, and C_6 alkyl.

[000485] Animal: As used herein, the term "animal" refers to any member of the animal kingdom. In some embodiments, "animal" refers to humans at any stage of development. In some embodiments, "animal" refers to non-human animals at any stage of development. In certain embodiments, the non-human animal is a mammal (e.g., a rodent, a mouse, a rat, a rabbit, a monkey, a dog, a cat, a sheep, cattle, a primate, or a pig). In some embodiments, animals include, but are not limited to, mammals, birds, reptiles, amphibians, fish, and worms. In some embodiments, the animal is a transgenic animal, genetically-engineered animal, or a clone.

[000486] Approximately: As used herein, the term "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term "approximately" or "about" refers to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100%, of a possible value).

[000487] Associated with: As used herein, the terms "associated with," "conjugated," "linked," "attached," and "tethered," when used with respect to two or more moieties, means that the moieties are physically associated or connected with one another, either directly or via one or more additional moieties that serves as a linking agent, to form a structure that is sufficiently stable so that the moieties remain physically associated under the conditions in which the structure is used, e.g., physiological conditions. An "association" need not be strictly through direct covalent chemical bonding. It may also suggest ionic or hydrogen bonding or a hybridization based connectivity sufficiently stable such that the "associated" entities remain physically associated.

[000488] Biologically active: As used herein, the phrase "biologically active" refers to a characteristic of any substance that has activity in a biological system and/or organism. For instance, a substance that, when administered to an organism, has a biological affect on that organism, is considered to be biologically active. In particular embodiments, a nucleic acid
molecule of the present invention may be considered biologically active if even a portion of the nucleic acid molecule is biologically active or mimics an activity considered biologically relevant.

[000489] **Chemical terms:** As used herein, the term "alkyl" is meant to refer to a saturated hydrocarbon group which is straight-chained or branched. Example alkyl groups include methyl (Me), ethyl (Et), propyl (e.g., n-propyl and isopropyl), butyl (e.g., n-butyl, isobutyl, t-butyl), pentyl (e.g., n-pentyl, isopentyl, neopentyl), and the like. An alkyl group can contain from 1 to about 20, from 2 to about 20, from 1 to about 12, from 1 to about 8, from 1 to about 6, from 1 to about 4, or from 1 to about 3 carbon atoms.

[000490] As used herein, "alkenyl" refers to an alkyl group having one or more double carbon-carbon bonds. Example alkenyl groups include ethenyl, propenyl, and the like.

[000491] As used herein, "alkoxy" refers to an -O-alkyl group. Example alkoxy groups include methoxy, ethoxy, propoxy (e.g., n-propoxy and isopropoxy), t-butoxy, and the like.

[000492] As used herein, "alkenyl" refers to an alkyl, as defined above, containing at least one double bond between adjacent carbon atoms. Alkenyls include both cis and trans isomers. Representative straight chain and branched alkenyls include ethylenyl, propylenyl, 1-butenyl, 2-butenyl, isobutylenyl, 1-pentenyl, 2-pentenyl, 3-methyl-1-butenyl, 2-methyl-2-butenyl, 2,3-dimethyl-2-butenyl, and the like.

[000493] As used herein, "alkynyl" refers to an alkyl group having one or more triple carbon-carbon bonds. Example alkynyl groups include ethynyl, propynyl, and the like.

[000494] As used herein, "aryl" refers to monocyclic or polycyclic (e.g., having 2, 3 or 4 fused rings) aromatic hydrocarbons such as, for example, phenyl, naphthyl, anthracenyl, phenanthrenyl, indanyl, indenyl, and the like. In some embodiments, aryl groups have from 6 to about 20 carbon atoms.

[000495] As used herein, "halo" or "halogen" includes fluoro, chloro, bromo, and iodo.

[000496] **Compound:** As used herein, the term "compound," is meant to include all stereoisomers, geometric isomers, tautomers, and isotopes of the structures depicted.

[000497] The compounds described herein can be asymmetric (e.g., having one or more stereocenters). All stereoisomers, such as enantiomers and diastereomers, are intended unless otherwise indicated. Compounds of the present disclosure that contain asymmetrically substituted carbon atoms can be isolated in optically active or racemic forms. Methods on how to prepare optically active forms from optically active starting materials are known in the art, such as by resolution of racemic mixtures or by stereoselective synthesis. Many geometric isomers of olefins,
C=N double bonds, and the like can also be present in the compounds described herein, and all such stable isomers are contemplated in the present disclosure. Cis and trans geometric isomers of the compounds of the present disclosure are described and may be isolated as a mixture of isomers or as separated isomeric forms.

[000498] Compounds of the present disclosure also include tautomeric forms. Tautomeric forms result from the swapping of a single bond with an adjacent double bond and the concomitant migration of a proton. Tautomeric forms include prototropic tautomers which are isomeric protonation states having the same empirical formula and total charge. Examples prototropic tautomers include ketone - enol pairs, amide - imidic acid pairs, lactam - lactim pairs, amide - imidic acid pairs, enamine - imine pairs, and annular forms where a proton can occupy two or more positions of a heterocyclic system, such as, 1H- and 3H-imidazole, 1H-, 2H- and 4H- 1,2,4-triazole, 1H- and 2H- isoindole, and 1H- and 2H-pyrazole. Tautomeric forms can be in equilibrium or sterically locked into one form by appropriate substitution.

[000499] Compounds of the present disclosure also include all of the isotopes of the atoms occurring in the intermediate or final compounds. "Isotopes" refers to atoms having the same atomic number but different mass numbers resulting from a different number of neutrons in the nuclei. For example, isotopes of hydrogen include tritium and deuterium.

[000500] The compounds and salts of the present disclosure can be prepared in combination with solvent or water molecules to form solvates and hydrates by routine methods.

[000501] Conserved: As used herein, the term "conserved" refers to nucleotides or amino acid residues of a polynucleotide sequence or polypeptide sequence, respectively, that are those that occur unaltered in the same position of two or more sequences being compared. Nucleotides or amino acids that are relatively conserved are those that are conserved amongst more related sequences than nucleotides or amino acids appearing elsewhere in the sequences.

