# MODIFIED NUCLEOSIDES, NUCLEOTIDES, AND NUCLEIC ACIDS, AND USES THEREOF

## REFERENCE TO THE SEQUENCE LISTING

**[0001]** The present application is being filed along with a Sequence Listing in electronic format. The Sequence Listing file, entitled M009SQLST.txt, was created on October 3, 2012 and is 9,859 bytes in size. The information in electronic format of the Sequence Listing is incorporated herein by reference in its entirety.

## CROSS REFERENCE TO RELATED APPLICATIONS

**[0002]** This application claims priority to U.S. Provisional Patent Application No. 61/542,533, filed October 3, 2011, entitled Modified Nucleosides, Nucleotides, and Nucleic Acids, and Uses Thereof, the contents of which are incorporated by reference in its entirety.

## TECHNICAL FIELD

**[0003]** The present disclosure provides compositions and methods using modified nucleic acids to modulate cellular function. The modified nucleic acids of the invention may encode peptides, polypeptides or multiple proteins. The encoded molecules may be used as therapeutics and/or diagnostics.

## BACKGROUND OF THE INVENTION

**[0004]** Naturally occurring RNAs are synthesized from four basic ribonucleotides: ATP, CTP, UTP and GTP, but may contain post-transcriptionally modified nucleotides. Further, approximately one hundred different nucleoside modifications have been identified in RNA (Rozenski, J, Crain, P, and McCloskey, J. (1999). The RNA Modification Database: 1999 update. Nucl Acids Res 27: 196-197). The role of nucleoside modifications on the immune-stimulatory potential and on the translation efficiency of RNA, however, is unclear.

**[0005]** There are multiple problems with prior methodologies of effecting protein expression. For example, heterologous 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.

**[0006]** There is a need in the art for biological modalities to address the modulation of intracellular translation of nucleic acids.

**SUMMARY OF THE INVENTION**

**[0007]** The present disclosure provides, *inter alia*, modified nucleosides, modified nucleotides, and modified nucleic acids which can exhibit a reduced innate immune response when introduced into a population of cells, both *in vivo* and *ex vivo*.

**[0008]** The present invention provides polynucleotides which may be isolated or purified. These polynucleotides may encode one or more polypeptides of interest and comprise a sequence of n number of linked nucleosides or nucleotides comprising at least one modified nucleoside or nucleotide as compared to the chemical structure of an A, G, U or C nucleoside or nucleotide. The polynucleotides may also contain a 5' UTR comprising at least one Kozak sequence, a 3' UTR, and at least one 5′ cap structure. The isolated polynucleotides may further contain a poly-A tail and may be purified.

**[0009]** The isolated polynucleotides of the invention also comprise at least one 5′ cap structure selected from the group consisting of Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2′fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

**[0010]** Modifications of the polynucleotides of the invention may be on the nucleoside base and/or sugar portion of the nucleosides which comprise the polynucleotide.

**[0011]** In some embodiments, the modification is on the nucleobase and is selected from the group consisting of pseudouridine or N1-methylpseudouridine.

**[0012]** In some embodiments, the modified nucleoside is not pseudouridine (ψ) or 5-methyl-cytidine (m5C).

**[0013]** In some embodiments, multiple modifications are included in the modified nucleic acid or in one or more individual nucleoside or nucleotide. For example, modifications to a nucleoside may include one or more modifications to the nucleobase and the sugar.

**[0014]** In some embodiments are provided novel building blocks, e.g., nucleosides and nucleotides for the preparation of modified polynucleotides and their method of synthesis and manufacture.

**[0015]** The present invention also provides for pharmaceutical compositions comprising the modified polynucleotides described herein. These may also further include one or more pharmaceutically acceptable excipients selected from a solvent, aqueous solvent, non-aqueous solvent, dispersion media, diluent, dispersion, suspension aid, surface active agent, isotonic agent, thickening or emulsifying agent, preservative, lipid, lipidoids liposome, lipid nanoparticle, core-shell nanoparticles, polymer, lipoplexe peptide, protein, cell, hyaluronidase, and mixtures thereof.

**[0016]** Methods of using the polynucleotides and modified nucleic acids of the invention are also provided. In this instance, the poynucleotides may be formulated by any means known in the art or administered via any of several routes including injection by intradermal, subcutaneous or intramuscular means.

**[0017]** Administration of the modified nucleic acids of the inventin may be via two or more equal or unequal split doses. In some embodiments, the level of the polypeptide produced by the subject by administering split doses of the polynucleotide is greater than the levels produced by administering the same total daily dose of polynucleotide as a single administration.

**[0018]** Detection of the modified nucleic acids or the encoded polypeptides may be performed in the bodily fluid of the subject or patient where the bodily fluid is selected from the group consisting of peripheral blood, serum, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, broncheoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucosal secretion, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyl cavity fluid, and umbilical cord blood.

**[0019]** In some embodiments, administration is according to a dosing regimen which occurs over the course of hours, days, weeks, months, or years and may be achieved by using one or more devices selected from multi-needle injection systems, catheter or lumen systems, and ultrasound, electrical or radiation based systems.

**[0020]** Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Methods and materials are described herein for use in the present disclosure; other, suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

**[0021]** Other features and advantages of the present disclosure will be apparent from the following detailed description and figures, and from the claims.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**[0022]** The foregoing and other objects, features and advantages will be apparent from the following description of particular embodiments of the invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating the principles of various embodiments of the invention.

**[0023]** FIG. 1 provides the spectrum and graphs of the analytical results for N4-Me-CTP (NTP of compound 1). Figure 1A provides the nuclear magnetic resonance (NMR) spectrum in DMSO and Figure 1B provides the NMR spectrum in $D_2O$, Figure 1C provides the mass spectrometry (MS) results, and Figure 1D is the high performance liquid chromatography (HPLC) results for N4-methylcytidine (N4-Me-cytidine, compound 1).

**[0024]** FIG. 2 shows the HPLC results for N4-Me-CTP (NTP of compound 1).

**[0025]** FIG.3 provides the analytical results for 2'-OMe-N, N-di-Me-CTP (NTP of compound 2). Figure 3A provides the NMR spectrum. Figure 3B provides the MS results. Figure 3C provides HPLC results for 2'-O-methyl-$N^4$, $N^4$-dimethylcytidine (2'-OMe-N,N-di-Me-cytidine, compound 2).

**[0026]** FIG. 4 shows the HPLC results for 2'-OMe-N, N-di-Me-CTP (NTP of compound 2).

**[0027]** FIG. 5 provides the HPLC results for 5-methoxycarbonylmethoxy-UTP (NTP of compound 3).

**[0028]** FIG. 6 provides the analytical results of 3-methyl pseudouridine (compound 4). Figure 6A provides the NMR spectrum of 3-methyl pseudouridine (compound 4) and Figure 6B provides the HPLC results for 3-methyl pseudouridine (compound 4).

**[0029]** FIG. 7 provides the analytical results of 5-TBDMS-$OCH_2$-cytidine (compound 6). Figure 7A provide the NMR spectrum, Figure 7B provides the MS results, and Figure 7C provides the HPLC results for 5-TBDMS-$OCH_2$-cytidine (compound 6).

**[0030]** FIG. 8 provides the analytical results of 5-trifluoromethyl uridine (compound 8). Figure 8A provides the NMR spectrum, Figure 8B provides MS results, and Figure 8C provides HPLC results for 5-trifluoromethyl uridine (compound 8).

**[0031]** FIG. 9 provides the NMR spectrum results for of 5-(methoxycarbonyl) methyl uridine (compound 9).

**[0032]** FIG. 10 provides a graph showing the variability of protein (GCSF; line B) and cytokine (interferon-alpha (IFNa); line A and tumor necrosis factor-alpha (TNFa); line C) expression as function of percent modification.

**DETAILED DESCRIPTION**

**[0033]** The present disclosure provides, *inter alia*, modified nucleosides, modified nucleotides, and modified nucleic acids that exhibit improved therapeutic properties including, but not limited to, a reduced innate immune response when introduced into a population of cells.

**[0034]** As there remains a need in the art for therapeutic modalities to address the myriad of barriers surrounding the efficacious modulation of intracellular translation and processing of nucleic acids encoding polypeptides or fragments thereof, the inventors have shown that certain modified mRNA sequences have the potential as therapeutics with benefits beyond just evading, avoiding or diminishing the immune response.

**[0035]** The present invention addresses this need by providing nucleic acid based compounds or polynucleotides which encode a polypeptide of interest (e.g., modified mRNA) and which have structural and/or chemical features that avoid one or more of the problems in the art, for example, features which are useful for optimizing nucleic acid-based therapeutics while retaining structural and functional integrity, overcoming the threshold of expression, improving expression rates, half life and/or protein concentrations, optimizing protein localization, and avoiding deleterious bio-responses such as the immune response and/or degradation pathways.

**[0036]** Provided herein, in part, are polynucleotides encoding polypeptides of interest which have been chemically modified to improve one or more of the stability and/or clearance in tissues, receptor uptake and/or kinetics, cellular access by the compositions, engagement with translational machinery, mRNA half-life, translation efficiency, immune evasion, protein production capacity,

secretion efficiency (when applicable), accessibility to circulation, protein half-life and/or modulation of a cell's status, function and/or activity.

**[0037]** The modified nucleosides, nucleotides and nucleic acids of the invention, including the combination of modifications taught herein have superior properties making them more suitable as therapeutic modalities.

**[0038]** It has been determined that the "all or none" model in the art is sorely insufficient to describe the biological phenomena associated with the therapeutic utility of modified mRNA. The present inventors have determined that to improve protein production, one may consider the nature of the modification, or combination of modifications, the percent modification and survey more than one cytokine or metric to determine the efficacy and risk profile of a particular modified mRNA.

**[0039]** In one aspect of the invention, methods of determining the effectiveness of a modified mRNA as compared to unmodified involves the measure and analysis of one or more cytokines whose expression is triggered by the administration of the exogenous nucleic acid of the invention. These values are compared to administration of an umodified nucleic acid or to a standard metric such as cytokine response, PolyIC, R-848 or other standard known in the art.

**[0040]** One example of a standard metric developed herein is the measure of the ratio of the level or amount of encoded polypeptide (protein) produced in the cell, tissue or organism to the level or amount of one or more (or a panel) of cytokines whose expression is triggered in the cell, tissue or organism as a result of administration or contact with the modified nucleic acid. Such ratios are referred to herein as the Protein:Cytokine Ratio or "PC" Ratio. The higher the PC ratio, the more efficacioius the modified nucleic acid (polynucleotide encoding the protein measured). Preferred PC Ratios, by cytokine, of the present invention may be greater than 1, greater than 10, greater than 100, greater than 1000, greater than 10,000 or more. Modified nucleic acids having higher PC Ratios than a modified nucleic acid of a different or unmodified construct are preferred.

**[0041]** The PC ratio may be further qualified by the percent modification present in the polynucleotide. For example, normalized to a 100% modified nucleic acid, the protein production as a function of cytokine (or risk) or cytokine profile can be determined.

**[0042]** In one embodiment, the present invention provides a method for determining, across chemistries, cytokines or percent modification, the relative efficacy of any particular modified polynucleotide by comparing the PC Ratio of the modified nucleic acid (polynucleotide).

**[0043]** In another embodiment, the chemically modified mRNA are substantially non toxic and non mutagenic.

**[0044]** In one embodiment, the modified nucleosides, modified nucleotides, and modified nucleic acids can be chemically modified on the major groove face, thereby disrupting major groove binding partner interactions, which may cause innate immune responses. Further, these modified nucleosides, modified nucleotides, and modified nucleic acids can be used to deliver a payload, *e.g.*, detectable or therapeutic agent, to a biological target. For example, the nucleic acids can be covalently linked to a payload, *e.g.* a detectable or therapeutic agent, through a linker attached to the nucleobase or the sugar moiety. The compositions and methods described herein can be used, *in vivo* and *in vitro*, both extracellarly or intracellularly, as well as in assays such as cell free assays.

**[0045]** In some embodiments, the present disclosure provides compounds comprising a nucleotide that 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 partners.

**[0046]** In another aspect, the present disclosure provides nucleotides that contain chemical modifications, wherein the nucleotide has altered binding to major groove interacting partners.

**[0047]** In some embodiments, the chemical modifications are located on the major groove face of the nucleobase, and wherein the chemical modifications can include replacing or substituting an atom of a pyrimidine nucleobase with an amine, an SH, an alkyl (*e.g.*, methyl or ethyl), or a halo (*e.g.*, chloro or fluoro).

**[0048]** In another aspect, the present disclosure provides chemical modifications located on the sugar moiety of the nucleotide.

**[0049]** In another aspect, the present disclosure provides chemical modifications located on the phosphate backbone of the nucleic acid.

**[0050]** In some embodiments, the chemical modifications alter the electrochemistry on the major groove face of the nucleic acid.

**[0051]** In another aspect, the present disclosure provides nucleotides that contain chemical modifications, wherein the nucleotide reduces the cellular innate immune response, as compared to the cellular innate immune induced by a corresponding unmodified nucleic acid.

**[0052]** In another aspect, the present disclosure provides nucleic acid sequences comprising at least two nucleotides, the nucleic acid sequence comprising a nucleotide that disrupts binding of a

major groove interacting partner with the nucleic acid sequence, wherein the nucleotide has decreased binding affinity to the major groove binding partner.

**[0053]** In another aspect, the present disclosure provides compositions comprising a compound as described herein. In some embodiments, the composition is a reaction mixture. In some embodiments, the composition is a pharmaceutical composition. In some embodiments, the composition is a cell culture. In some embodiments, the composition further comprises an RNA polymerase and a cDNA template. In some embodiments, the composition further comprises a nucleotide selected from the group consisting of adenosine, cytosine, guanosine, and uracil.

**[0054]** In a further aspect, the present disclosure provides methods of making a pharmaceutical formulation comprising a physiologically active secreted protein, comprising transfecting a first population of human cells with the pharmaceutical nucleic acid made by the methods described herein, wherein the secreted protein is active upon a second population of human cells.

**[0055]** In some embodiments, the secreted protein is capable of interacting with a receptor on the surface of at least one cell present in the second population.

**[0056]** In some embodiments, the secreted protein is Granulocyte-Colony Stimulating Factor (G-CSF).

**[0057]** In some embodiments, the second population contains myeloblast cells that express the G-CSF receptor.

**[0058]** In certain embodiments, provided herein 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 toxitity. 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)).

**[0059]** In one embodiment, it is intended that the compounds of the present disclosure are stable. It is further appreciated that certain features of the present disclosure, which are, for clarity, described in the context of separate embodiments, can also be provided in combination in a single embodiment. Conversely, various features of the present disclosure which are, for brevity, described in the context of a single embodiment, can also be provided separately or in any suitable subcombination.

**Modified Nucleotides, Nucleosides and Polynucleotides of the invention**

**[0060]** Herein, in a nucleotide, nucleoside or polynucleotide (such as the nucleic acids of the invention, e.g., mRNA molecule), the terms "modification" or, as appropriate, "modified" refer to modification with respect to A, G, U or C ribonucleotides. Generally, herein, these terms are not intended to refer to the ribonucleotide modifications in naturally occurring 5′-terminal mRNA cap moieties. In a polypeptide, the term "modification" refers to a modification as compared to the canonical set of 20 amino acids, moiety)

**[0061]** The modifications may be various distinct modifications. In some embodiments, where the nucleic acid is an mRNA, the coding region, the flanking regions and/or the terminal regions may contain one, two, or more (optionally different) nucleoside or nucleotide modifications. In some embodiments, a modified polynucleotide introduced to a cell may exhibit reduced degradation in the cell, as compared to an unmodified polynucleotide.

**[0062]** The polynucleotides can include any useful modification, such as to the sugar, the nucleobase, or the internucleoside linkage (*e.g.* to a linking phosphate / to a phosphodiester linkage / to the phosphodiester backbone). For example, the major groove of a polynucleotide, or the major groove face of a nucleobase may comprise one or more modifications. One or more atoms of a pyrimidine nucleobase (e.g. on the major groove face) may be replaced or substituted with optionally substituted amino, optionally substituted thiol, optionally substituted alkyl (e.g., methyl or ethyl), or halo (e.g., chloro or fluoro). In certain embodiments, modifications (e.g., one or more modifications) are present in each of the sugar and the internucleoside linkage. Modifications according to the present invention may be modifications of ribonucleic acids (RNAs) to deoxyribonucleic acids (DNAs), e.g., the substitution of the 2'OH of the ribofuranysyl ring to 2'H, threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs) or hybrids thereof). Additional modifications are described herein.

**[0063]** As described herein, the polynucleotides of the invention do not substantially induce an innate immune response of a cell into which the polynucleotide (e.g., mRNA) is introduced. Features of an induced innate immune response include 1) increased expression of pro-inflammatory cytokines, 2) activation of intracellular PRRs (RIG-I, MDA5, etc, and/or 3) termination or reduction in protein translation.

**[0064]** In certain embodiments, it may desirable for a modified nucleic acid molecule introduced into the cell to be degraded intracellulary. For example, degradation of a modified nucleic acid

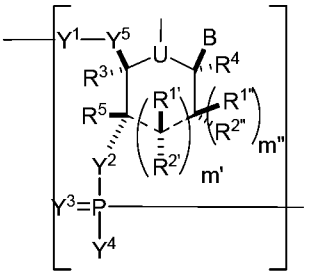
molecule may be preferable if precise timing of protein production is desired. Thus, in some embodiments, the invention provides a modified nucleic acid molecule containing a degradation domain, which is capable of being acted on in a directed manner within a cell. In another aspect, the present disclosure provides polynucleotides comprising a nucleoside or nucleotide that can disrupt the binding of a major groove interacting, e.g. binding, partner with the polynucleotide (e.g., where the modified nucleotide has decreased binding affinity to major groove interacting partner, as compared to an unmodified nucleotide).

**[0065]** The polynucleotides can optionally include other agents (e.g., RNAi-inducing agents, RNAi agents, siRNAs, shRNAs, miRNAs, antisense RNAs, ribozymes, catalytic DNA, tRNA, RNAs that induce triple helix formation, aptamers, vectors, etc.). In some embodiments, the polynucleotides may include one or more messenger RNAs (mRNAs) having one or more modified nucleoside or nucleotides (i.e., modified mRNA molecules). Details for these polynucleotides follow.

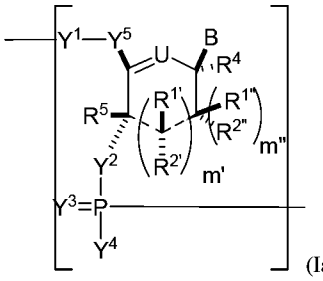
*Polynucleotides*

**[0066]** The polynucleotides of the invention includes a first region of linked nucleosides encoding a polypeptide of interest, a first flanking region located at the 5′ terminus of the first region, and a second flanking region located at the 3′ terminus of the first region.