[000502] In some embodiments, two or more sequences are said to be "completely conserved" if they are 100% identical to one another. In some embodiments, two or more sequences are said to be "highly conserved" if they are at least 70% identical, at least 80% identical, at least 90% identical, or at least 95% identical to one another. In some embodiments, two or more sequences are said to be "highly conserved" if they are about 70% identical, about 80% identical, about 90% identical, about 95%, about 98%, or about 99% identical to one another. In some embodiments, two or more sequences are said to be "conserved" if they are at least 30% identical, at least 40% identical, at least 50% identical, at least 60% identical, at least 70% identical, at least 80% identical, at least 90%
identical, or at least 95% identical to one another. In some embodiments, two or more sequences are said to be "conserved" if they are about 30% identical, about 40% identical, about 50% identical, about 60% identical, about 70% identical, about 80% identical, about 90% identical, about 95% identical, about 98% identical, or about 99% identical to one another. Conservation of sequence may apply to the entire length of an oligonucleotide or polypeptide or may apply to a portion, region or feature thereof.

[000503] **Delivery:** As used herein, "delivery" refers to the act or manner of delivering a compound, substance, entity, moiety, cargo or payload.

[000504] **Delivery Agent:** As used herein, "delivery agent" refers to any substance which facilitates, at least in part, the in vivo delivery of a nucleic acid molecule to targeted cells.

[000505] **Detectable label:** As used herein, "detectable label" refers to one or more markers, signals, or moieties which are attached, incorporated or associated with another entity that is readily detected by methods known in the art including radiography, fluorescence, chemiluminescence, enzymatic activity, absorbance and the like. Detectable labels include radioisotopes, fluorophores, chromophores, enzymes, dyes, metal ions, ligands such as biotin, avidin, strepavidin and haptens, quantum dots, and the like. Detectable labels may be located at any position in the peptides or proteins disclosed herein. They may be within the amino acids, the peptides, or proteins, or located at the N- or C- termini.

[000506] **Expression:** As used herein, "expression" of a nucleic acid sequence refers to one or more of the following events: (1) production of an RNA template from a DNA sequence (e.g., by transcription); (2) processing of an RNA transcript (e.g., by splicing, editing, 5' cap formation, and/or 3' end processing); (3) translation of an RNA into a polypeptide or protein; and (4) post-translational modification of a polypeptide or protein.

[000507] **Formulation:** As used herein, a "formulation" includes at least a modified nucleic acid molecule and a delivery agent.

[000508] **Functional:** As used herein, a "functional" biological molecule is a biological molecule in a form in which it exhibits a property and/or activity by which it is characterized.

[000509] **Homology:** As used herein, the term "homology" refers to the overall relatedness between polymeric molecules, e.g. between nucleic acid molecules (e.g. DNA molecules and/or RNA molecules) and/or between polypeptide molecules. In some embodiments, polymeric molecules are considered to be "homologous" to one another if their sequences are at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, 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%, or at least 99% identical. In some embodiments, polymeric molecules are considered to be "homologous" to one another if their sequences are at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, 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%, or at least 99% similar. The term "homologous" necessarily refers to a comparison between at least two sequences (polynucleotide or polypeptide sequences).

[000510] In accordance with the invention, two polynucleotide sequences are considered to be homologous if the polypeptides they encode are at least about 50% identical, at least about 60% identical, at least about 70% identical, at least about 80% identical, or at least about 90% identical for at least one stretch of at least about 20 amino acids.

[000511] In some embodiments, homologous polynucleotide sequences are characterized by the ability to encode a stretch of at least 4-5 uniquely specified amino acids. For polynucleotide sequences less than 60 nucleotides in length, homology is determined by the ability to encode a stretch of at least 4-5 uniquely specified amino acids. In accordance with the invention, two protein sequences are considered to be homologous if the proteins are at least about 50% identical, at least about 60% identical, at least about 70% identical, at least about 80% identical, or at least about 90% identical for at least one stretch of at least about 20 amino acids.

[000512] Identity: As used herein, the term "identity" refers to the overall relatedness between polymeric molecules, e.g., between oligonucleotide molecules (e.g. DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Calculation of the percent identity of two polynucleotide sequences, for example, can be performed by aligning the two sequences for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second nucleic acid sequences for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In certain embodiments, the length of a sequence aligned for comparison purposes is at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or 100%, of the length of the reference sequence. The nucleotides at corresponding nucleotide positions are then compared. When a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which needs to be introduced for optimal alignment of the two sequences. The comparison of sequences and determination of percent identity between two sequences can be
accomplished using a mathematical algorithm. For example, the percent identity between two nucleotide sequences can be determined using methods such as those described in Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; each of which is incorporated herein by reference. For example, the percent identity between two nucleotide sequences can be determined using the algorithm of Meyers and Miller (CABIOS, 1989, 4:1 1-17), which has been incorporated into the ALIGN program (version 2.0) using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. The percent identity between two nucleotide sequences can, alternatively, be determined using the GAP program in the GCG software package using an NWSgpdna.CMP matrix. Methods commonly employed to determine percent identity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., SIAM J Applied Math., 48:1073 (1988); incorporated herein by reference.

Techniques for determining identity are codified in publicly available computer programs. Exemplary computer software to determine homology between two sequences include, but are not limited to, GCG program package, Devereux, J., et al., Nucleic Acids Research, 12(1), 387 (1984)), BLASTP, BLASTN, and FASTA Atschul, S. F. et al., J. Molec. Biol. 215, 403 (1990)).

[000513] Inhibit expression of a gene: As used herein, the phrase "inhibit expression of a gene" means to cause a reduction in the amount of an expression product of the gene. The expression product can be an RNA transcribed from the gene (e.g., an mRNA) or a polypeptide translated from an mRNA transcribed from the gene. Typically a reduction in the level of an mRNA results in a reduction in the level of a polypeptide translated therefrom. The level of expression may be determined using standard techniques for measuring mRNA or protein.

[000514] In vitro: As used herein, the term "in vitro" refers to events that occur in an artificial environment, e.g., in a test tube or reaction vessel, in cell culture, in a Petri dish, etc., rather than within an organism (e.g., animal, plant, or microbe).

[000515] In vivo: As used herein, the term "in vivo" refers to events that occur within an organism (e.g., animal, plant, or microbe or cell or tissue thereof).
Isolated: As used herein, the term "isolated" refers to a substance or entity that has been separated from at least some of the components with which it was associated (whether in nature or in an experimental setting). Isolated substances may have varying levels of purity in reference to the substances from which they have been associated. Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, or more of the other components with which they were initially associated. In some embodiments, isolated agents are more than about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% pure. As used herein, a substance is "pure" if it is substantially free of other components. Substantially isolated: By "substantially isolated" is meant that the compound is substantially separated from the environment in which it was formed or detected. Partial separation can include, for example, a composition enriched in the compound of the present disclosure. Substantial separation can include compositions containing at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 97%, or at least about 99% by weight of the compound of the present disclosure, or salt thereof. Methods for isolating compounds and their salts are routine in the art.

Modified: As used herein "modified" refers to a changed state or structure of a molecule of the invention. Molecules may be modified in many ways including chemically, structurally, and functionally. In one embodiment, the mRNA molecules of the present invention are modified by the introduction of non-natural nucleosides and/or nucleotides.

Naturally occurring: As used herein, "naturally occurring" means existing in nature without artificial aid.

Patient: As used herein, "patient" refers to a subject who may seek or be in need of treatment, requires treatment, is receiving treatment, will receive treatment, or a subject who is under care by a trained professional for a particular disease or condition.

Peptide: As used herein, "peptide" is less than or equal to 50 amino acids long, e.g., about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 amino acids long.

Prodrug: The present disclosure also includes prodrugs of the compounds described herein. As used herein, "prodrugs" refer to any substance, molecule or entity which is in a form predicate for that substance, molecule or entity to act as a therapeutic upon chemical or physical alteration. Prodrugs may by covalently bonded or sequestered in some way and which release or are converted into the active drug moiety prior to, upon or after administered to a mammalian subject.
Prodrugs can be prepared by modifying functional groups present in the compounds in such a way that the modifications are cleaved, either in routine manipulation or in vivo, to the parent compounds. Prodrugs include compounds wherein hydroxyl, amino, sulfhydryl, or carboxyl groups are bonded to any group that, when administered to a mammalian subject, cleaves to form a free hydroxyl, amino, sulfhydryl, or carboxyl group respectively. Preparation and use of prodrugs is discussed in T. Higuchi and V. Stella, "Pro-drugs as Novel Delivery Systems," Vol. 14 of the A.C.S. Symposium Series, and in Bioreversible Carriers in Drug Design, ed. Edward B. Roche, American Pharmaceutical Association and Pergamon Press, 1987, both of which are hereby incorporated by reference in their entirety.

**[000522] Pharmaceutically acceptable:** The phrase "pharmaceutically acceptable" is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

**[000523] Pharmaceutically acceptable salts:** The present disclosure also includes pharmaceutically acceptable salts of the compounds described herein. As used herein, "pharmaceutically acceptable salts" refers to derivatives of the disclosed compounds wherein the parent compound is modified by converting an existing acid or base moiety to its salt form. Examples of pharmaceutically acceptable salts include, but are not limited to, mineral or organic acid salts of basic residues such as amines; alkali or organic salts of acidic residues such as carboxylic acids; and the like. The pharmaceutically acceptable salts of the present disclosure include the conventional non-toxic salts of the parent compound formed, for example, from non-toxic inorganic or organic acids. The pharmaceutically acceptable salts of the present disclosure can be synthesized from the parent compound which contains a basic or acidic moiety by conventional chemical methods. Generally, such salts can be prepared by reacting the free acid or base forms of these compounds with a stoichiometric amount of the appropriate base or acid in water or in an organic solvent, or in a mixture of the two; generally, nonaqueous media like ether, ethyl acetate, ethanol, isopropanol, or acetonitrile are preferred. Lists of suitable salts are found in Remington's Pharmaceutical Sciences, 17th ed., Mack Publishing Company, Easton, Pa., 1985, p. 1418 and Journal of Pharmaceutical Science, 66, 2 (1977), each of which is incorporated herein by reference in its entirety.

**[000524] Polypeptide:** As used herein, "polypeptide" means a polymer of amino acid residues linked together by peptide bonds. The term, as used herein, refers to proteins, polypeptides, and
peptides of any size, structure, or function. Typically, however, a polypeptide will be at least 50 amino acids long. In some instances the polypeptide encoded is smaller than about 50 amino acids and the polypeptide is termed a peptide. If the polypeptide is a peptide, it will be at least about 5 amino acid residues long. Thus, polypeptides include gene products, naturally occurring polypeptides, synthetic polypeptides, homologs, orthologs, paralogs, fragments and other equivalents, variants, and analogs of the foregoing. A polypeptide may be a single molecule or may be a multi-molecular complex such as a dimer, trimer or tetramer. The term polypeptide may also apply to amino acid polymers in which one or more amino acid residues are an artificial chemical analogue of a corresponding naturally occurring amino acid.

**Sample:** As used herein, the term "sample" refers to a subset of its tissues, cells or component parts (e.g., body fluids, including but not limited to blood, mucus, lymphatic fluid, synovial fluid, cerebrospinal fluid, saliva, amniotic fluid, amniotic cord blood, urine, vaginal fluid and semen). A sample further may include a homogenate, lysate or extract prepared from a whole organism or a subset of its tissues, cells or component parts, or a fraction or portion thereof, including but not limited to, for example, plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, blood cells, tumors, organs. A sample further refers to a medium, such as a nutrient broth or gel, which may contain cellular components, such as proteins or nucleic acid molecule.

**Similarity:** As used herein, the term "similarity" refers to the overall relatedness between polymeric molecules, e.g. between polynucleotide molecules (e.g. DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Calculation of percent similarity of polymeric molecules to one another can be performed in the same manner as a calculation of percent identity, except that calculation of percent similarity takes into account conservative substitutions as is understood in the art.

**Stable:** As used herein "stable" refers to a compound that is sufficiently robust to survive isolation to a useful degree of purity from a reaction mixture, and preferably capable of formulation into an efficacious therapeutic agent.