**[0067]** In some embodiments, the polynucleotide (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (Ia) or Formula (Ia-1):



(Ia)



(Ia-1) or a pharmaceutically acceptable salt or stereoisomer thereof, wherein U is O, S, $N(R^U)_{nu}$, or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

**[0068]** - - - is a single bond or absent;

**[0069]** each of $R^{1'}$, $R^{2'}$, $R^{1''}$, $R^{2''}$, $R^1$, $R^2$, $R^3$, $R^4$, and $R^5$, if present, is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent; wherein the combination of $R^3$ with one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, $R^{2''}$, or $R^5$ (e.g., the combination of $R^{1'}$ and $R^3$, the combination of $R^{1''}$ and $R^3$, the combination of $R^{2'}$ and $R^3$, the combination of $R^{2''}$ and $R^3$, or the combination of $R^5$ and $R^3$) can join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl (e.g., a bicyclic, tricyclic, or tetracyclic heterocyclyl); wherein the combination of $R^5$ with one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, or $R^{2''}$ (e.g., the combination of $R^{1'}$ and $R^5$, the combination of $R^{1''}$ and $R^5$, the combination of $R^{2'}$ and $R^5$, or the combination of $R^{2''}$ and $R^5$) can join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl (e.g., a bicyclic, tricyclic, or tetracyclic heterocyclyl); and wherein the combination of $R^4$ and one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, $R^{2''}$, $R^3$, or $R^5$ can join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl (e.g., a bicyclic, tricyclic, or tetracyclic heterocyclyl);

**[0070]** each of m' and m" is, independently, an integer from 0 to 3 (e.g., from 0 to 2, from 0 to 1, from 1 to 3, or from 1 to 2);

**[0071]** each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, -$NR^{N1}$-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, or absent;

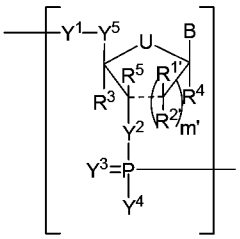
**[0072]** each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino;

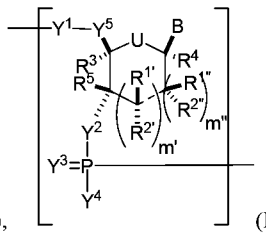
**[0073]** each $Y^5$ is, independently, O, S, Se, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

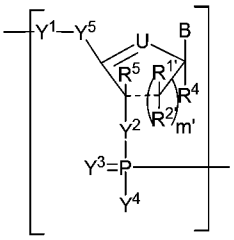
**[0074]** n is an integer from 1 to 100,000; and

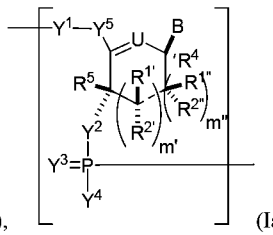
**[0075]** B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof), wherein the combination of B and $R^{1'}$, the combination of B and $R^{2'}$, the combination of B and $R^{1''}$, or the combination of B and $R^{2''}$ can, taken together with the carbons to which they are attached, optionally form a bicyclic group (e.g., a bicyclic heterocyclyl) or wherein the combination of B, $R^{1''}$, and $R^3$ or the combination of B, $R^{2''}$, and $R^3$ can optionally form a tricyclic or tetracyclic group (e.g., a tricyclic or tetracyclic heterocyclyl, such as in Formula (IIo)-(IIp) herein).

**[0076]** In some embodiments, the polynucleotide includes a modified ribose. In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (Ia-2)-(Ia-5) or a pharmaceutically acceptable salt or stereoisomer thereof.

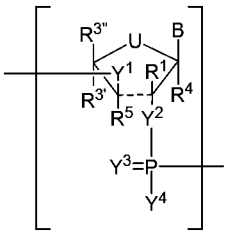
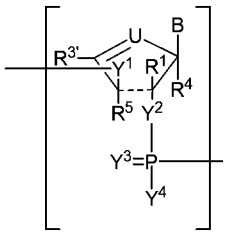

(Ia-2),


(Ia-3),


(Ia-4),


(Ia-5).

**[0077]** In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (Ib) or Formula (Ib-1):


(Ib),

(Ib-1)

**[0078]** or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[0079]** U is O, S, $N(R^U)_{nu}$, or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

**[0080]** - - - is a single bond or absent;

**[0081]** each of $R^1$, $R^{3'}$, $R^{3''}$, and $R^4$ is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent; and wherein the combination of $R^1$ and $R^{3'}$ or the combination of $R^1$ and $R^{3''}$ can be taken together to form optionally substituted alkylene or optionally substituted heteroalkylene (e.g., to produce a locked nucleic acid);

**[0082]** each $R^5$ is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, or absent;

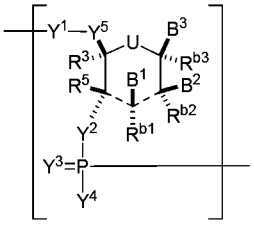
**[0083]** each of $Y^1$, $Y^2$, and $Y^3$ is, independently, O, S, Se, $NR^{N1}$-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl;

**[0084]** each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted alkoxyalkoxy, or optionally substituted amino;

**[0085]** n is an integer from 1 to 100,000; and

**[0086]** B is a nucleobase.

**[0087]** In some embodiments, the polynucleotide (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (Ic):



(Ic), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[0088]** U is O, S, $N(R^U)_{nu}$, or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

**[0089]** - - - is a single bond or absent;

**[0090]** each of $B^1$, $B^2$, and $B^3$ is, independently, a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof, as described herein), H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl, wherein one and only one of $B^1$, $B^2$, and $B^3$ is a nucleobase;

**[0091]** each of $R^{b1}$, $R^{b2}$, $R^{b3}$, $R^3$, and $R^5$ is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl;

**[0092]** each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, $-NR^{N1}-$, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl;

**[0093]** each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy,

optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino;

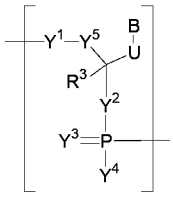
**[0094]** each $Y^5$ is, independently, O, S, Se, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

**[0095]** n is an integer from 1 to 100,000; and

**[0096]** wherein the ring including U can include one or more double bonds.

**[0097]** In particular embodiments, the ring including U does not have a double bond between U-$CB^3R^{b3}$ or between $CB^3R^{b3}$-$C^{B2}R^{b2}$.

**[0098]** In some embodiments, the polynucleotide (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (Id):




(Id), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein U is O, S, $N(R^U)_{nu}$, or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

**[0099]** each $R^3$ is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl;

**[00100]** each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, -$NR^{N1}$-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl;

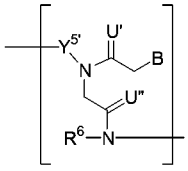
**[00101]** each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino;

**[00102]** each $Y^5$ is, independently, O, S, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

**[00103]** n is an integer from 1 to 100,000; and

**[00104]** B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof).

**[00105]** In some embodiments, the polynucleotide (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (Ie):



(Ie), or a pharmaceutically acceptable salt or stereoisomer thereof,

**[00106]** wherein each of U' and U" is, independently, O, S, $N(R^U)_{nu}$, or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

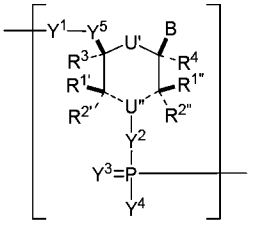
**[00107]** each $R^6$ is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl;

**[00108]** each $Y^{5'}$ is, independently, O, S, optionally substituted alkylene (e.g., methylene or ethylene), or optionally substituted heteroalkylene;

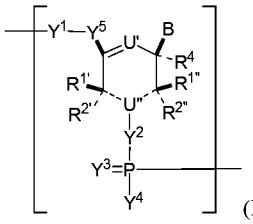
**[00109]** n is an integer from 1 to 100,000; and

**[00110]** B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof).

**[00111]** In some embodiments, the polynucleotide (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (If) or (If-1):



(If),



(If-1), or a pharmaceutically acceptable salt or stereoisomer thereof,

**[00112]** wherein each of U' and U" is, independently, O, S, N, $N(R^U)_{nu}$, or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl (e.g., U' is O and U" is N);

**[00113]** - - - is a single bond or absent;

**[00114]** each of $R^{1'}$, $R^{2'}$, $R^{1"}$, $R^{2"}$, $R^3$, and $R^4$ is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent; and wherein the combination of $R^{1'}$ and $R^3$, the combination of $R^{1"}$ and $R^3$, the combination of $R^{2'}$ and $R^3$, or the combination of $R^{2"}$ and $R^3$ can be taken together to form optionally substituted alkylene or optionally substituted heteroalkylene (e.g., to produce a locked nucleic acid);each of m' and m" is, independently, an integer from 0 to 3 (e.g., from 0 to 2, from 0 to 1, from 1 to 3, or from 1 to 2);

**[00115]** each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, $-NR^{N1}-$, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, or absent;

**[00116]** each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino;

**[00117]** each $Y^5$ is, independently, O, S, Se, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

**[00118]** n is an integer from 1 to 100,000; and

**[00119]** B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof).

**[00120]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), the ring including U has one or two double bonds.

**[00121]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), each of $R^1$, $R^{1'}$, and $R^{1"}$, if present, is H. In further embodiments, each of $R^2$, $R^{2'}$, and $R^{2"}$, if present, is,

independently, H, halo (e.g., fluoro), hydroxy, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy. In particular embodiments, alkoxyalkoxy is -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl). In some embodiments, s2 is 0, s1 is 1 or 2, s3 is 0 or 1, and R' is $C_{1-6}$ alkyl.

**[00122]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), each of $R^2$, $R^{2'}$, and $R^{2''}$, if present, is H. In further embodiments, each of $R^1$, $R^{1'}$, and $R^{1''}$, if present, is, independently, H, halo (e.g., fluoro), hydroxy, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy. In particular embodiments, alkoxyalkoxy is -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl). In some embodiments, s2 is 0, s1 is 1 or 2, s3 is 0 or 1, and R' is $C_{1-6}$ alkyl.

**[00123]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), each of $R^3$, $R^4$, and $R^5$ is, independently, H, halo (e.g., fluoro), hydroxy, optionally substituted alkyl, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy. In particular embodiments, $R^3$ is H, $R^4$ is H, $R^5$ is H, or $R^3$, $R^4$, and $R^5$ are all H. In particular embodiments, $R^3$ is $C_{1-6}$ alkyl, $R^4$ is $C_{1-6}$ alkyl, $R^5$ is $C_{1-6}$ alkyl, or $R^3$, $R^4$, and $R^5$ are all $C_{1-6}$ alkyl. In particular embodiments, $R^3$ and $R^4$ are both H, and $R^5$ is $C_{1-6}$ alkyl.

**[00124]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), $R^3$ and $R^5$ join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl (e.g., a bicyclic, tricyclic, or tetracyclic heterocyclyl, such as trans-3',4' analogs, wherein $R^3$ and $R^5$ join together to form heteroalkylene (e.g., -$(CH_2)_{b1}O(CH_2)_{b2}O(CH_2)_{b3}$-, wherein each of b1, b2, and b3 are, independently, an integer from 0 to 3).

**[00125]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), $R^3$ and one or

more of $R^{1'}$, $R^{1''}$, $R^{2'}$, $R^{2''}$, or $R^5$ join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl (e.g., a bicyclic, tricyclic, or tetracyclic heterocyclyl, $R^3$ and one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, $R^{2''}$, or $R^5$ join together to form heteroalkylene (e.g., -$(CH_2)_{b1}O(CH_2)_{b2}O(CH_2)_{b3}$-, wherein each of b1, b2, and b3 are, independently, an integer from 0 to 3).

**[00126]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), $R^5$ and one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, or $R^{2''}$ join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl (e.g., a bicyclic, tricyclic, or tetracyclic heterocyclyl, $R^5$ and one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, or $R^{2''}$ join together to form heteroalkylene (e.g., -$(CH_2)_{b1}O(CH_2)_{b2}O(CH_2)_{b3}$-, wherein each of b1, b2, and b3 are, independently, an integer from 0 to 3).

**[00127]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), each $Y^2$ is, independently, O, S, or -$NR^{N1}$-, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl. In particular embodiments, $Y^2$ is $NR^{N1}$-, wherein $R^{N1}$ is H or optionally substituted alkyl (e.g., $C_{1-6}$ alkyl, such as methyl, ethyl, isopropyl, or n-propyl).

**[00128]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), each $Y^3$ is, independently, O or S.

**[00129]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), $R^1$ is H; each $R^2$ is, independently, H, halo (e.g., fluoro), hydroxy, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy (e.g., -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, such as wherein s2 is 0, s1 is 1 or 2, s3 is 0 or 1, and R' is $C_{1-6}$ alkyl); each $Y^2$ is, independently, O or -$NR^{N1}$-, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted

alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein $R^{N1}$ is H or optionally substituted alkyl (e.g., $C_{1-6}$ alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each $Y^3$ is, independently, O or S (e.g., S). In further embodiments, $R^3$ is H, halo (e.g., fluoro), hydroxy, optionally substituted alkyl, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy. In yet further embodiments, each $Y^1$ is , independently, O or -$NR^{N1}$-, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein $R^{N1}$ is H or optionally substituted alkyl (e.g., $C_{1-6}$ alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each $Y^4$ is, independently, H, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino.

**[00130]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), each $R^1$ is, independently, H, halo (e.g., fluoro), hydroxy, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy (e.g., $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, such as wherein s2 is 0, s1 is 1 or 2, s3 is 0 or 1, and R' is $C_{1-6}$ alkyl); $R^2$ is H; each $Y^2$ is, independently, O or -$NR^{N1}$-, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein $R^{N1}$ is H or optionally substituted alkyl (e.g., $C_{1-6}$ alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each $Y^3$ is, independently, O or S (e.g., S). In further embodiments, $R^3$ is H, halo (e.g., fluoro), hydroxy, optionally substituted alkyl, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy. In yet further embodiments, each $Y^1$ is , independently, O or -$NR^{N1}$-, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein $R^{N1}$ is H or optionally substituted alkyl (e.g., $C_{1-6}$ alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each $Y^4$ is, independently, H, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino.

**[00131]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), the ring including U is in the β-D (e.g., β-D-ribo) configuration.

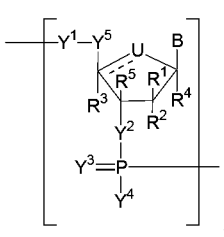
**[00132]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), the ring including U is in the α-L (e.g., α-L-ribo) configuration.

**[00133]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), one or more B is not pseudouridine (ψ) or 5-methyl-cytidine ($m^5C$).

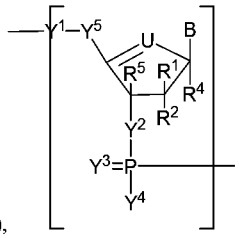
**[00134]** In some embodiments, about 10% to about 100% of n number of B nucleobases is not ψ or $m^5C$ (e.g., from 10% to 20%, from 10% to 35%, from 10% to 50%, from 10% to 60%, from 10% to 75%, from 10% to 90%, from 10% to 95%, from 10% to 98%, from 10% to 99%, from 20% to 35%, from 20% to 50%, from 20% to 60%, from 20% to 75%, from 20% to 90%, from 20% to 95%, from 20% to 98%, from 20% to 99%, from 20% to 100%, from 50% to 60%, from 50% to 75%, from 50% to 90%, from 50% to 95%, from 50% to 98%, from 50% to 99%, from 50% to 100%, from 75% to 90%, from 75% to 95%, from 75% to 98%, from 75% to 99%, and from 75% to 100% of n number of B is not ψ or $m^5C$). In some embodiments, B is not ψ or $m^5C$.

**[00135]** In some embodiments of the polynucleotides (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-1), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr)), when B is an unmodified nucleobase selected from cytosine, guanine, uracil and adenine, then at least one of $Y^1$, $Y^2$, or $Y^3$ is not O.

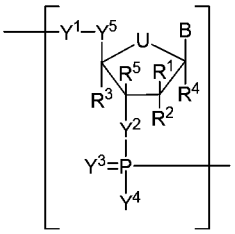
**[00136]** In some embodiments, the polynucleotide includes a modified ribose. In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIa)-(IIc):




. (IIa),

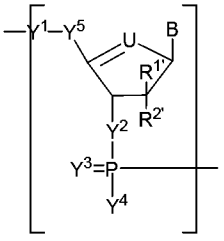



(IIb), or

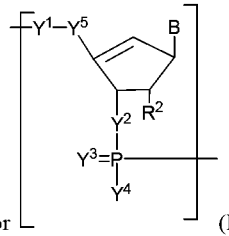


(IIc), or a pharmaceutically acceptable salt or stereoisomer thereof. In particular embodiments, U is O or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl (e.g., U is $-CH_2-$ or –CH–). In other embodiments, each of $R^1$, $R^2$, $R^3$, $R^4$, and $R^5$ is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent (e.g., each $R^1$ and $R^2$ is, independently H, halo, hydroxy, optionally substituted alkyl, or optionally substituted alkoxy; each $R^3$ and $R^4$ is, independently, H or optionally substituted alkyl; and $R^5$ is H or hydroxy), and - - - is a single bond or double bond.

**[00137]** In particular embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIb-1)-(IIb-2):



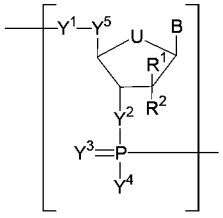
(IIb-1) or



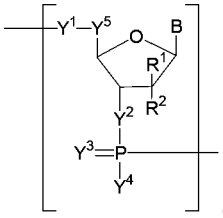
(IIb-2) or a pharmaceutically acceptable salt or stereoisomer thereof. In some embodiments, U is O or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl (e.g., U is $-CH_2-$ or –CH–). In other embodiments, each of $R^1$ and $R^2$ is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally

substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent (e.g., each $R^1$ and $R^2$ is, independently, H, halo, hydroxy, optionally substituted alkyl, or optionally substituted alkoxy, e.g., H, halo, hydroxy, alkyl, or alkoxy). In particular embodiments, $R^2$ is hydroxy or optionally substituted alkoxy (e.g., methoxy, ethoxy, or any described herein).