**Subject:** As used herein, the term "subject" or "patient" refers to any organism to which a composition in accordance with the invention may be administered, e.g., for experimental, diagnostic, prophylactic, and/or therapeutic purposes. Typical subjects include animals (e.g., mammals such as mice, rats, rabbits, non-human primates, and humans) and/or plants.
Substantially: As used herein, the term "substantially" refers to the qualitative condition of exhibiting total or near-total extent or degree of a characteristic or property of interest. One of ordinary skill in the biological arts will understand that biological and chemical phenomena rarely, if ever, go to completion and/or proceed to completeness or achieve or avoid an absolute result. The term "substantially" is therefore used herein to capture the potential lack of completeness inherent in many biological and chemical phenomena.

Suffering from: An individual who is "suffering from" a disease, disorder, and/or condition has been diagnosed with or displays one or more symptoms of a disease, disorder, and/or condition.

Susceptible to: An individual who is "susceptible to" a disease, disorder, and/or condition has not been diagnosed with and/or may not exhibit symptoms of the disease, disorder, and/or condition but harbors a propensity to develop a disease or its symptoms. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition (for example, cancer) may be characterized by one or more of the following: (1) a genetic mutation associated with development of the disease, disorder, and/or condition; (2) a genetic polymorphism associated with development of the disease, disorder, and/or condition; (3) increased and/or decreased expression and/or activity of a protein and/or nucleic acid associated with the disease, disorder, and/or condition; (4) habits and/or lifestyles associated with development of the disease, disorder, and/or condition; (5) a family history of the disease, disorder, and/or condition; and (6) exposure to and/or infection with a microbe associated with development of the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will develop the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will not develop the disease, disorder, and/or condition.

Synthetic: The term "synthetic" means produced, prepared, and/or manufactured by the hand of man. Synthesis of polynucleotides or polypeptides or other molecules of the present invention may be chemical or enzymatic.

Targeted Cells: As used herein, "targeted cells" refers to any one or more cells of interest. The cells may be found in vitro, in vivo, in situ or in the tissue or organ of an organism. The organism may be an animal, preferably a mammal, more preferably a human and most preferably a patient.
Therapeutic Agent: The term "therapeutic agent" refers to any agent that, when administered to a subject, has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect.

Therapeutically effective amount: As used herein, the term "therapeutically effective amount" means an amount of an agent to be delivered (e.g., nucleic acid, drug, therapeutic agent, diagnostic agent, prophylactic agent, etc.) that is sufficient, when administered to a subject suffering from or susceptible to a disease, disorder, and/or condition, to treat, improve symptoms of, diagnose, prevent, and/or delay the onset of the disease, disorder, and/or condition.

Transcription factor: As used herein, the term "transcription factor" refers to a DNA-binding protein that regulates transcription of DNA into RNA, for example, by activation or repression of transcription. Some transcription factors effect regulation of transcription alone, while others act in concert with other proteins. Some transcription factor can both activate and repress transcription under certain conditions. In general, transcription factors bind a specific target sequence or sequences highly similar to a specific consensus sequence in a regulatory region of a target gene. Transcription factors may regulate transcription of a target gene alone or in a complex with other molecules.

Treating: As used herein, the term "treating" refers to partially or completely alleviating, ameliorating, improving, relieving, delaying onset of, inhibiting progression of, reducing severity of, and/or reducing incidence of one or more symptoms or features of a particular disease, disorder, and/or condition. For example, "treating" cancer may refer to inhibiting survival, growth, and/or spread of a tumor. Treatment may be administered to a subject who does not exhibit signs of a disease, disorder, and/or condition and/or to a subject who exhibits only early signs of a disease, disorder, and/or condition for the purpose of decreasing the risk of developing pathology associated with the disease, disorder, and/or condition.

Unmodified: As used herein, "unmodified" refers to any substance, compound or molecule prior to being changed in any way. Unmodified may, but does not always, refer to the wild type or native form of a biomolecule. Molecules may undergo a series of modifications whereby each modified molecule may serve as the "unmodified" starting molecule for a subsequent modification.
**Equivalents and Scope**

[000539] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments in accordance with the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the appended claims.

[000540] In the claims, articles such as "a," "an," and "the" may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.

[000541] It is also noted that the term "comprising" is intended to be open and permits the inclusion of additional elements or steps.

[000542] Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

[000543] In addition, it is to be understood that any particular embodiment of the present invention that falls within the prior art may be explicitly excluded from any one or more of the claims. Since such embodiments are deemed to be known to one of ordinary skill in the art, they may be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the compositions of the invention (e.g., any nucleic acid or protein encoded thereby; any method of production; any method of use; etc.) can be excluded from any one or more claims, for any reason, whether or not related to the existence of prior art.

[000544] All cited sources, for example, references, publications, databases, database entries, and art cited herein, are incorporated into this application by reference, even if not expressly stated in the citation. In case of conflicting statements of a cited source and the instant application, the statement in the instant application shall control.
EXAMPLES

Example 1. Modified mRNA Production

[000545] Modified mRNAs according to the invention are made using standard laboratory methods and materials.

[000546] The open reading frame with various upstream or downstream additions (β-globin, tags, etc.) is ordered from DNA2.0 (Menlo Park, CA) and typically contains a multiple cloning site with Xbal recognition. Upon receipt of the construct, it is reconstituted and transformed into chemically competent E. coli. For the present invention, NEB DH5-alpha Competent E. coli are used. A typical clone map is shown in Figure 1. Transformations are performed according to NEB instructions using 100 ng of plasmid. The protocol is as follows:

1. Thaw a tube of NEB 5-alpha Competent E. coli cells on ice for 10 minutes.
2. Add 1-5 µl containing 1 pg-100 ng of plasmid DNA to the cell mixture. Carefully flick the tube 4-5 times to mix cells and DNA. Do not vortex.
3. Place the mixture on ice for 30 minutes. Do not mix.
5. Place on ice for 5 minutes. Do not mix.
6. Pipette 950 µl of room temperature SOC into the mixture.
7. Place at 37°C for 60 minutes. Shake vigorously (250 rpm) or rotate.
8. Warm selection plates to 37°C.
9. Mix the cells thoroughly by flicking the tube and inverting.

[000547] Spread 50-100 µl of each dilution onto a selection plate and incubate overnight at 37°C. Alternatively, incubate at 30°C for 24-36 hours or 25°C for 48 hours.

[000548] A single colony is then used to inoculate 5 ml of LB growth media using the appropriate antibiotic and then allowed to grow (250 RPM, 37°C) for 5 hours. This is then used to inoculate a 200 ml culture medium and allowed to grow overnight under the same conditions.

[000549] To isolate the plasmid (up to 850 µg), a maxi prep is performed using the Invitrogen PureLink™ HiPure Maxiprep Kit (Carlsbad, CA), following the manufacturer's instructions.

[000550] In order to generate cDNA for In Vitro Transcription (IVT), the plasmid is first linearized using a restriction enzyme such as Xbal. A typical restriction digest with Xbal will comprise the following: Plasmid 1.0 µg; 10x Buffer 1.0 µl; Xbal 1.5 µl; dH₂O up to 10 µl; incubated at 37°C for 1 hr. If performing at lab scale (< 5µg), the reaction is cleaned up using Invitrogen's PureLink™ PCR
Micro Kit (Carlsbad, CA) per manufacturer’s instructions. Larger scale purifications may need to be done with a product that has a larger load capacity such as Invitrogen’s standard PureLink PCR Kit (Carlsbad, CA). Following the cleanup, the linearized vector is quantified using the NanoDrop and analyzed to confirm linearization using agarose gel electrophoresis.

**Example 2: PCR for cDNA Production**

The reverse primer of the instant invention incorporates a poly-Ti₂₀ for a poly-Ai₂₀ in the mRNA. Other reverse primers with longer or shorter poly(T) tracts can be used to adjust the length of the poly(A) tail in the mRNA.

**Example 3. In vitro Transcription**

The *in vitro* transcription reaction generates mRNA containing modified nucleotides or modified RNA. The input nucleotide triphosphate (NTP) mix is made in-house using natural and unnatural NTPs.

A typical *in vitro* transcription reaction includes the following:

1. Template cDNA 1.0 µg
2. 10x transcription buffer (400 mM Tris-HCl pH 8.0, 190 mM MgCl₂, 50 mM DTT, 10 mM Spermidine) 2.0 µl
3. Custom NTPs (25mM each) 7.2 µl
4. RNase Inhibitor 20 U
5. T7 RNA polymerase 3000 U
6. d40  Up to 20.0 µl. and
7. Incubation at 37° C for 3 hr-5 hrs.

[000556] The crude IVT mix may be stored at 4° C overnight for cleanup the next day. 1 U of RNase-free DNase is then used to digest the original template. After 15 minutes of incubation at 37° C, the mRNA is purified using Ambion's MEGAgear™ Kit (Austin, TX) following the manufacturer's instructions. This kit can purify up to 500 µg of RNA. Following the cleanup, the RNA is quantified using the NanoDrop and analyzed by agarose gel electrophoresis to confirm the RNA is the proper size and that no degradation of the RNA has occurred.

Example 4. Enzymatic Capping of mRNA

[000557] Capping of the mRNA is performed as follows where the mixture includes: IVT RNA 60 µg-180µg and dH2O up to 72 µl. The mixture is incubated at 65° C for 5 minutes to denature RNA, then transfer immediately to ice.

[000558] The protocol then involves the mixing of lOx Capping Buffer (0.5 M Tris-HCl (pH 8.0), 60 mM KCl, 12.5 mM MgCl2 (10.0 µl); 20 mM GTP (5.0 µl); 20 mM S-Adenosyl Methionine (2.5 µl); RNase Inhibitor (100 U); 2'-0-Methyltransferase (400U); Vaccinia capping enzyme (Guanylyl transferase) (40 U); dH2O (Up to 28 µl); and incubation at 37° C for 30 minutes for 60 µg RNA or up to 2 hours for 180 µg of RNA.

[000559] The mRNA is then purified using Ambion's MEGAgear™ Kit (Austin, TX) following the manufacturer's instructions. Following the cleanup, the RNA is quantified using the NanoDrop (ThermoFisher, Waltham, MA) and analyzed by agarose gel electrophoresis to confirm the RNA is the proper size and that no degradation of the RNA has occurred. The RNA product may also be sequenced by running a reverse-transcription-PCR to generate the cDNA for sequencing.

Example 5. PolyA Tailing Reaction

[000560] Without a poly-T in the cDNA, a poly-A tailing reaction must be performed before cleaning the final product. This is done by mixing Capped IVT RNA(100 µl); RNase Inhibitor (20 U); lOx Tailing Buffer (0.5 M Tris-HCl (pH 8.0), 2.5 M NaCl, 100 mM MgCl2(12.0 µl); 20 mM ATP (6.0 µl); Poly-A Polymerase (20 U); dH2O up to 123.5 µl and incubation at 37° C for 30 min. If the poly-A tail is already in the transcript, then the tailing reaction may be skipped and proceed directly to cleanup with Ambion's MEGAgear™ kit (up to 500 µg). Poly-A Polymerase is preferably a recombinant enzyme expressed in yeast.
Example 6. Enzymatic vs. Chemical Caps

Exemplary capping structures.

5'-capping of modified RNA may be completed concomitantly during the in vitro-transcription reaction using the following chemical RNA cap analogs to generate the 5'-guanosine cap structure according to manufacturer protocols: 3'-0-Me-m7G(5')ppp(5')G; G(5')ppp(5')A; G(5')ppp(5')G; m7G(5')ppp(5')A; m7G(5')ppp(5')G (New England BioLabs, Ipswich, MA). 5'-capping of modified RNA may be completed post-transcriptionally using a Vaccinia Virus Capping Enzyme to generate the "Cap 0" structure: m7G(5')ppp(5')G (New England BioLabs, Ipswich, MA). Cap 1 structure may be generated using both Vaccinia Virus Capping Enzyme and a 2'-0 methyl-transferase to generate: m7G(5')ppp(5')G-2'-0-methyl. Cap 2 structure may be generated from the Cap 1 structure followed by the 2'-0-methylation of the 5'-antepenultimate nucleotide using a 2'-0 methyl-transferase. Cap 3 structure may be generated from the Cap 2 structure followed by the 2'-O-methylation of the 5'-preantepenultimate nucleotide using a 2'-0 methyl-transferase. Enzymes are preferably derived from a recombinant source.