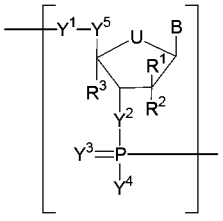
**[00138]** In particular embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIc-1)-(IIc-4):



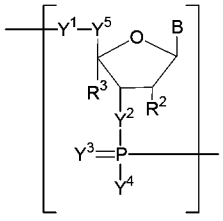

(IIc-1),




(IIc-2),




(IIc-3), or

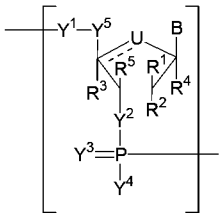



(IIc-4), or a pharmaceutically acceptable salt or stereoisomer thereof.

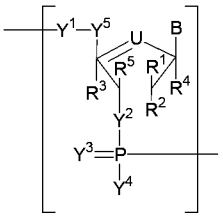
**[00139]** In some embodiments, U is O or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl (e.g., U is $–CH_2–$ or –CH–). In some embodiments, each of $R^1$, $R^2$, and $R^3$ is, independently, H, halo, hydroxy, thio[l], optionally substituted alkyl, optionally substitut[e]d alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, opti[o]nal[l]y subst[i]tuted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent (e.g., each $R^1$ and $R^2$ is, independently, H, halo, hydroxy, optionally substituted alkyl, or optionally substituted alkoxy, e.g., H, halo, hydroxy, alkyl, or alkoxy; and each $R^3$ is, independently, H or optionally substituted alkyl)). In particular embodiments, $R^2$ is optionally

substituted alkoxy (e.g., methoxy or ethoxy, or any described herein). In particular embodiments, $R^1$ is optionally substituted alkyl, and $R^2$ is hydroxy. In other embodiments, $R^1$ is hydroxy, and $R^2$ is optionally substituted alkyl. In further embodiments, $R^3$ is optionally substituted alkyl.

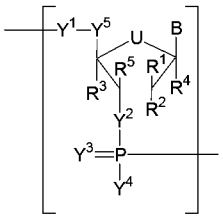
**[00140]** In some embodiments, the polynucleotide includes an acyclic modified ribose. In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IId)-(IIf):



(IId),

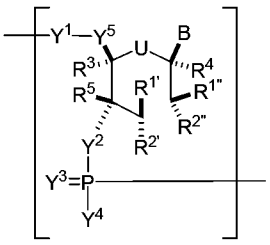


(IIe), or

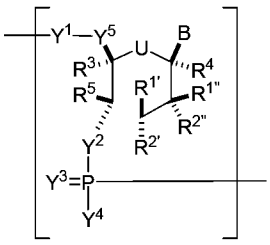


(IIf), or a pharmaceutically acceptable salt or stereoisomer thereof.

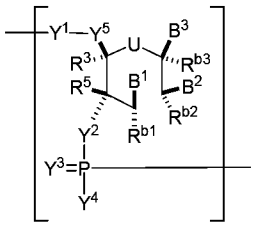
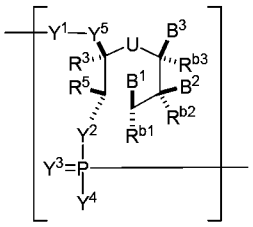
**[00141]** In some embodiments, the polynucleotide includes an acyclic modified hexitol. In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIg)-(IIj):



(IIg),

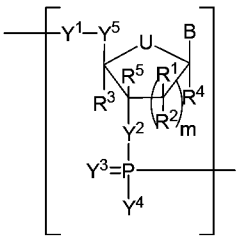
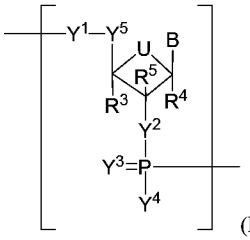


(IIh),

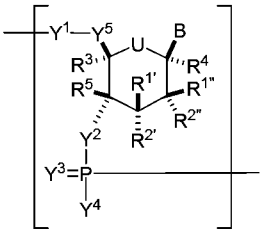



(IIi), or (IIj), or a pharmaceutically acceptable salt or stereoisomer thereof.

**[00142]** In some embodiments, the polynucleotide includes a sugar moiety having a contracted or an expanded ribose ring. In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIk)-(IIm):

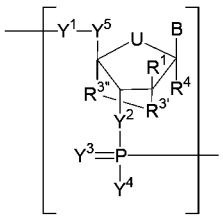



(IIk), (III), or



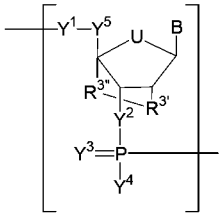
(IIm), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each of $R^{1'}$, $R^{1''}$, $R^{2'}$, and $R^{2''}$ is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, or absent; and wherein the combination of $R^{2'}$ and $R^3$ or the combination of $R^{2''}$ and $R^3$ can be taken together to form optionally substituted alkylene or optionally substituted heteroalkylene.

**[00143]** In some embodiments, the polynucleotide includes a locked modified ribose. In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIn):

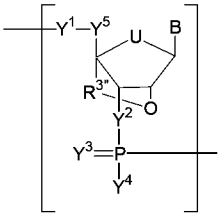


(IIn), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $R^{3'}$ is O, S, or $-NR^{N1}-$, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl and $R^{3''}$ is optionally substituted alkylene (e.g., $-CH_2-$, $-CH_2CH_2-$, or $-CH_2CH_2CH_2-$) or optionally substituted heteroalkylene (e.g., $-CH_2NH-$, $-CH_2CH_2NH-$, $-CH_2OCH_2-$, or $-CH_2CH_2OCH_2-$) (e.g., $R^{3'}$ is O and $R^{3''}$ is optionally substituted alkylene (e.g., $-CH_2-$, $-CH_2CH_2-$, or $-CH_2CH_2CH_2-$)).

**[00144]** In some embodiments, the polynucleotide (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIn-1)-(IIn2):



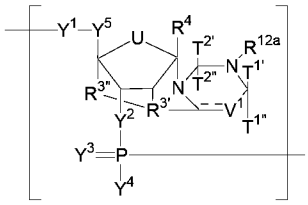
(IIn-1) or



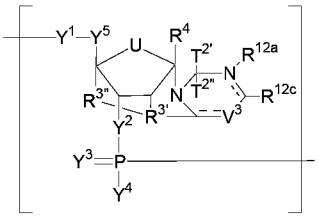
(IIn-2), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $R^{3'}$ is O, S, or $-NR^{N1}-$, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl and $R^{3''}$ is optionally substituted alkylene (e.g., $-CH_2-$, $-CH_2CH_2-$, or $-CH_2CH_2CH_2-$) or optionally substituted heteroalkylene (e.g., $-CH_2NH-$, $-CH_2CH_2NH-$, $-CH_2OCH_2-$, or $-CH_2CH_2OCH_2-$) (e.g., $R^{3'}$ is O and $R^{3''}$ is optionally substituted alkylene (e.g., $-CH_2-$, $-CH_2CH_2-$, or $-CH_2CH_2CH_2-$)).

**[00145]** In some embodiments, the polynucleotide includes a locked modified ribose that forms a tetracyclic heterocyclyl. In some embodiments, the polynucleotide (e.g., the first region, the first

flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (IIo):



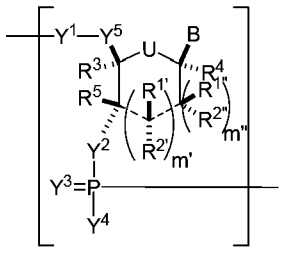
(IIo) or



(IIp), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $R^{12a}$, $R^{12c}$, $T^{1'}$, $T^{1''}$,$T^{2'}$, $T^{2''}$, $V^1$, and $V^3$ are as described herein.

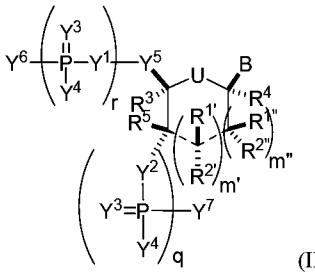
**[00146]** Any of the formulas for the polynucleotides can include one or more nucleobases described herein (e.g., Formulas (b1)-(b43)).

**[00147]** In one embodiment, the present invention provides methods of preparing a polynucleotide comprising at least one nucleotide that disrupts binding of a major groove interacting partner with the nucleic acid, wherein the polynucleotide comprises n number of nucleosides having Formula (Ia), as defined herein:



(Ia),

the method comprising reacting a compound of Formula (IIIa), as defined herein:

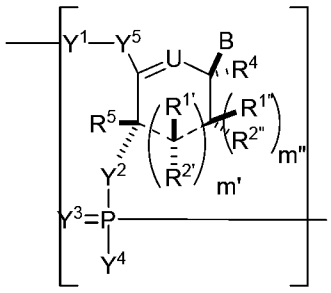


(IIIa),

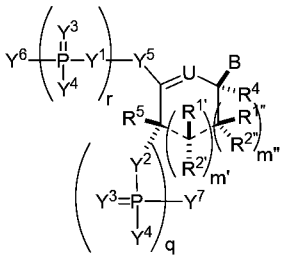
**[00148]** with an RNA polymerase, and a cDNA template.

**[00149]** In a further embodiment, the present invention provides methods of amplifying a polynucleotide comprising at least one nucleotide that disrupts binding of a major groove binding partner with the polynucleotide sequence, the method comprising: reacting a compound of Formula (IIIa), as defined herein, with a primer, a cDNA template, and an RNA polymerase.

**[00150]** In one embodiment, the present invention provides methods of preparing a polynucleotide comprising at least one nucleotide that disrupts binding of a major groove interacting partner with the nucleic acid, wherein the polynucleotide comprises n number of nucleosides having Formula (Ia-1), as defined herein:



(Ia-1), the method comprising reacting a compound of Formula (IIIa-1), as defined herein:

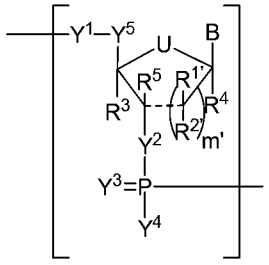


(IIIa-1), with an RNA polymerase, and a cDNA template.

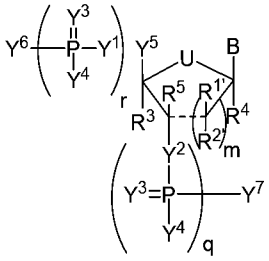
**[00151]** In a further embodiment, the present invention provides methods of amplifying a polynucleotide comprising at least one nucleotide (e.g., modified mRNA molecule) that disrupts binding of a major groove binding partner with the polynucleotide sequence, the method comprising: reacting a compound of Formula (IIIa-1), as defined herein, with a primer, a cDNA template, and an RNA polymerase.

**[00152]** In one embodiment, the present invention provides methods of preparing a polynucleotide comprising at least one nucleotide that disrupts binding of a major groove interacting partner with

the nucleic acid sequence, wherein the polynucleotide comprises n number of nucleosides having Formula (Ia-2), as defined herein:



(Ia-2), the method comprising reacting a compound of Formula (IIIa-2), as defined herein:



(IIIa-2), with an RNA polymerase, and a cDNA template.

**[00153]** In a further embodiment, the present invention provides methods of amplifying a polynucleotide comprising at least one nucleotide (e.g., modified mRNA molecule) that disrupts binding of a major groove binding partner with the polynucleotide, the method comprising reacting a compound of Formula (IIIa-2), as defined herein, with a primer, a cDNA template, and an RNA polymerase.

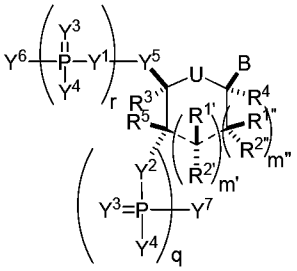
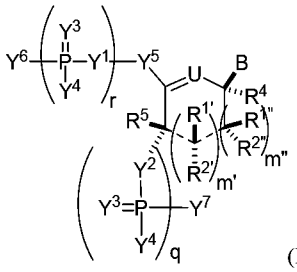
**[00154]** In some embodiments, the reaction may be repeated from 1 to about 7,000 times. In any of the embodiments herein, B may be a nucleobase of Formula (b1)-(b43).

**[00155]** The polynucleotides can optionally include 5′ and/or 3′ flanking regions, which are described herein.

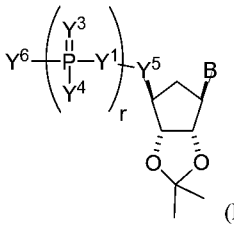
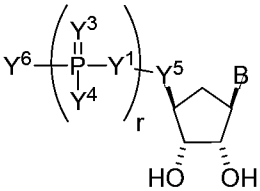
*Modified Nucleotides and Nucleosides*

**[00156]** The present invention also includes the building blocks, e.g., modified ribonucleosides, modified ribonucleotides, of the polynucleotides, e.g., modified RNA ( or mRNA) molecules. For example, these building blocks can be useful for preparing the polynucleotides of the invention.

**[00157]** In some embodiments, the building block molecule has Formula (IIIa) or (IIIa-1):

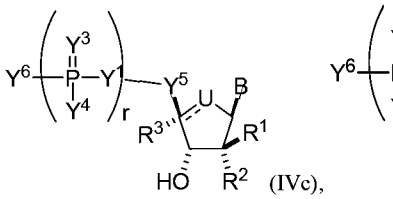

(IIIa),

(IIIa-1) or a pharmaceutically acceptable salt or stereoisomer thereof, wherein the substituents are as described herein (e.g., for Formula (Ia) and (Ia-1)), and wherein when B is an unmodified nucleobase selected from cytosine, guanine, uracil and adenine, then at least one of $Y^1$, $Y^2$, or $Y^3$ is not O.

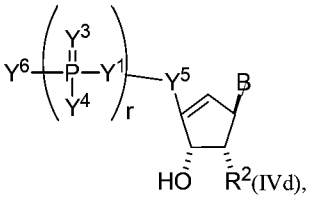
**[00158]** In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, has Formula (IVa)-(IVb):

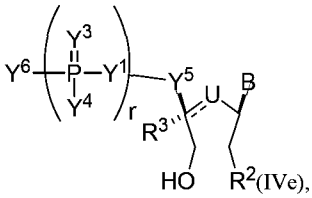

(IVa) or

(IVb), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (b1)-(b43)).

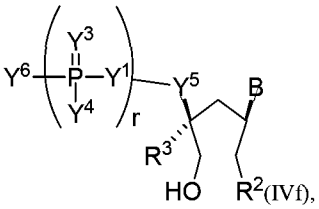
**[00159]** In particular embodiments, Formula (IVa) or (IVb) is combined with a modified uracil (e.g., any one of formulas (b1)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (b1), (b8), (b28), (b29), or (b30)). In particular embodiments, Formula (IVa) or (IVb) is combined with a modified cytosine (e.g., any one of formulas (b10)-(b14), (b24), (b25), and (b32)-(b36), such as formula (b10) or (b32)). In particular embodiments, Formula (IVa) or (IVb) is combined with a modified guanine (e.g., any one of formulas (b15)-(b17) and (b37)-(b40)). In particular embodiments, Formula (IVa) or (IVb) is combined with a modified adenine (e.g., any one of formulas (b18)-(b20) and (b41)-(b43)).

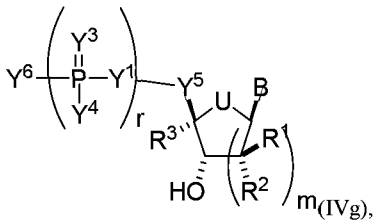
**[00160]** In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, has Formula (IVc)-(IVk):

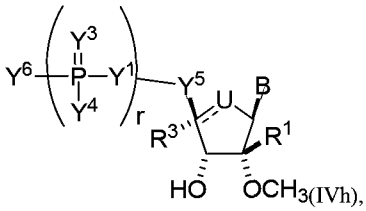

(IVc),

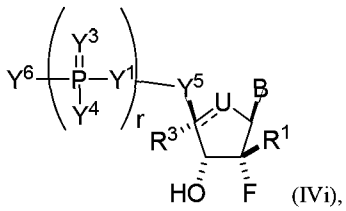

(IVd),

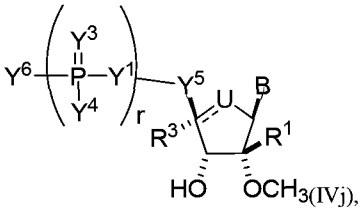

(IVe),

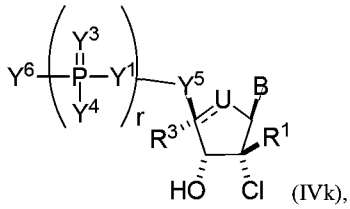

(IVf),

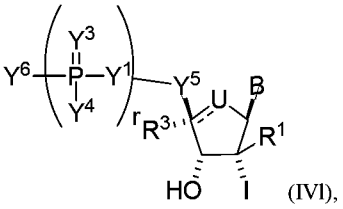

(IVg),


(IVh),


(IVi),


(IVj),


(IVk), or


(IVl),

or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (b1)-(b43)).

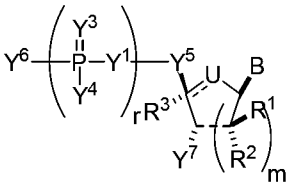
**[00161]** In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified uracil (e.g., any one of formulas (b1)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (b1), (b8), (b28), (b29), or (b30)).

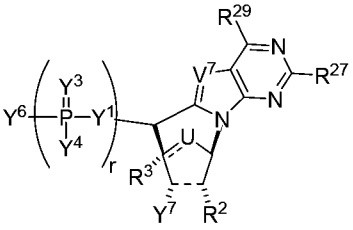
**[00162]** In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified cytosine (e.g., any one of formulas (b10)-(b14), (b24), (b25), and (b32)-(b36), such as formula (b10) or (b32)).

**[00163]** In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified guanine (e.g., any one of formulas (b15)-(b17) and (b37)-(b40)).

**[00164]** In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified adenine (e.g., any one of formulas (b18)-(b20) and (b41)-(b43)).

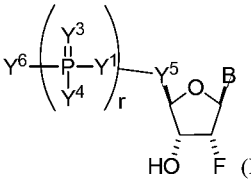
**[00165]** In other embodiments, the building block molecule, which may be incorporated into a polynucleotide has Formula (Va) or (Vb):

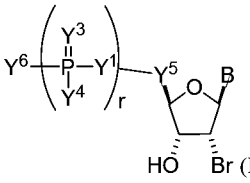

(Va) or

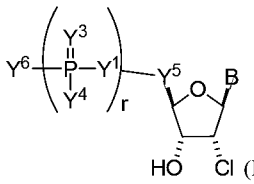
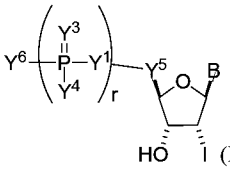

(Vb), or a

pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (b1)-(b43)).