When transfected into mammalian cells, the modified mRNAs may have a stability of between 12-18 hours or more than 18 hours, e.g., 24, 36, 48, 60, 72 or greater than 72 hours.

Example 7. Chemical Cap vs. Enzymatically-Derived Cap Protein Expression Assay

Synthetic mRNAs encoding human G-CSF containing the ARCA cap analog or the Capl structure can be transfected into human primary keratinocytes at equal concentrations. 6, 12, 24 and 36 hours post-transfection the amount of G-CSF secreted into the culture medium can be assayed by ELISA. Synthetic mRNAs that secrete higher levels of G-CSF into the medium would correspond to a synthetic mRNA with a higher translationally-competent Cap structure.

Example 8. Chemical Cap vs. Enzymatically-Derived Cap Purity Analysis

Synthetic mRNAs encoding human G-CSF containing the ARCA cap analog or the Capl structure crude synthesis products can be compared for purity using denaturing Agarose-Urea gel electrophoresis or HPLC analysis. Synthetic mRNAs with a single, consolidated band by electrophoresis correspond to the higher purity product compared to a synthetic mRNA with multiple bands or streaking bands. Synthetic mRNAs with a single HPLC peak would also correspond to a higher purity product. The capping reaction with a higher efficiency would provide a more pure mRNA population.
Example 9. Chemical Cap vs. Enzymatically-Derived Cap Cytokine Analysis

[000565] Synthetic mRNAs encoding human G-CSF containing the ARCA cap analog or the Capl structure can be transfected into human primary keratinocytes at multiple concentrations. 6, 12, 24 and 36 hours post-transfection the amount of pro-inflammatory cytokines such as TNF-alpha and IFN-beta secreted into the culture medium can be assayed by ELISA. Synthetic mRNAs that secrete higher levels of pro-inflammatory cytokines into the medium would correspond to a synthetic mRNA containing an immune-activating cap structure.

Example 10. Chemical Cap vs. Enzymatically-Derived Cap Capping Reaction Efficiency

[000566] Synthetic mRNAs encoding human G-CSF containing the ARCA cap analog or the Capl structure can be analyzed for capping reaction efficiency by LC-MS after capped mRNA nuclease treatment. Nuclease treatment of capped mRNAs would yield a mixture of free nucleotides and the capped 5'-5'-triphosphate cap structure detectable by LC-MS. The amount of capped product on the LC-MS spectra can be expressed as a percent of total mRNA from the reaction and would correspond to capping reaction efficiency. The cap structure with a higher capping reaction efficiency would have a higher amount of capped product by LC-MS.

Example 11. Agarose Gel Electrophoresis of Modified RNA or RT PCR Products

[000567] Individual modRNAs (200-400 ng in a 20 µl volume) or reverse transcribed PCR products (200-400 ng) are loaded into a well on a non-denaturing 1.2% Agarose E-Gel (Invitrogen, Carlsbad, CA) and run for 12-15 minutes according to the manufacturer protocol.

Example 12. Nanodrop Modified RNA Quantification and UV Spectral Data:

[000568] Modified RNAs in TE buffer (1 µl) are used for Nanodrop UV absorbance readings to quantitate the yield of each modified RNA from an in vitro transcription reaction.

Example 13. Formulation of Modified mRNA Using Lipidoids

[000569] Modified mRNAs (mmRNAs) were made using standard laboratory methods and materials for in vitro transcription with the exception that the nucleotide mix contained modified nucleotides. The open reading frame (ORF) of the gene of interest is flanked by a 5' untranslated region (UTR) containing a strong Kozak translational initiation signal and an alpha-globin 3' UTR terminating with an oligo(dT) sequence for templated addition of a polyA tail for mmRNAs not incorporating Adenosine analogs. Adenosine-containing mmRNAs were synthesized without an oligo (dT) sequence to allow for post-transcription poly (A) polymerase poly-(A) tailing. In some
cases, the mmRNAs were modified by incorporating chemically modified nucleotides from the list indicated in Table 2 during the in vitro transcription with 100% replacement of the corresponding natural nucleotide or partial replacement of the corresponding natural nucleotide at the indicated percentage.

**[000570]** Modified mRNA are formulated for in vitro experiments by mixing the mmRNA with the lipidoid at a set ratio prior to addition to cells. In vivo formulation requires the addition of extra ingredients to facilitate circulation throughout the body. To test the ability of these lipidoids to form particles suitable for in vivo work, a standard formulation process used for siRNA-lipidoid formulations was used as a starting point. Initial mmRNA-lipidoid formulations consist of particles composed of 42% lipidoid, 48% cholesterol and 10% PEG, with further optimization of ratios possible. After formation of the particle, mmRNA was added and allowed to integrate with the complex. The encapsulation efficiency was determined using a standard dye exclusion assays.


**[000572]** The ratio of mmRNA to lipidoid used to test for in vitro transfection is tested empirically at different lipidoid:mmRNA ratios. Previous work using siRNA and lipidoids have utilized 2.5:1, 5:1, 10:1, and 15:1 lipidoid:siRNA wtwt ratios. Given the longer length of mmRNA relative to siRNA, a lower wtwt ratio of lipidoid to mmRNA may be effective. In addition, for comparison mmRNA were also formulated using RNAiMax (Invitrogen) or TRANSIT-mRNA (Minis Bio) cationic lipid delivery vehicles. The ability of lipidoid-formulated Luciferase, GFP, G-CSF, and EPO mmRNA to express the desired protein product can be confirmed by luminescence for
luciferase expression, flow cytometry for GFP expression, and by ELISA for G-CSF and Erythropoietin (EPO) secretion.