**[00166]** In other embodiments, the building block molecule, which may be incorporated into a polynucleotide has Formula (IXa)-(IXd):


(IXa),


(IXb),

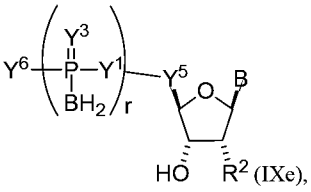
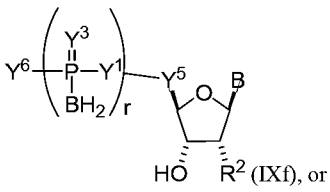
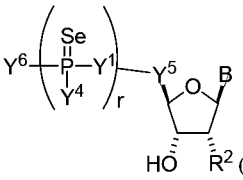

(IXc), or

(IXd), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (b1)-(b43)).

In particular embodiments, one of Formulas (IXa)-(IXd) is combined with a modified uracil (e.g., any one of formulas (b1)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (b1), (b8), (b28), (b29), or (b30)). In particular embodiments, one of Formulas (IXa)-(IXd) is combined with a modified cytosine (e.g., any one of formulas (b10)-(b14), (b24), (b25), and (b32)-(b36), such as formula (b10) or (b32)).

In particular embodiments, one of Formulas (IXa)-(IXd) is combined with a modified guanine (e.g., any one of formulas (b15)-(b17) and (b37)-(b40)).

In particular embodiments, one of Formulas (IXa)-(IXd) is combined with a modified adenine (e.g., any one of formulas (b18)-(b20) and (b41)-(b43)).

**[00167]** In other embodiments, the building block molecule, which may be incorporated into a polynucleotide has Formula (IXe)-(IXg):


(IXe),

(IXf), or

(IXg), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (b1)-(b43)).

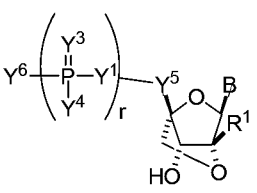
**[00168]** In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified uracil (e.g., any one of formulas (b1)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (b1), (b8), (b28), (b29), or (b30)).

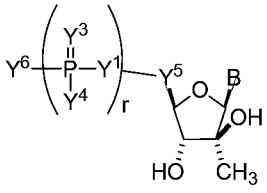
**[00169]** In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified cytosine (e.g., any one of formulas (b10)-(b14), (b24), (b25), and (b32)-(b36), such as formula (b10) or (b32)).

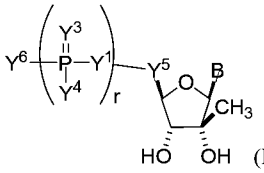
**[00170]** In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified guanine (e.g., any one of formulas (b15)-(b17) and (b37)-(b40)).

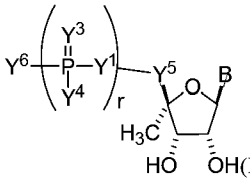
**[00171]** In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified adenine (e.g., any one of formulas (b18)-(b20) and (b41)-(b43)).

**[00172]** In other embodiments, the building block molecule, which may be incorporated into a polynucleotide has Formula (IXh)-(IXk):


(IXh),

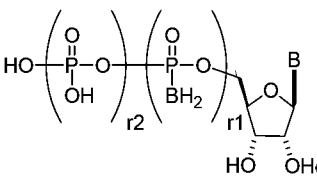

(IXi),

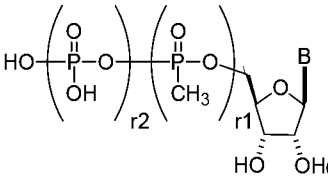

(IXj), or

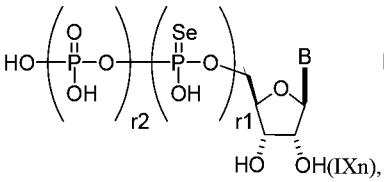

(IXk), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (b1)-(b43)). In particular embodiments, one of Formulas (IXh)-(IXk) is combined with a modified uracil (e.g., any one of formulas (b1)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (b1), (b8), (b28), (b29), or (b30)). In particular embodiments, one of Formulas (IXh)-(IXk) is combined with a modified cytosine (e.g., any one of formulas (b10)-(b14), (b24), (b25), and (b32)-(b36), such as formula (b10) or (b32)).

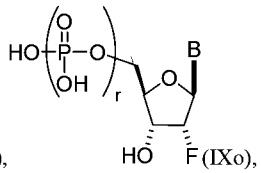
**[00173]** In particular embodiments, one of Formulas (IXh)-(IXk) is combined with a modified guanine (e.g., any one of formulas (b15)-(b17) and (b37)-(b40)). In particular embodiments, one of Formulas (IXh)-(IXk) is combined with a modified adenine (e.g., any one of formulas (b18)-(b20) and (b41)-(b43)).

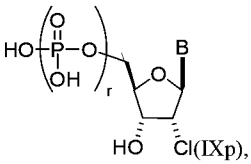
**[00174]** In other embodiments, the building block molecule, which may be incorporated into a polynucleotide has Formula (IXl)-(IXr):

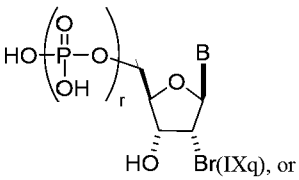

(IXl),

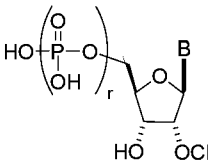

(IXm),


(IXn),


(IXo),


(IXp),


(IXq), or


(IXr) or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r1 and r2 is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5) and B is as described herein (e.g., any one of (b1)-(b43)).

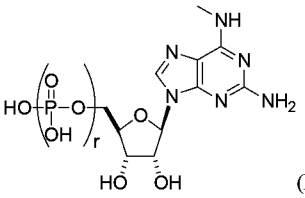
**[00175]** In particular embodiments, one of Formulas (IXl)-(IXr) is combined with a modified uracil (e.g., any one of formulas (b1)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (b1), (b8), (b28), (b29), or (b30)).

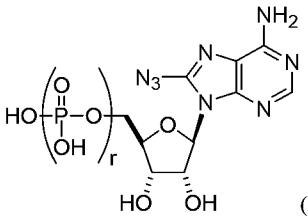
**[00176]** In particular embodiments, one of Formulas (IXl)-(IXr) is combined with a modified cytosine (e.g., any one of formulas (b10)-(b14), (b24), (b25), and (b32)-(b36), such as formula (b10) or (b32)).

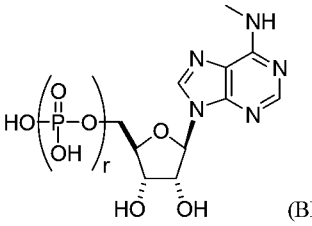
**[00177]** In particular embodiments, one of Formulas (IXl)-(IXr) is combined with a modified guanine (e.g., any one of formulas (b15)-(b17) and (b37)-(b40)). In particular embodiments, one of

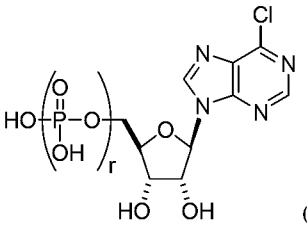
Formulas (IXI)-(IXr) is combined with a modified adenine (e.g., any one of formulas (b18)-(b20) and (b41)-(b43)).

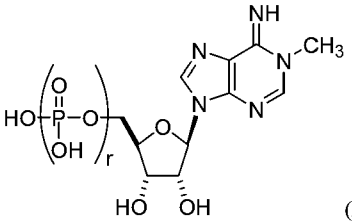
**[00178]** In some embodiments, the building block molecule, which may be incorporated into a polynucleotide can be selected from the group consisting of:

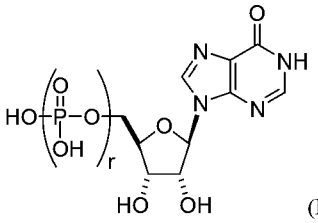

(BB- 1),

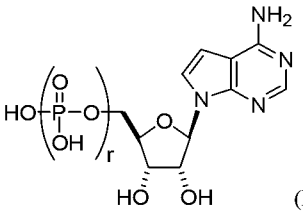

(BB- 2),

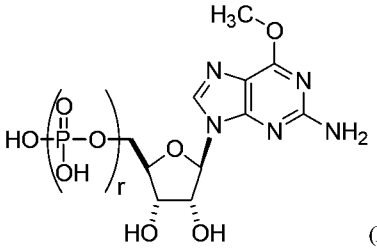

(BB- 3),

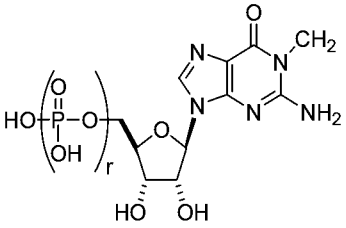

(BB- 4),

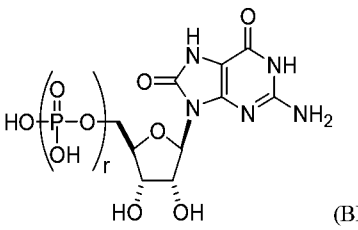

(BB- 5),

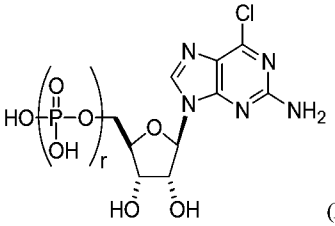

(BB- 6),

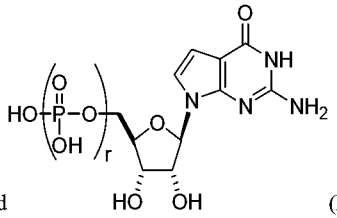

(BB- 7),


(BB- 8),

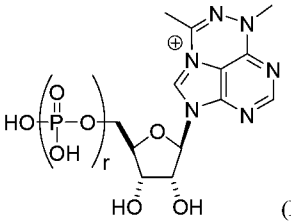

(BB- 9),


(BB- 10),

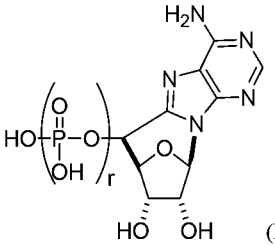

(BB- 11), and


(BB- 12), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5).

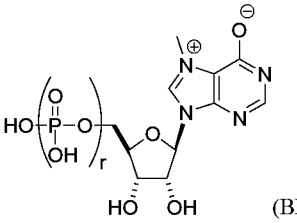
**[00179]** In some embodiments, the building block molecule, which may be incorporated into a polynucleotide can be selected from the group consisting of:



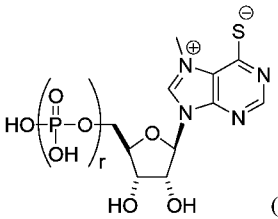
(BB- 13),



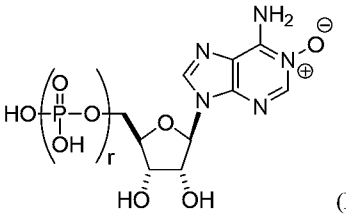
(BB- 14),



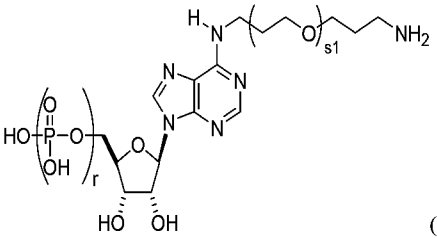
(BB- 15),



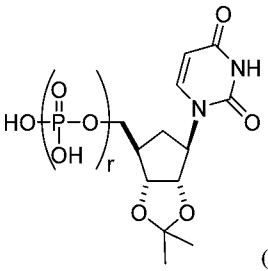
(BB- 16),



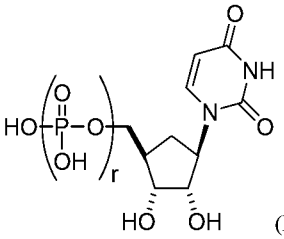
(BB- 17),



(BB- 18),



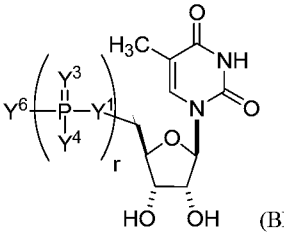
(BB- 19), and



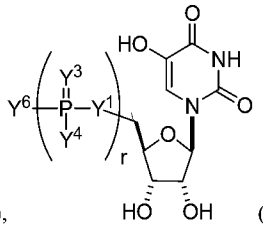
(BB- 20), or a pharmaceutically acceptable salt or stereoisomer

thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5) and s1 is as described herein.

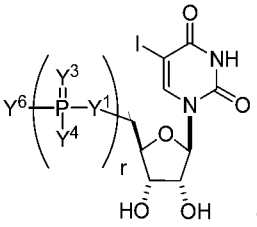
**[00180]** In some embodiments, the building block molecule, which may be incorporated into a nucleic acid (e.g., RNA, mRNA, polynucleotide), is a modified uridine (e.g., selected from the group consisting of:



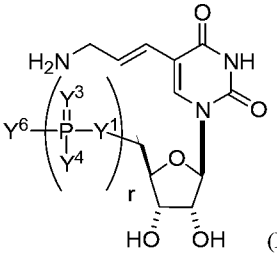
(BB- 21),



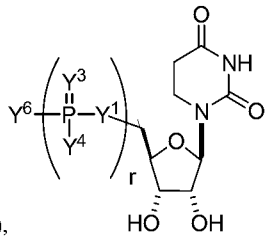
(BB- 22),



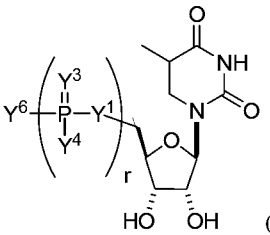
(BB- 23),



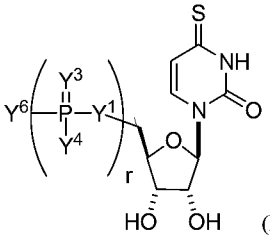
(BB- 24),



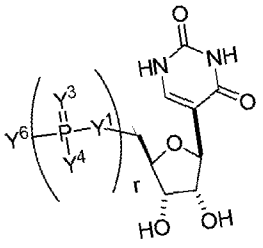
(BB- 25),



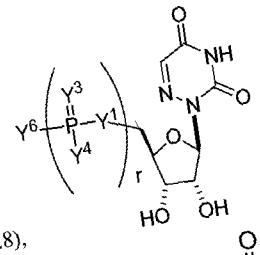
(BB- 26),



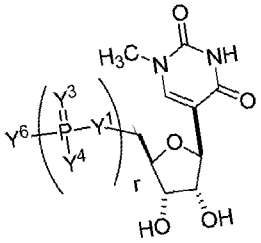
(BB- 27),



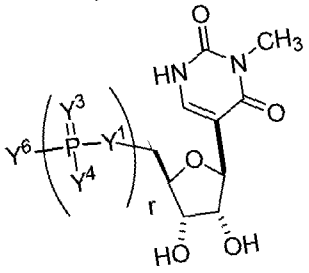
(BB- 28),



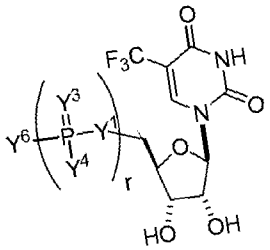
(BB- 29),



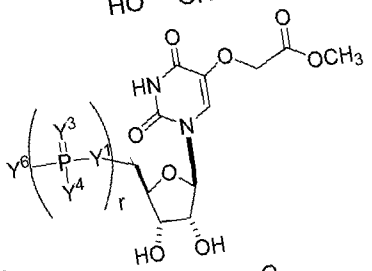
(BB- 30),



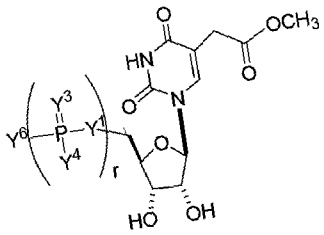
(BB- 31),



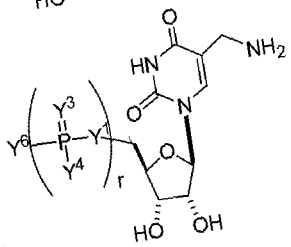
(BB- 32),



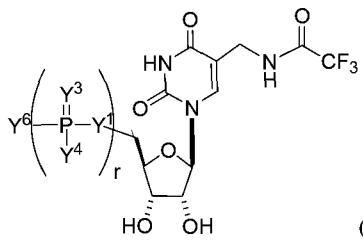
(BB- 33),



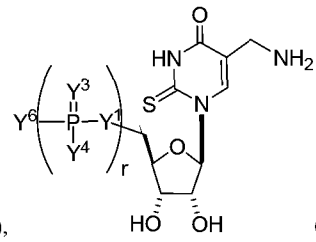
(BB- 34),



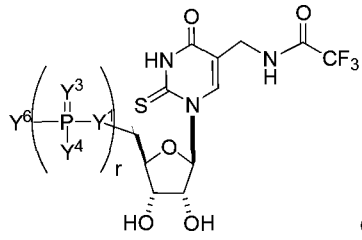
(BB- 35),



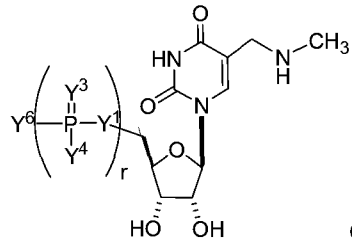
(BB- 36),



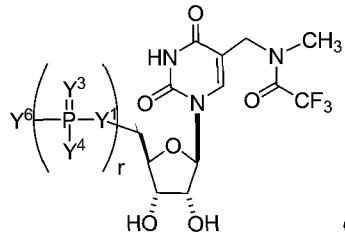
(BB- 37),



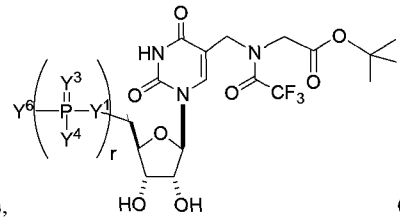
(BB- 38),



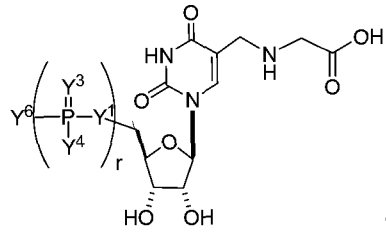
(BB- 39),



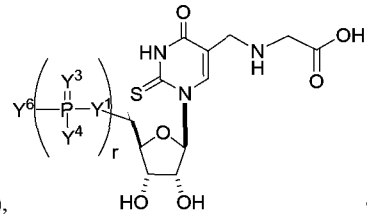
(BB- 40),



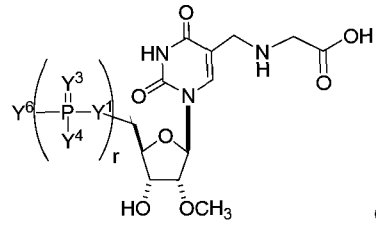
(BB- 41),



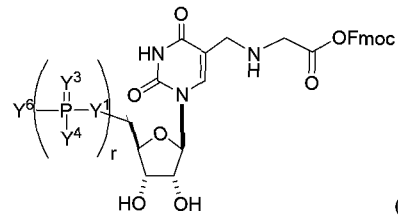
(BB- 42),



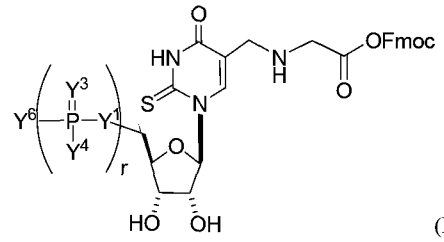
(BB- 43),



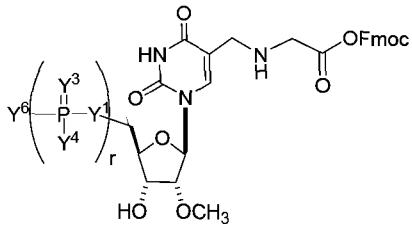
(BB- 44),



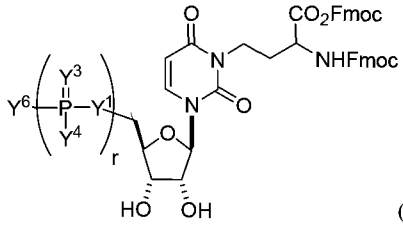
(BB- 45),



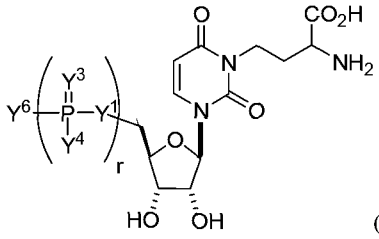
(BB-46),



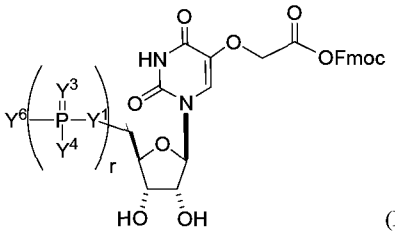
(BB- 47),



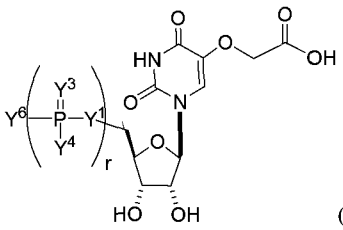
(BB- 48),



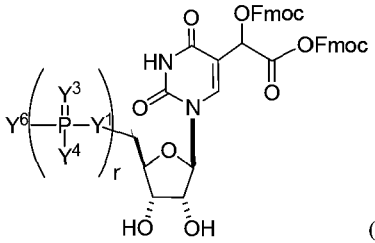
(BB- 49),



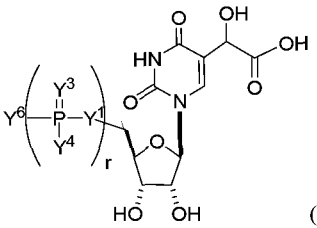
(BB- 50),



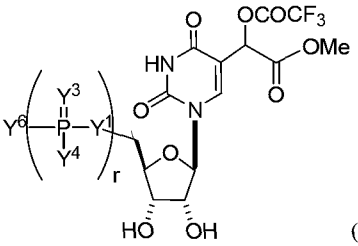
(BB- 51),



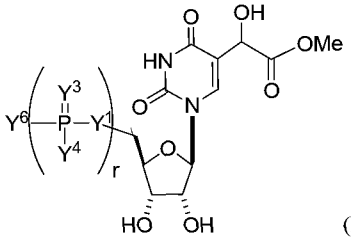
(BB- 52),



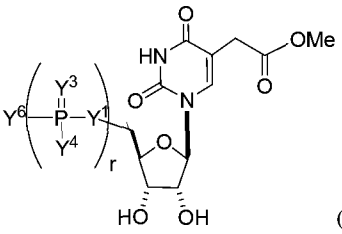
(BB- 53),



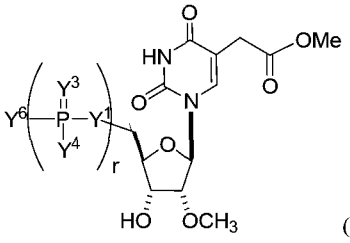
(BB- 54),



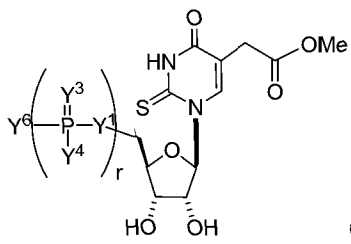
(BB- 55),



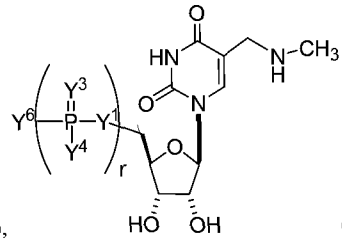
(BB-56),



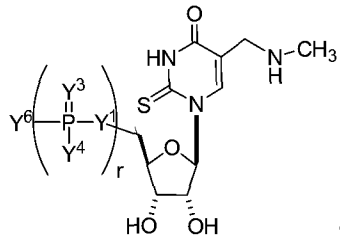
(BB-57),



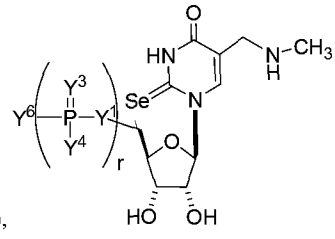
(BB- 58),



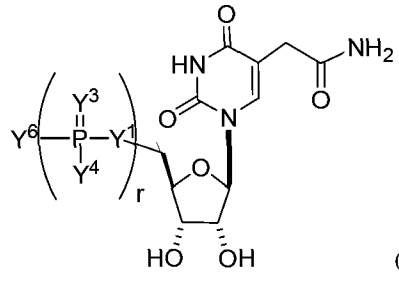
(BB- 59),



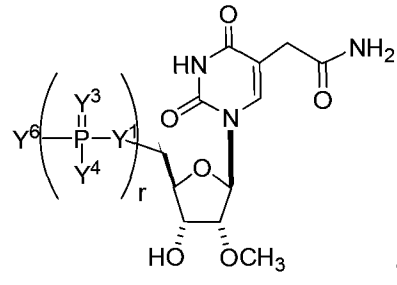
(BB- 60),



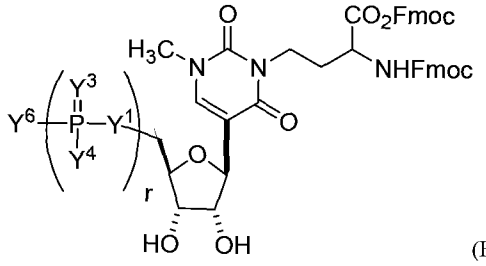
(BB- 61),



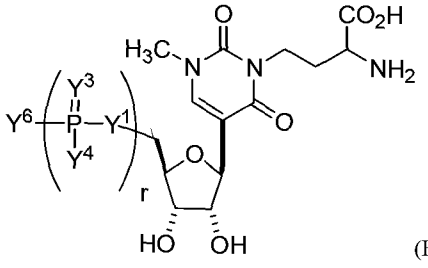
(BB- 62),



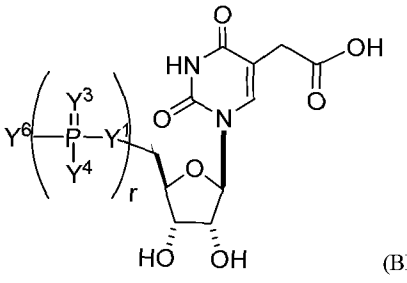
(BB- 63),



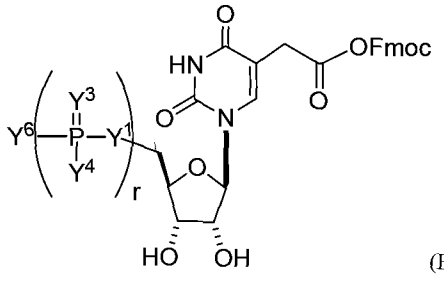
(BB- 64),



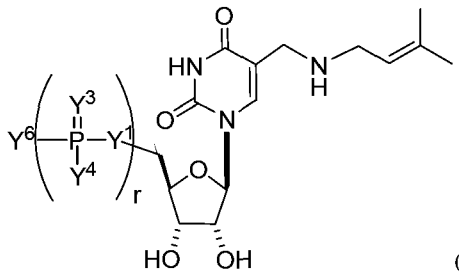
(BB- 65),



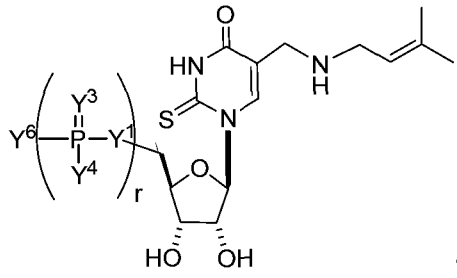
(BB- 66),



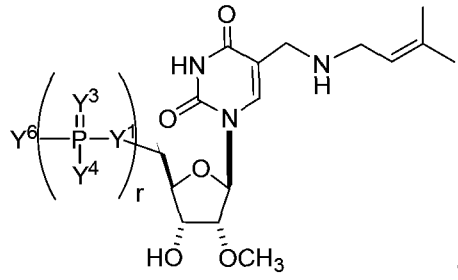
(BB- 67),



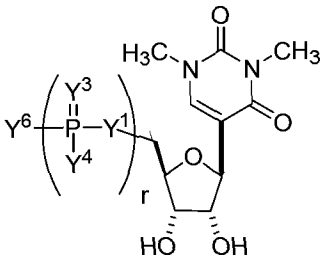
(BB- 68),



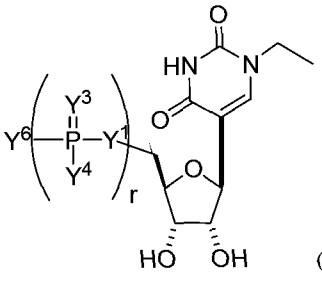
(BB- 69),



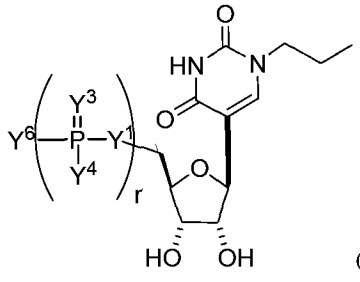
(BB- 70),



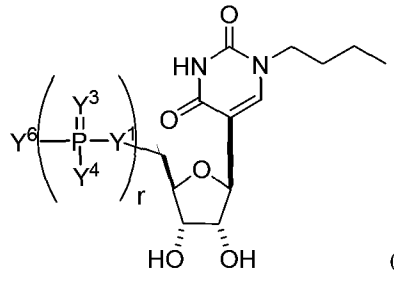
(BB- 71),



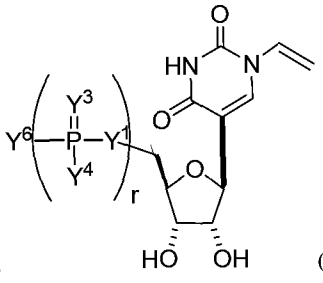
(BB- 72),



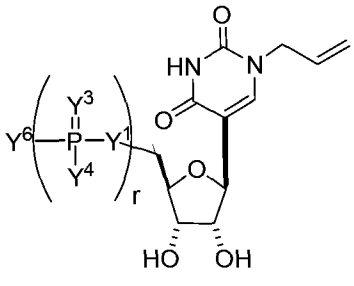
(BB- 73),



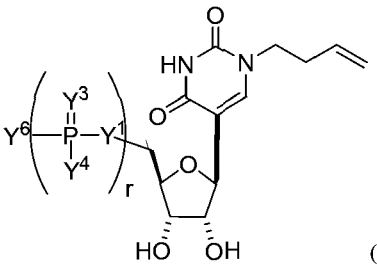
(BB- 74),

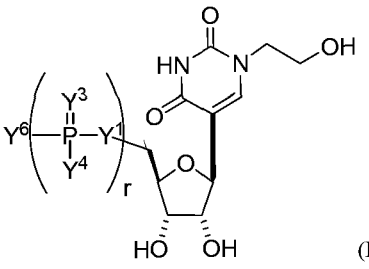


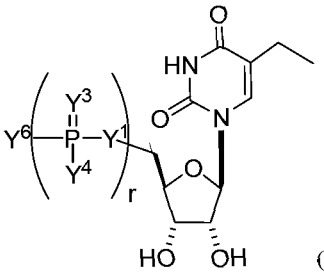
(BB- 75),

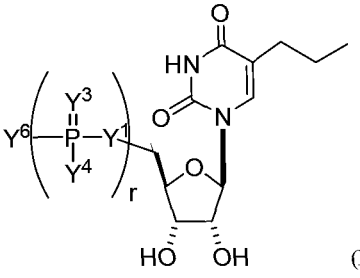


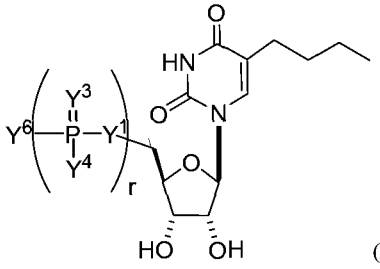