**Example 15. In vivo Expression of Modified RNA-Encoded Proteins Following Intravenous Injection Using Lipidoid Formulations**


[000574] Lipidoid formulations containing mmRNA can be injected intravenously into animals. The expression of the mmRNA-encoded proteins can be assessed in blood and other organs samples such as the liver and spleen collected from the animal. Conducting single dose intravenous studies will also allow an assessment of the magnitude, dose responsiveness, and longevity of expression of the desired product. In a study, lipidoid based formulations 98N12-5, C12-200, MD1 and other lipidoid-based formulations, may be used to deliver luciferase, green fluorescent protein (GFP), human G-CSF, or human Erythropoietin (EPO) mmRNA into the animal. After formulation of mmRNA with the lipidoid formulations as described previously, animals are divided into groups to receive either a saline formulation, or a lipidoid-formulation containing one of four different mmRNA selected from luciferase, GFP, human G-CSF and human EPO. Prior to injection into the animal, mmRNA-containing lipidoid formulations are diluted in PBS. Animals are then administered a single dose of formulated mmRNA ranging from a dose of 10 mg/kg to doses as low as 1 mg/kg, with a preferred range to be 10 mg/kg to 100 ng/kg, depending on the amount of mmRNA injected per animal body weight. If the animal is a mouse, the volume of an intravenous injection of the lipidoid formulation is a maximum of 0.2 ml for a 20 gram mouse. At various points in time following the administration of the mmRNA-lipidoid, serum, tissues, and tissue lysates can be obtained and the level of the mmRNA-encoded product determined. The ability of lipidoid-formulated Luciferase, GFP, G-CSF, and EPO mmRNA to express the desired protein product can
be confirmed by luminescence for luciferase expression, flow cytometry for GFP expression, and by ELISA for G-CSF and Erythropoietin (EPO) secretion.

Additional studies for a multi-dose regimen can also be performed to determine the maximal expression using mmRNA, to evaluate the saturability of the mmRNA-driven expression (achieved by giving a control and active mmRNA formulation in parallel or in sequence), and to determine the feasibility of repeat drug administration (by giving mmRNA in doses separated by weeks or months and then determining whether expression level is affected by factors such as immunogenicity). In addition to detection of the expressed protein product, an assessment of the physiological function of proteins such as G-CSF and EPO can also be determined through analyzing samples from the animal tested and detecting increases in granulocyte and red blood cell counts, respectively.

**Example 16. In vivo Expression of Modified RNA-Encoded Proteins Following Intramuscular and/or Subcutaneous Injection Using Lipidoid Formulations**

The use of lipidoid formulations to deliver oligonucleotides, including siRNA, via an intramuscular route or a subcutaneous route of injection needs to be evaluated as it has not been previously reported. The intramuscular and/or subcutaneous injection of mmRNA-containing lipidoid formulations will be evaluated to determine if they are capable to produce both localized and systemic expression of the desired proteins.

Lipidoid formulations containing mmRNA can be injected intramuscularly and/or subcutaneously into animals. The expression of mmRNA-encoded proteins can be assessed both within the muscle or subcutaneous tissue and systemically in blood and other organs such as the liver and spleen. The ability of 98N12-5, C12-200, and MD1-based lipidoid formulations, and possibly other lipidoid-based formulations, to deliver either luciferase, green fluorescent protein (GFP), human G-CSF, or human Erythropoietin (EPO) mmRNA will be evaluated. Conducting single dose studies will also allow an assessment of the magnitude, dose responsiveness, and longevity of expression of the desired product. After the formulation of mmRNA with the lipidoid formulations, as described previously, animals will be divided into groups receiving either a saline formulation, or a lipidoid-formulation containing one of four different mmRNA selected from, luciferase, GFP, human G-CSF, human EPO. Prior to injection, mmRNA-containing lipidoid formulations are diluted in PBS and animals administered a single intramuscular dose of formulated mmRNA ranging from 50 mg/kg to doses as low as 1 ng/kg with a preferred range to be 10 mg/kg to 100 ng/kg. If
the animal tested is a mouse the maximum dose can be roughly 1 mg mmRNA or as low as 0.02 ng
mmRNA if administered once into the hind limb. Likewise for subcutaneous administration,
mmRNA-containing lipidoid formulations are diluted in PBS before the animals are administered a
single subcutaneous dose of formulated mmRNA ranging from 400 mg/kg- to doses as low as 1
ng/kg. A preferred dosage range may be 80 mg/kg to 100 ng/kg. If the animal tested is a mouse, the
maximum dose administered can be roughly 8 mg mmRNA or as low as 0.02ng mmRNA if the dose
is administered once subcutaneously.

[000578] It is preferred that the volume of a single intramuscular injection is maximally 0.025 ml
and of a single subcutaneous injection is maximally 0.2 ml for a 20 gram mouse. The dose of the
mmRNA administered to the animal is calculated depending on the body weight of the animal. At
various points in time points following the administration of the mmRNA-lipidoid, serum, tissues,
and tissue lysates can be obtained and the level of the mmRNA-encoded product determined. The
ability of lipidoid-formulated Luciferase, GFP, G-CSF, and EPO mmRNA to express the desired
protein product can be confirmed by luminescence for luciferase expression, flow cytometry for GFP
expression, and by ELISA for G-CSF and Erythropoietin (EPO) secretion.

[000579] Additional studies for a multi-dose regimen can also be performed to determine the
maximal expression using mmRNA, to evaluate the saturability of the mmRNA-driven expression
(achieved by giving a control and active mmRNA formulation in parallel or in sequence), and to
determine the feasibility of repeat drug administration (by giving mmRNA in doses separated by
weeks or months and then determining whether expression level is affected by factors such as
immunogenicity). Studies utilizing multiple subcutaneous or intramuscular injection sites at one
time point, can also be utilized to further increase mmRNA drug exposure and improve protein
production. In addition to detection of the expressed protein product, an assessment of the
physiological function of proteins such as G-CSF and EPO can also be determined through
analyzing samples from the animal tested and detecting increases in granulocyte and red blood cell
counts, respectively.

[000580] It is to be understood that the words which have been used are words of description rather
than limitation, and that changes may be made within the purview of the appended claims without
departing from the true scope and spirit of the invention in its broader aspects.

[000581] While the present invention has been described at some length and with some
particularity with respect to the several described embodiments, it is not intended that it should be
limited to any such particulars or embodiments or any particular embodiment, but it is to be
construed with references to the appended claims so as to provide the broadest possible interpretation of such claims in view of the prior art and, therefore, to effectively encompass the intended scope of the invention.