(BB- 76),

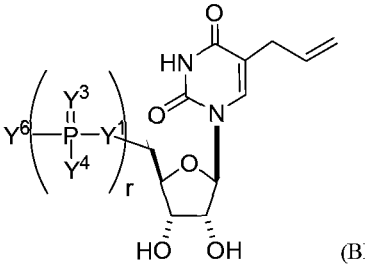

(BB- 77),

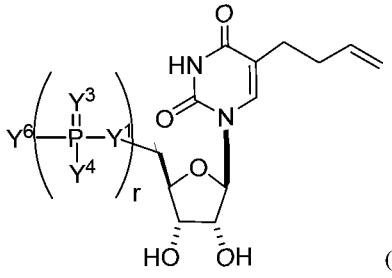

(BB- 78),

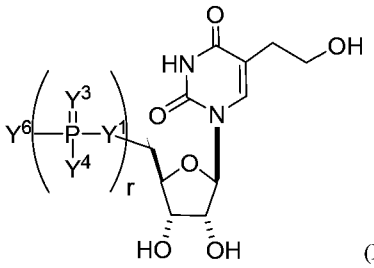

(BB- 79),


(BB- 80),

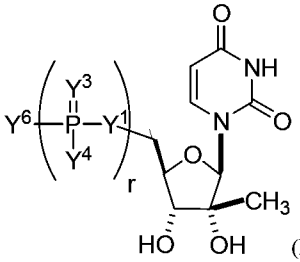

(BB- 81),


(BB- 82),

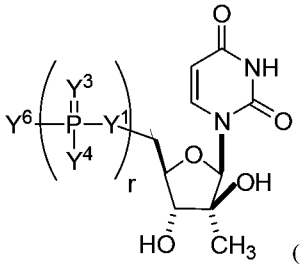

(BB- 83),



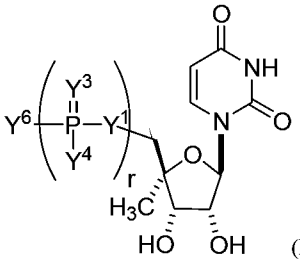
(BB- 84),



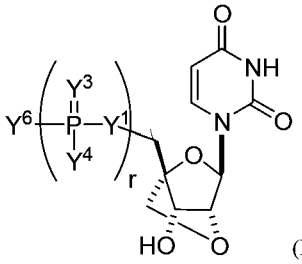
(BB- 85),



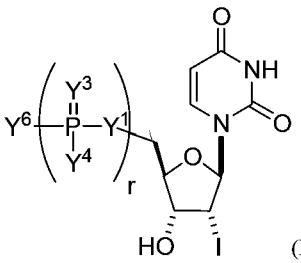
(BB- 86),



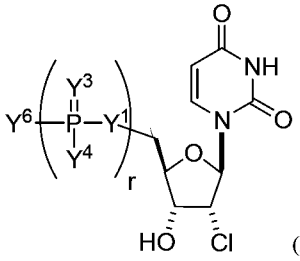
(BB- 87),



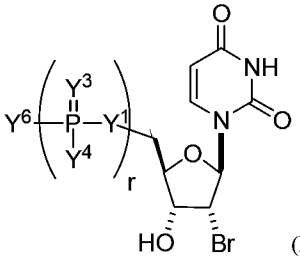
(BB- 88),



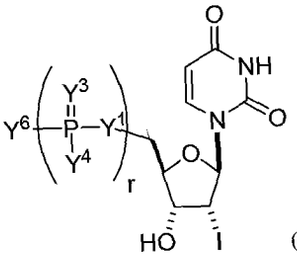
(BB- 89),

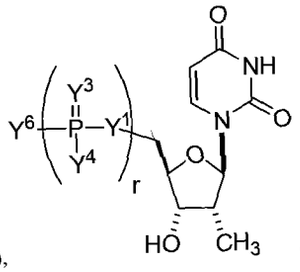


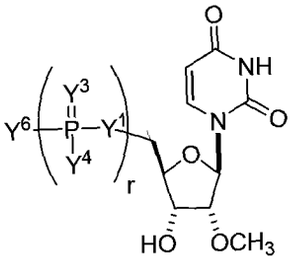
(BB- 90),

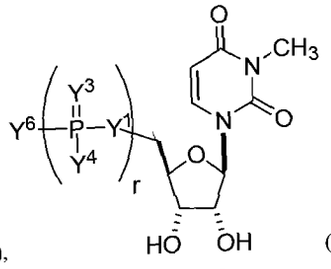


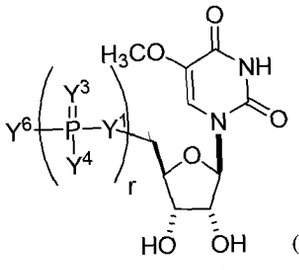
(BB- 91),

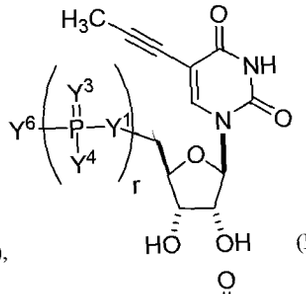

(BB- 92),

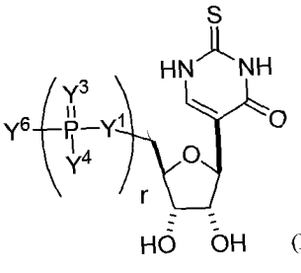

(BB- 93),

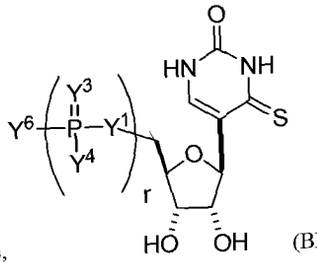

(BB- 94),

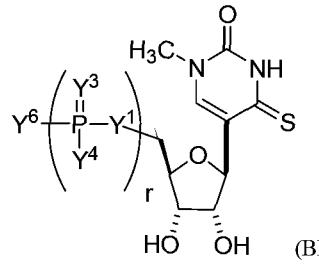

(BB- 95),


(BB- 96),

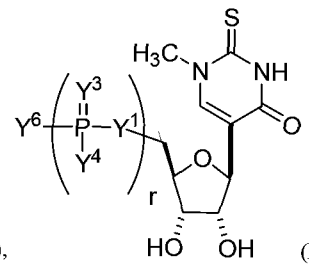

(BB- 97),


(BB- 98),

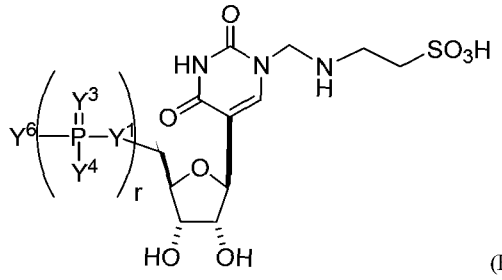

(BB- 99),



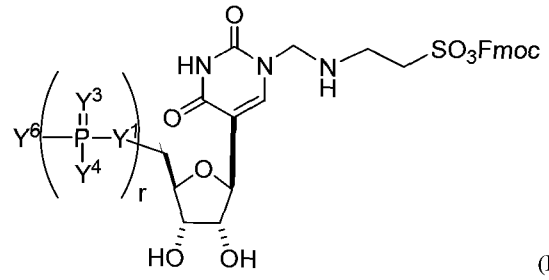
(BB- 100),



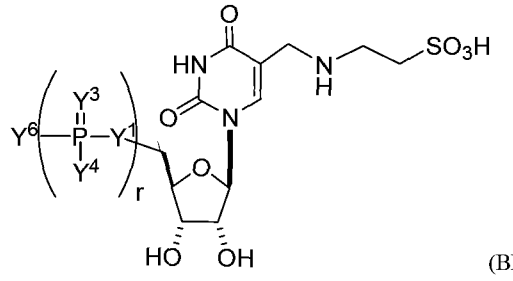
(BB- 101),



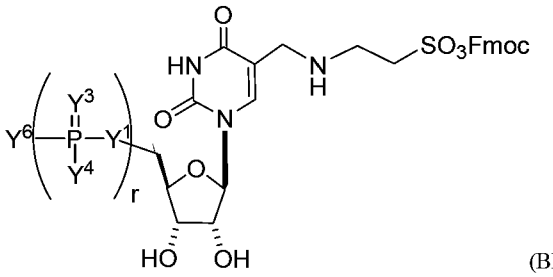
(BB- 102),



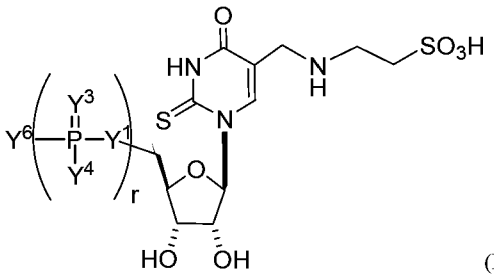
(BB- 103),



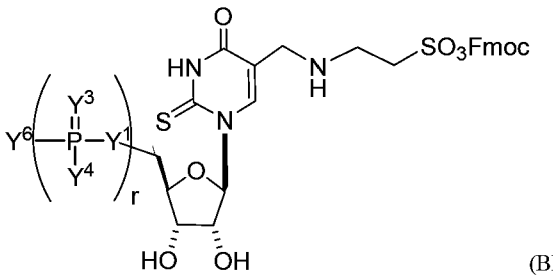
(BB- 104),



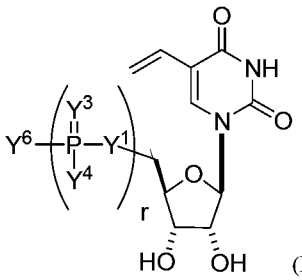
(BB- 105),



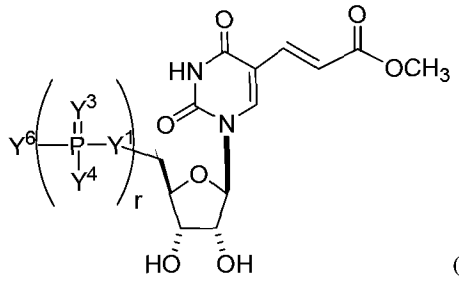
(BB- 106),



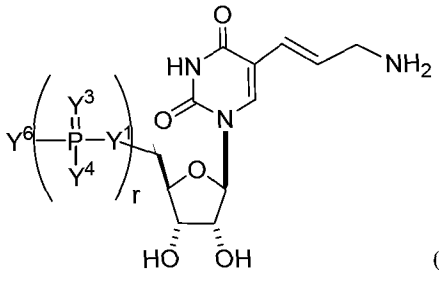
(BB- 107),



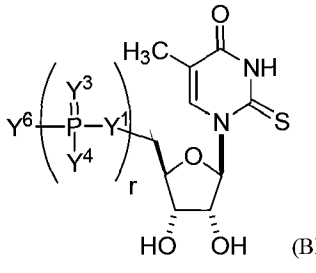
(BB- 108),



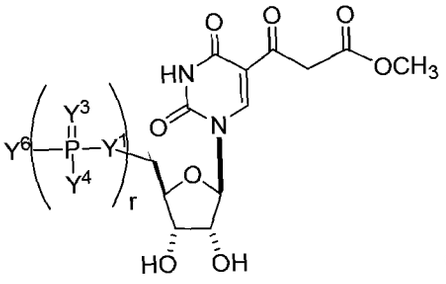
(BB- 109),



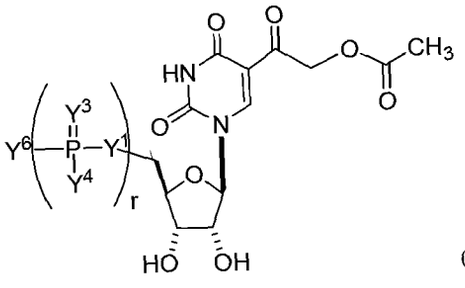
(BB- 110),



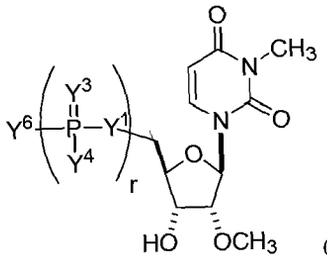
(BB- 111),



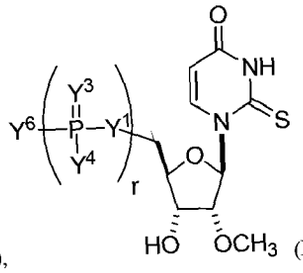
(BB- 112),



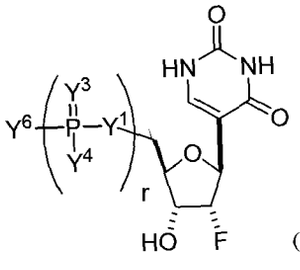
(BB- 113),



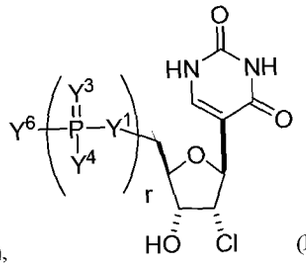
(BB- 114),



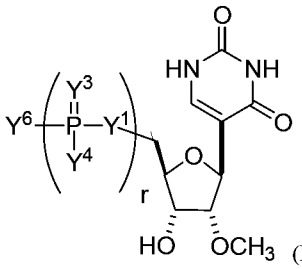
(BB- 115),

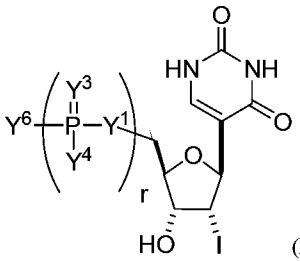


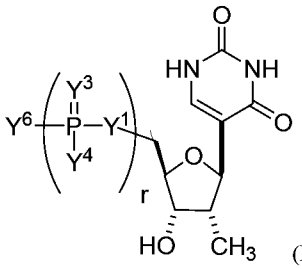
(BB- 116),

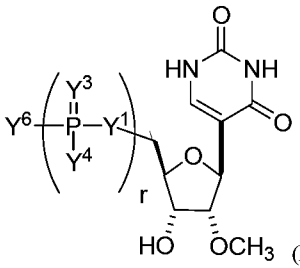


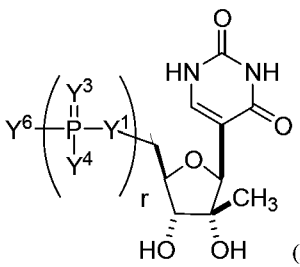
(BB- 117),

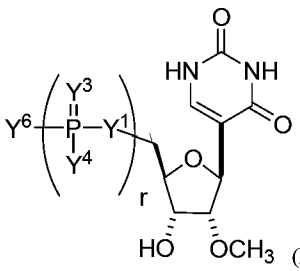

(BB- 118),

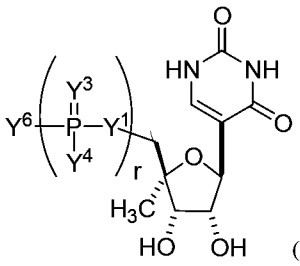

(BB- 119),

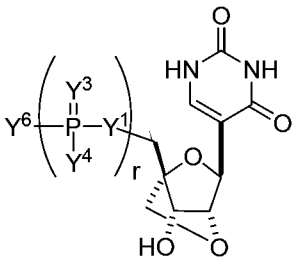

(BB- 120),


(BB- 121),

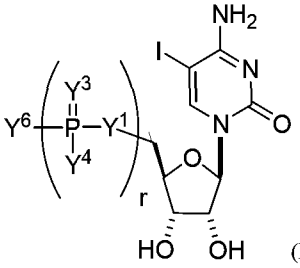

(BB- 122),

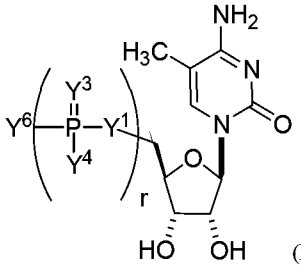

(BB- 123),

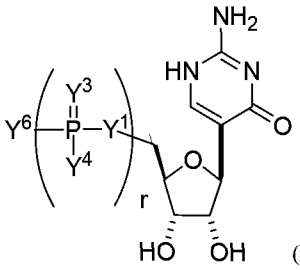

(BB- 124), and

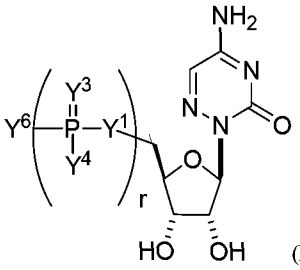

(BB- 125), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $Y^1$, $Y^3$, $Y^4$, $Y^6$, and r are as described herein (e.g., each r is, independently, an integer from 0 to 5, such as from 0 to 3, from 1 to 3, or from 1 to 5)).

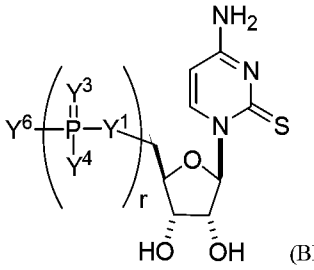
**[00181]** In some embodiments, the building block molecule, which may be incorporated into a polynucleotide is a modified cytidine (e.g., selected from the group consisting of:

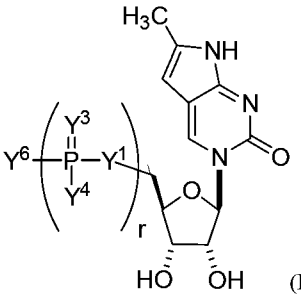

(BB- 126),

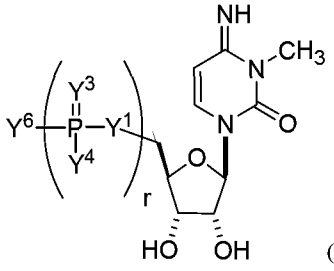

(BB- 127),

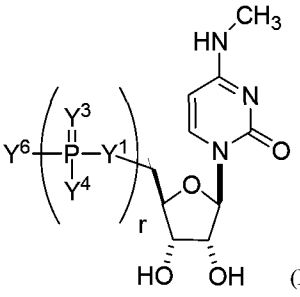

(BB- 128),

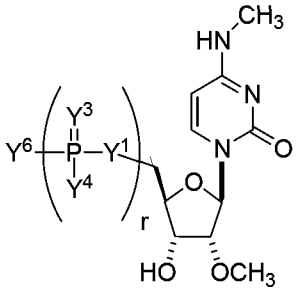

(BB- 129),

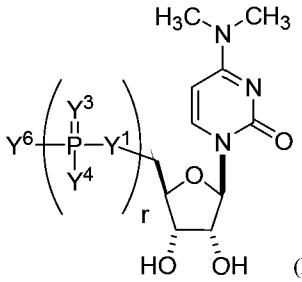

(BB- 130),

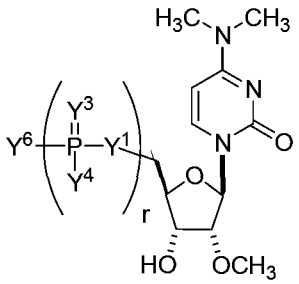

(BB- 131),

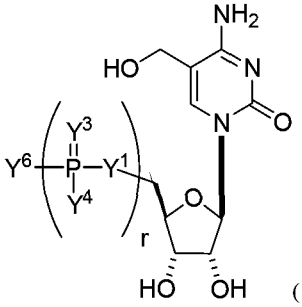

(BB- 132),

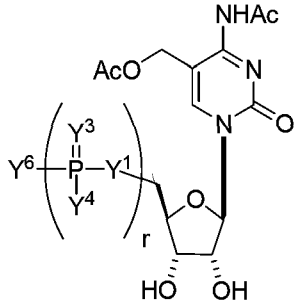

(BB- 133),

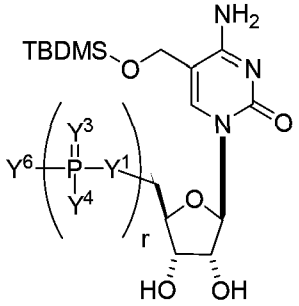

(BB- 134),

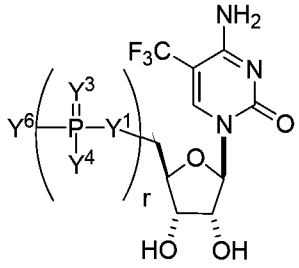

(BB- 135),

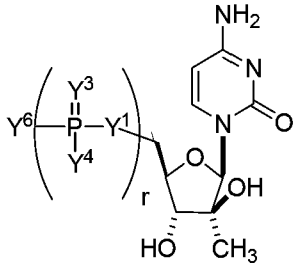

(BB- 136),

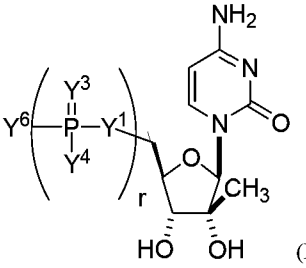

( BB- 137),

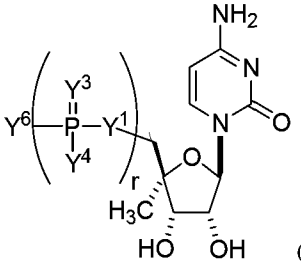

(BB- 138),

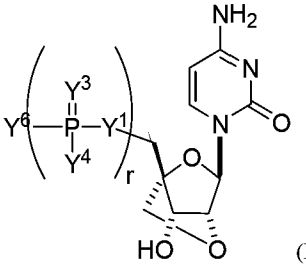

(BB- 139),

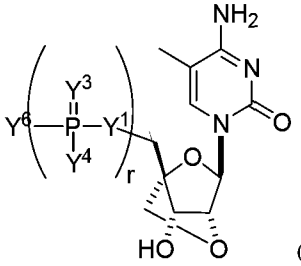

(BB- 140),

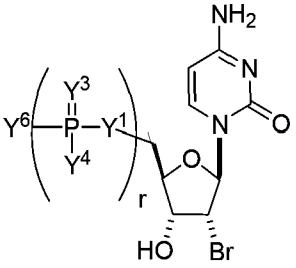

(BB- 141),

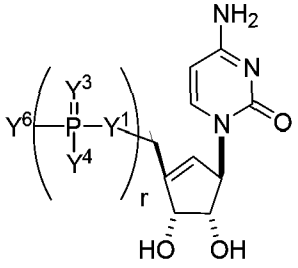

(BB- 142),

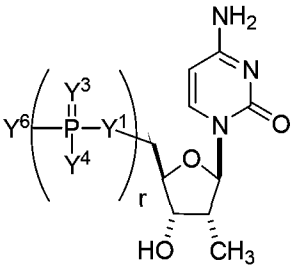

(BB- 143),

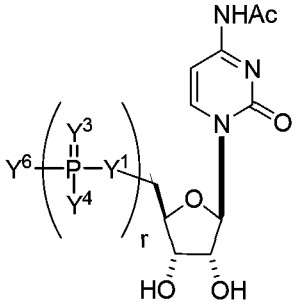

(BB- 144),

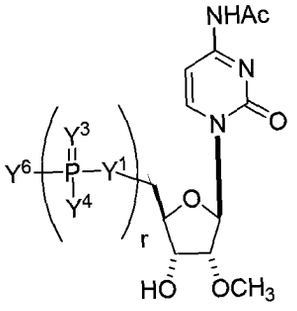

(BB- 145),

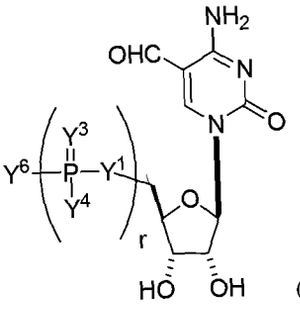

(BB- 146),

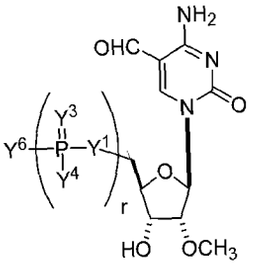

(BB- 147),

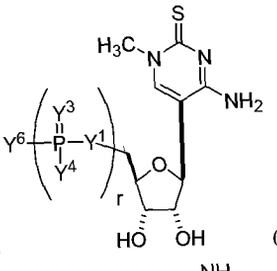

(BB- 148),

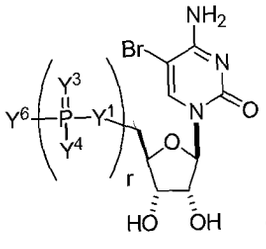

(BB- 149),

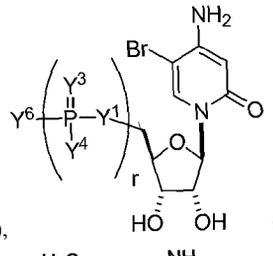

(BB- 150),

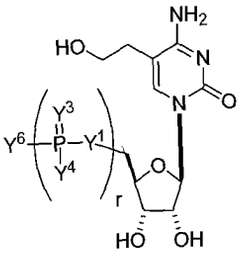

(BB- 151),

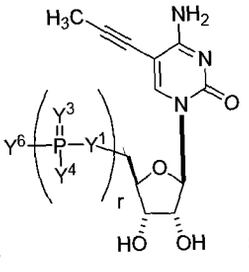

(BB- 152),

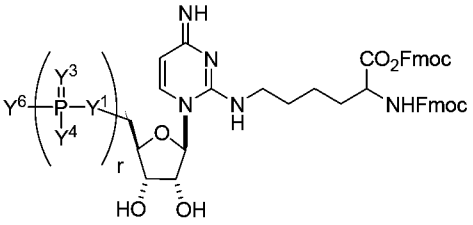

(BB- 153),


(BB- 154),

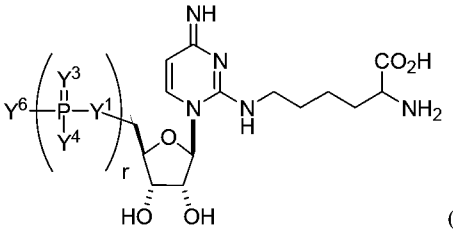

(BB- 155),


(BB- 156),

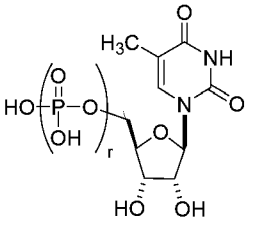

(BB- 157),



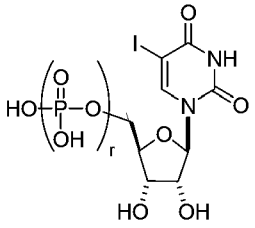
(BB- 158), and



(BB- 159) , or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $Y^1$, $Y^3$, $Y^4$, $Y^6$, and r are as described herein (e.g., each r is, independently, an integer from 0 to 5, such as from 0 to 3, from 1 to 3, or from 1 to 5)). For example, the building block molecule, which may be incorporated into a polynucleotide can be:

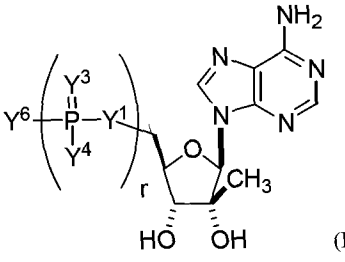


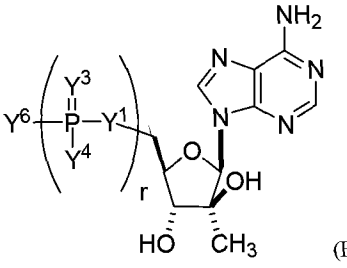
(BB- 160) or

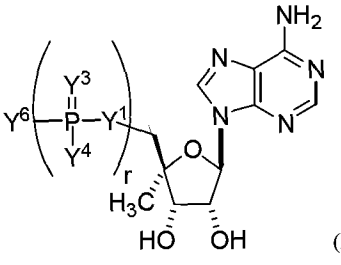


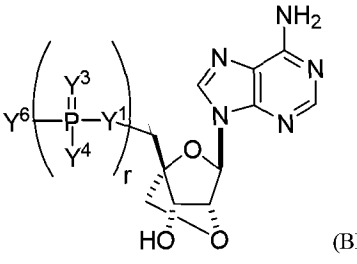
(BB- 161), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5).

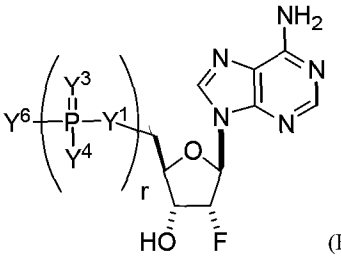
**[00182]** In some embodiments, the building block molecule, which may be incorporated into a polynucleotide is a modified adenosine (e.g., selected from the group consisting of:

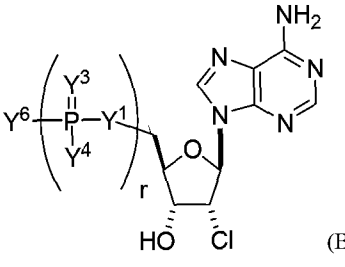

(BB- 162),

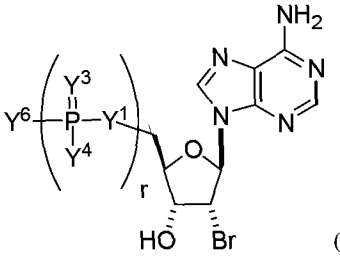

(BB- 163),

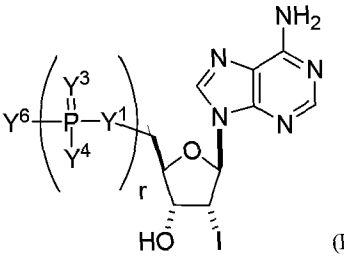

(BB- 164),

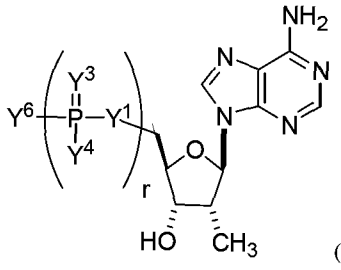

(BB- 165),


(BB- 166),

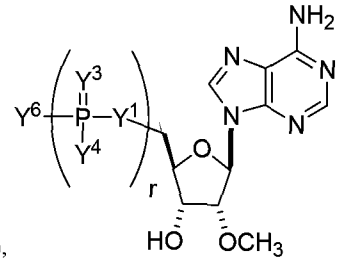

(BB- 167),


(BB- 168),

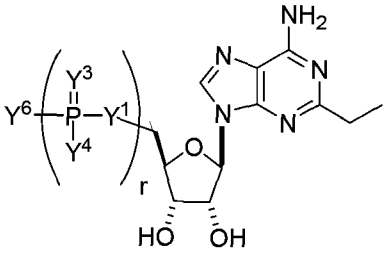

(BB- 169),



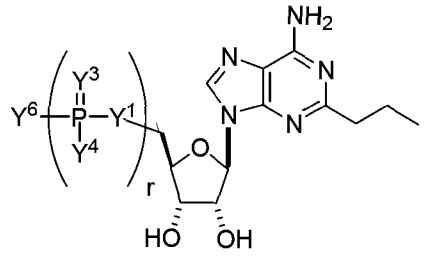
(BB- 170),



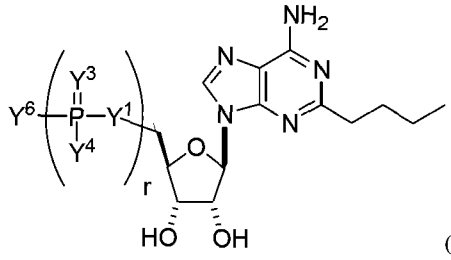
(BB- 171),



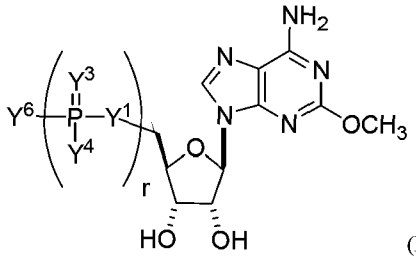
(BB- 172),



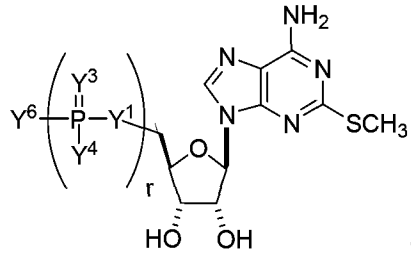
(BB- 173),



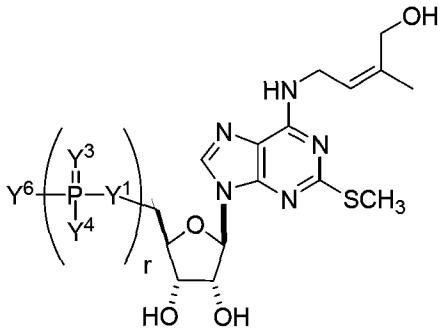
(BB- 174),



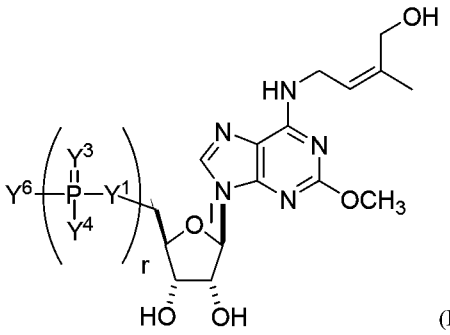
(BB- 175),



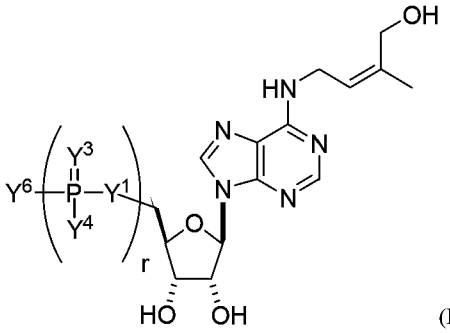
(BB- 176),



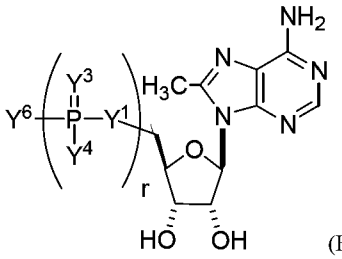
(BB- 177),



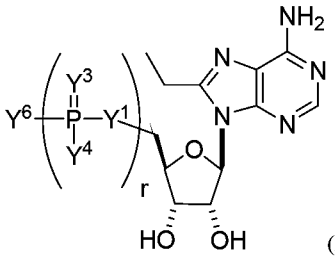
(BB- 178),



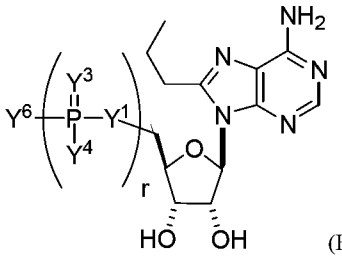
(BB- 179),



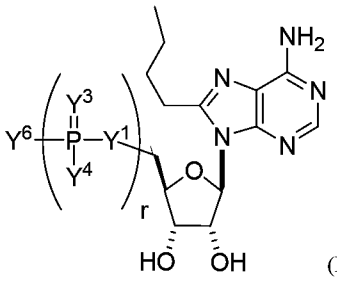
(BB- 180),



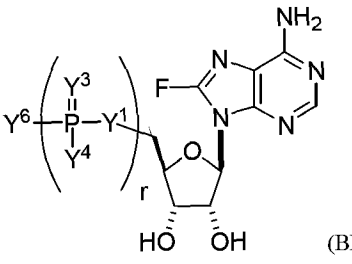
(BB- 181),

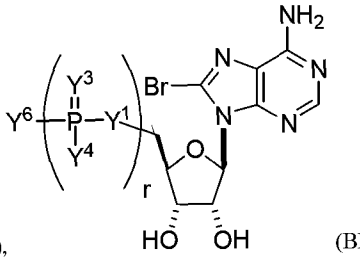


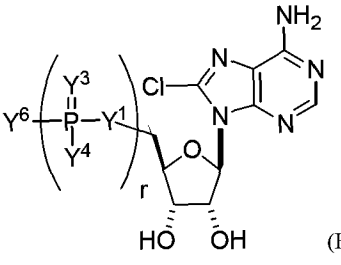
(BB- 182),

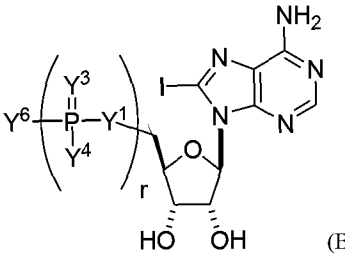


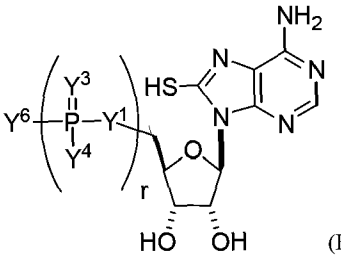
(BB- 183),

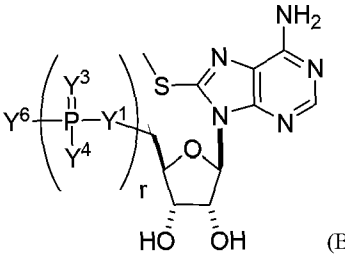

(BB- 184),

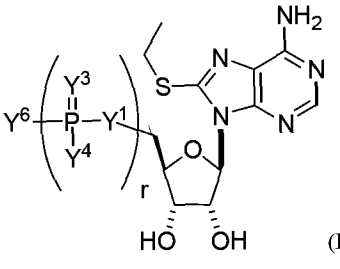

(BB- 185),

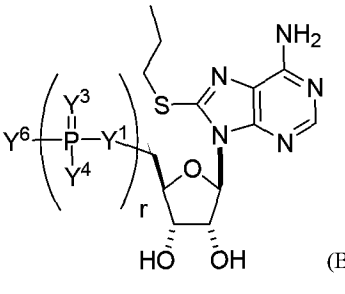

(BB- 186),

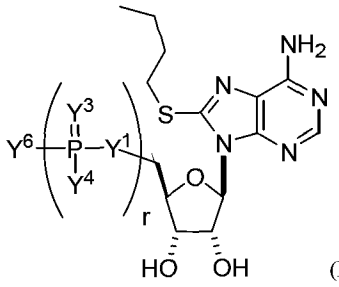

(BB- 187),


(BB- 188),

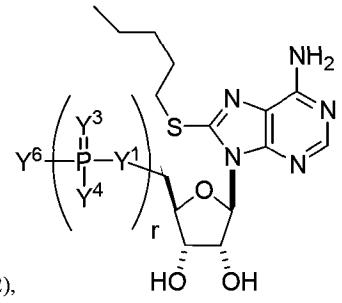

(BB- 189),


(BB- 190),

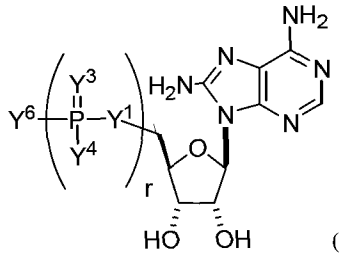

(BB- 191),



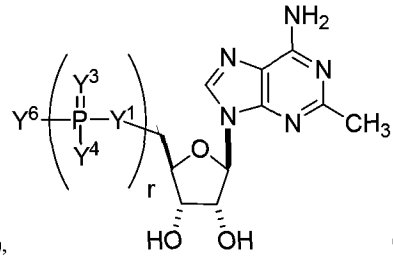
(BB- 192),



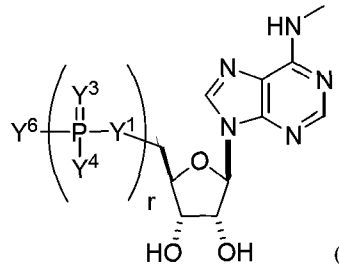
(BB- 193),



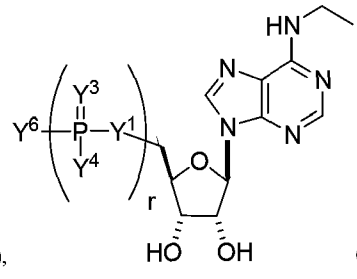
(BB- 194),



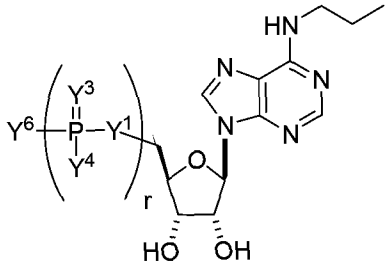
(BB- 195),



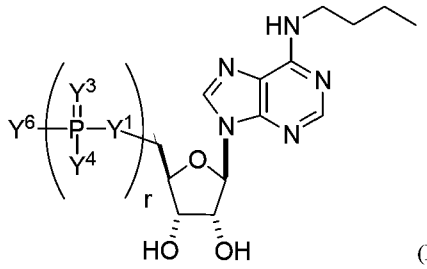
(BB- 196),



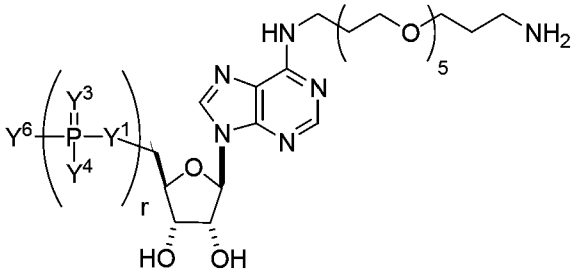
(BB- 197),



(BB- 198),

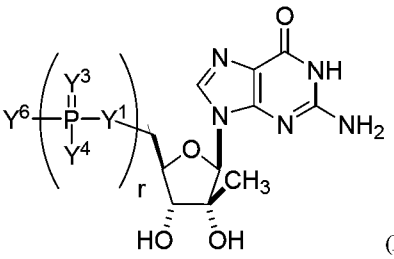


(BB- 199), and

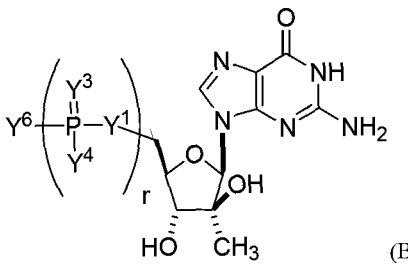


(BB- 200) or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $Y^1$, $Y^3$, $Y^4$, $Y^6$, and r are as described herein (e.g., each r is, independently, an integer from 0 to 5, such as from 0 to 3, from 1 to 3, or from 1 to 5)).