[000582] All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, section headings, the materials, methods, and examples are illustrative only and not intended to be limiting.
Claims

We claim:

1. A composition comprising a modified nucleic acid encoding a cell-penetrating polypeptide, wherein the modified nucleic acid provokes a reduced innate immune response as compared to an unmodified nucleic acid encoding the same cell-penetrating polypeptide when introduced into a first cell.

2. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a complex comprising a charged protein operably linked to a therapeutic protein.

3. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a fusion protein comprising a charged protein covalently linked to a therapeutic protein.

4. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a therapeutic protein comprising a surface charge.

5. The composition of claim 1, wherein the cell-penetrating polypeptide further comprises a detectable label.

6. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a supercharged polypeptide.

7. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a signal sequence for secretion of the cell-penetrating polypeptide.

8. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a ratio of surface charge to surface amino acids is approximately 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9.

9. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a first domain comprising a supercharged polypeptide and a second domain comprising a protein-binding partner.

10. The composition of claim 9, wherein the cell-penetrating polypeptide is capable of penetrating a second cell.

11. The composition of claim 10, wherein the first cell and the second cell are cells from the same tissue.

12. The composition of claim 10, wherein the first cell and the second cell are cells from different tissues.

13. The composition of claim 10, wherein the first cell is located proximal to the second cell.
14. The composition of claim 10, wherein the first cell is located distal to the second cell.
15. The composition of claim 10, wherein the cell-penetrating polypeptide penetrating the second cell comprises an intracellular binding partner for the protein-binding partner.
16. The composition of claim 9, wherein the cell-penetrating polypeptide is capable of penetrating the first cell.
17. The composition of claim 1, wherein the cell-penetrating polypeptide comprises a protein-binding partner.
18. The composition of claim 17, wherein the protein-binding partner is selected from a supercharged antibody and a functional fragment thereof.
19. The composition of claim 17 where the modified nucleic acid is introduced to a cell where the cell-penetrating polypeptide is secreted from.
20. The composition of claim 19, wherein the cell-penetrating polypeptide is capable of penetrating the second cell comprising an intracellular binding partner for the supercharged antibody or functional fragment thereof.
21. The composition of claim 19, wherein the cell-penetrating polypeptide is capable of penetrating the first cell.
22. A formulation comprising the composition of claim 1 and a delivery agent.
23. The formulation of claim 22, wherein the delivery agent comprises a cationic material.
24. The formulation of claim 22, wherein the delivery agent comprises a cell-penetrating polypeptide.
25. A method of delivering a protein-binding partner into a target cell comprising an intracellular binding partner for the protein-binding partner, comprising the step of introducing into a first cell a composition comprising a modified nucleic acid encoding a cell-penetrating polypeptide comprising the protein-binding partner under conditions wherein the cell-penetrating polypeptide is produced in and secreted from the first cell, and subsequently penetrates into the target cell.
26. The method of claim 25, wherein the modified nucleic acid provokes a reduced innate immune response as compared to an unmodified nucleic acid encoding the same cell-penetrating polypeptide when introduced into the first cell.
27. The method of claim 25, wherein the first cell and the target cell are the same cell.
28. The method of claim 25, wherein the first cell and the target cell are co-located in a tissue of a subject.
29. The method of claim 25, wherein the first cell is located proximal to the target cell.
30. The method of claim 25, wherein the first cell is located distal to the target cell.
31. The method of claim 25, wherein the modified nucleic acid is retained in an extracellular depot adjacent to the first cell prior to introduction in the first cell.
# INTERNATIONAL SEARCH REPORT

<table>
<thead>
<tr>
<th>Classification of Subject Matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPC(8) - A61K 48/00; C07H 21/00; C12N 15/00 (2013.01)</td>
</tr>
<tr>
<td>USPC - 514/44R; 536/23.1; 23.4; 435/455</td>
</tr>
</tbody>
</table>

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

## B. Fields Searched

<table>
<thead>
<tr>
<th>Minimum documentation searched (classification system followed by classification symbols)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPC(8) - A61K 48/00; C07H 21/00; C12N 15/00 (2013.01)</td>
</tr>
<tr>
<td>USPC - 514/44R; 536/23.1; 23.4; 435/455</td>
</tr>
</tbody>
</table>

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

| USPC - 536/23.53; 530/350, 387.1 |

(Text Search)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

- PubMed: (PGP); USPT; USOC; EPAB; IPAB; Google Scholar and PubBase Search Terms: penetrateS, enter, enterS, secrets, delivS, entS, sequence, encodiS, cell, modiS, neighbors, adjacenS, chargS, superchargS, oligoribonucleotS, ribonucleotS, nucleotS, polyribonucleotS, dna, nucleic, nucleotS, peptidS, polypeptidS, oligopeptidS.

## C. Documents Considered to be Relevant

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to Claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>US 2010/0233141 A1 (POLACH et al.) 16 September 2010 (16.09.2010) abstract; Fig. 1; para [0001]-[0026], [0027]-[0030], [0031]-[0033], [0034]-[0036], [0037]-[0039], [0040]-[0042]</td>
<td>5 and 27-31</td>
</tr>
<tr>
<td>Y</td>
<td>US 201 0120140 A1 (LIU et al.) 12 May 2011 (12.05.2011) abstract; para [0001]-[0010], [0011]-[0014], [0015]-[0018], [0019]-[0022], [0023]-[0026], [0027]-[0030]</td>
<td>1-24 and 26</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>

- **A** Special categories of cited documents:
  - "A" document defining the general state of the art which is not considered to be of particular relevance
  - "E" earlier application or patent but published on or after the international filing date
  - "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  - "O" document referring to an oral disclosure, use, exhibition or other means
  - "P" document published prior to the international filing date but later than the priority date claimed

Recent documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

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

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

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

- **&** document member of the same patent family

- **&** document member of the same patent family

---

Date of the actual completion of the international search: 19 March 2013 (19.03.2013)

Date of mailing of the international search report: 05 APR 2013

Name and mailing address of the ISA/US:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents and Trademarks, P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 71-272-3201

Authorized officer: Lee W. Young
PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)