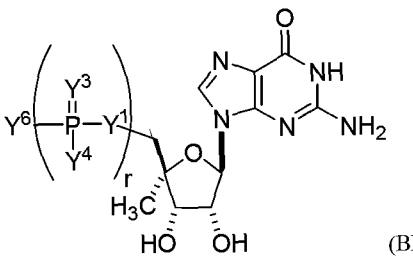
**[00183]** In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, is a modified guanosine (e.g., selected from the group consisting of:



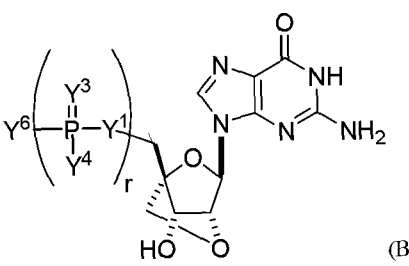
(BB- 201),



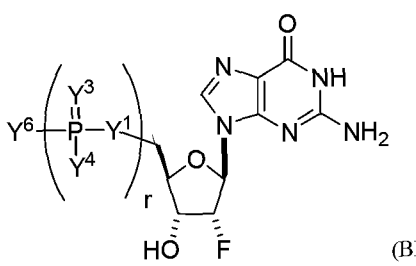
(BB- 202),



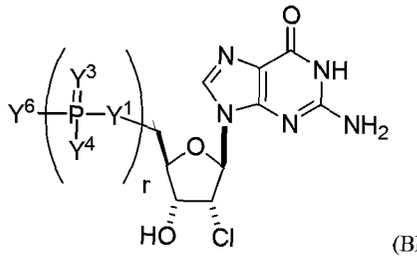
(BB- 203),



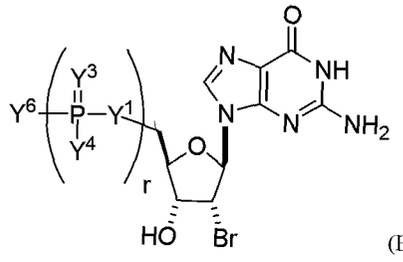
(BB- 204),



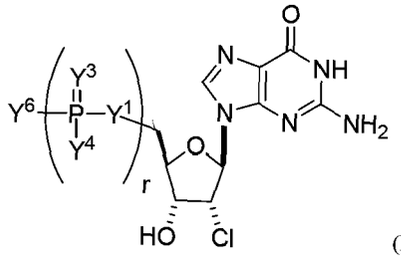
(BB- 205),



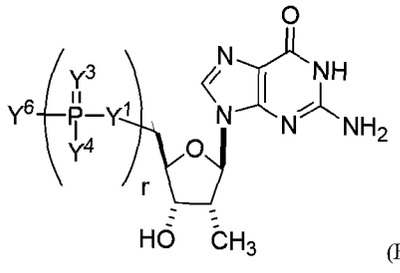
(BB- 206),



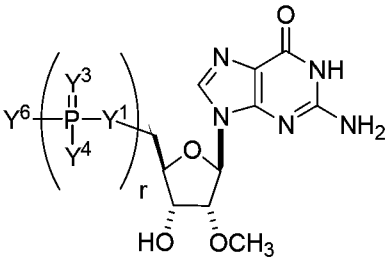
(BB- 207),



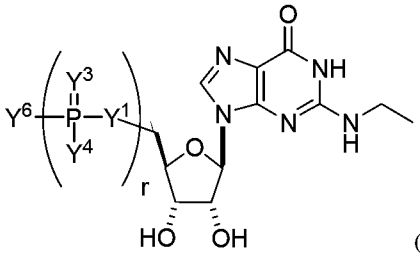
(BB- 208),



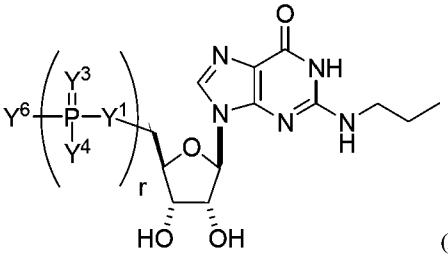
(BB- 209),



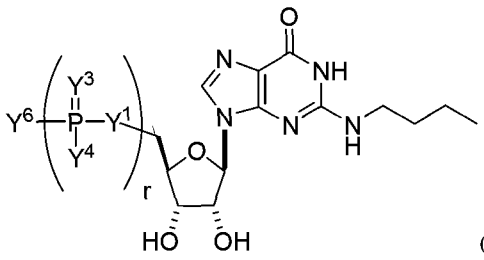
(BB- 210),



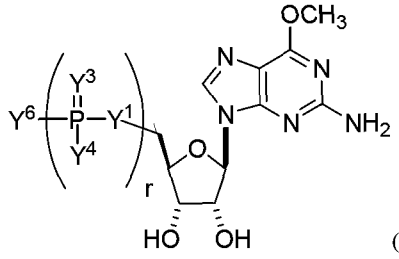
(BB- 211),



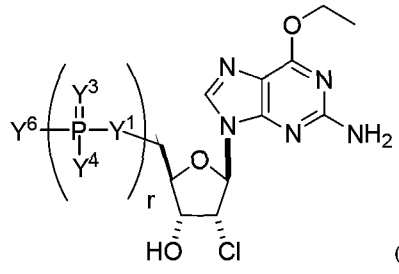
(BB- 212),



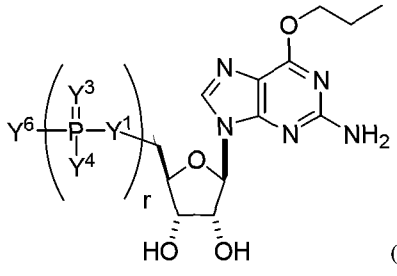
(BB- 213),



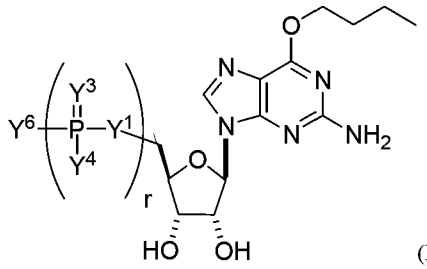
(BB- 214),



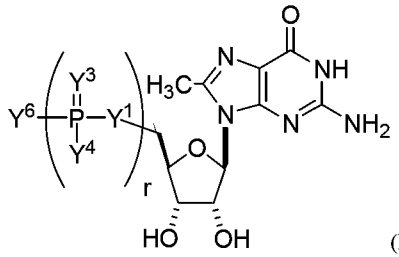
(BB- 215),



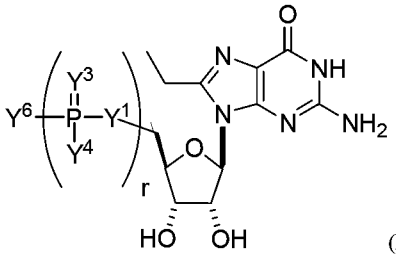
(BB- 216),



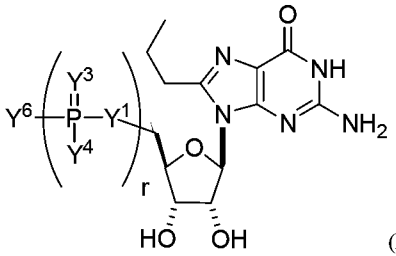
(BB- 217),



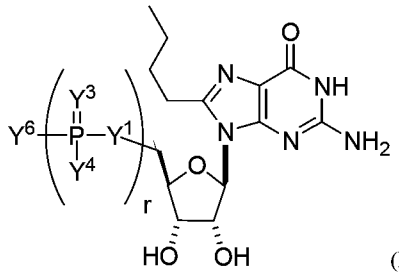
(BB- 218),



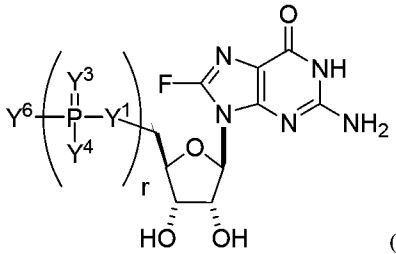
(BB- 219),



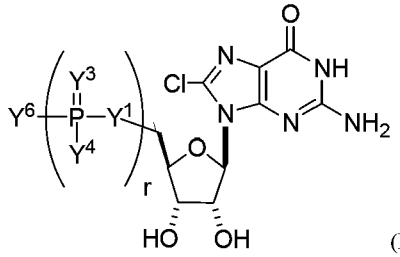
(BB- 220),



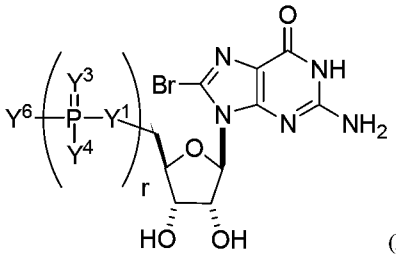
(BB- 221),



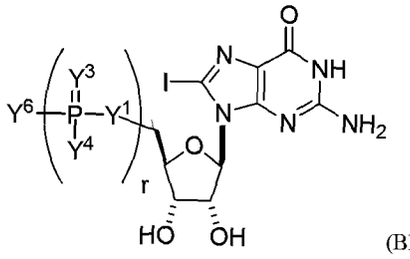
(BB- 222),



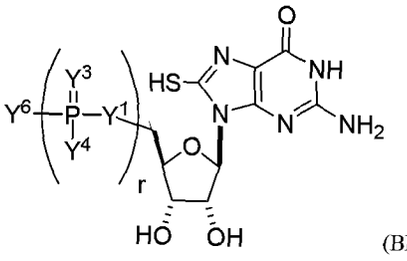
(BB- 223),



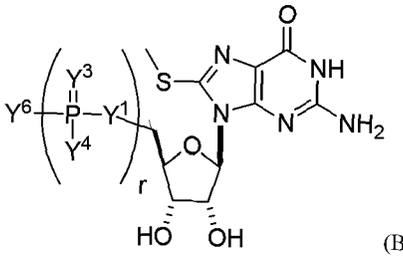
(BB- 224),



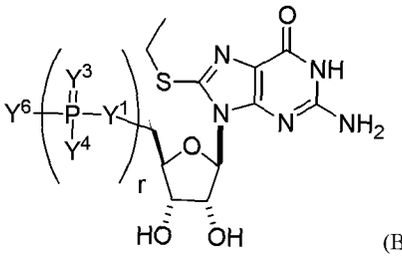
(BB- 225),



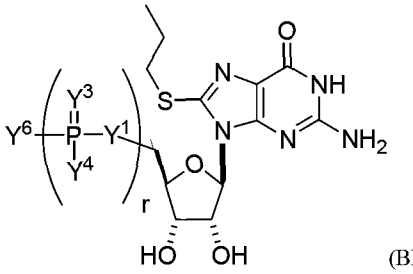
(BB- 226),



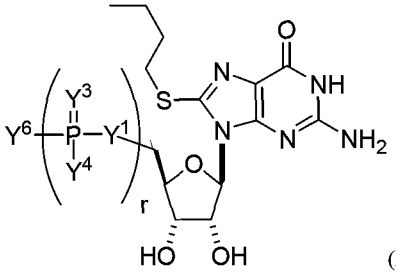
(BB- 227),



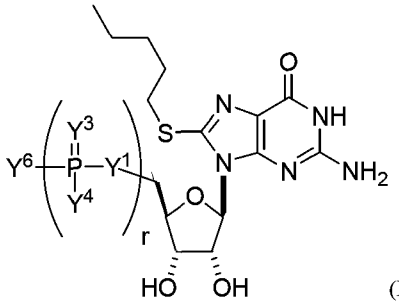
(BB- 228),



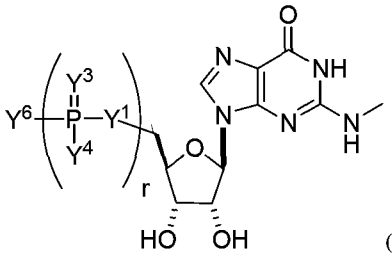
(BB- 229),



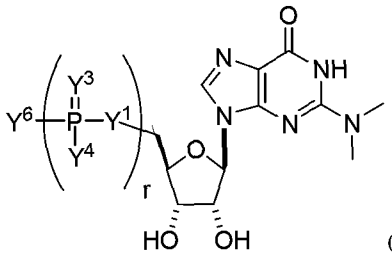
(BB- 230),



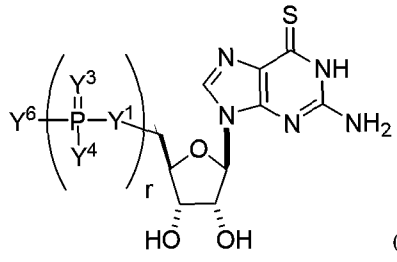
(BB- 231),



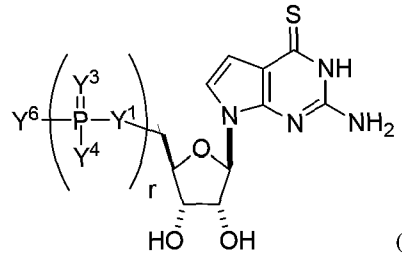
(BB- 232),



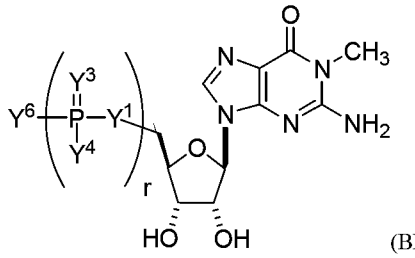
(BB- 233),



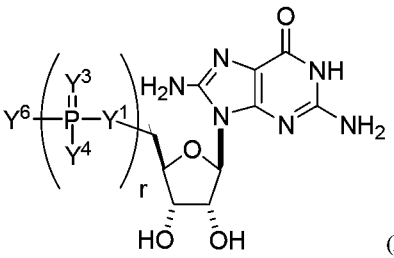
(BB- 234),



(BB- 235),

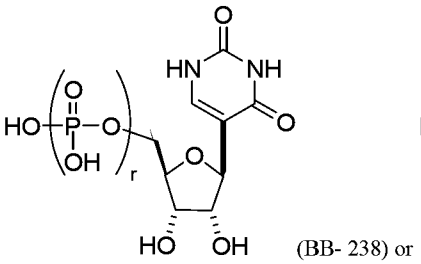


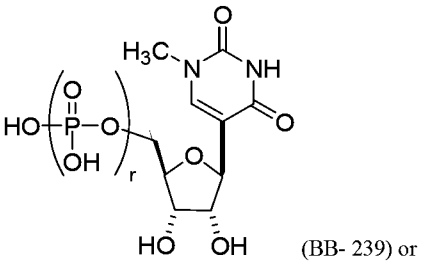
(BB- 236), and

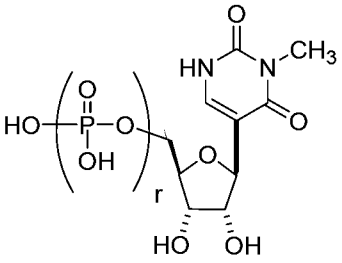


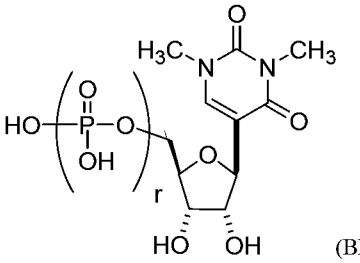
(BB- 237), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $Y^1$, $Y^3$, $Y^4$, $Y^6$, and r are as described herein (e.g., each r is, independently, an integer from 0 to 5, such as from 0 to 3, from 1 to 3, or from 1 to 5)).

**[00184]** In some embodiments, the major groove chemical modification can include replacement of C group at C-5 of the ring (e.g., for a pyrimidine nucleoside, such as cytosine or uracil) with N (e.g., replacement of the >CH group at C-5 with $>NR^{N1}$ group, wherein $R^{N1}$ is H or optionally substituted alkyl). For example, the building block molecule, which may be incorporated into a polynucleotide can be:


(BB- 238) or

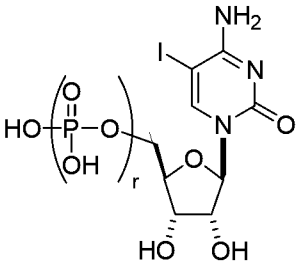

(BB- 239) or

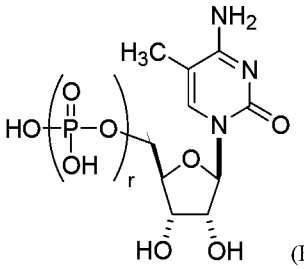

(BB- 240) or

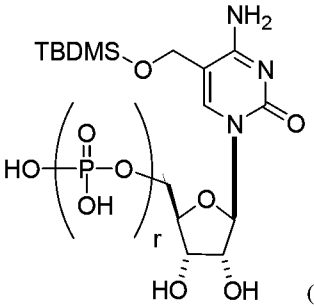

(BB- 241),

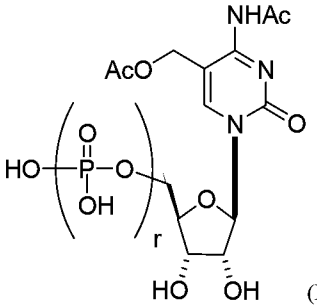
or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5).

**[00185]** In another embodiment, the major groove chemical modification can include replacement of the hydrogen at C-5 of cytosine with halo (e.g., Br, Cl, F, or I) or optionally substituted alkyl (e.g., methyl). For example, the building block molecule, which may be incorporated into a polynucleotide can be:

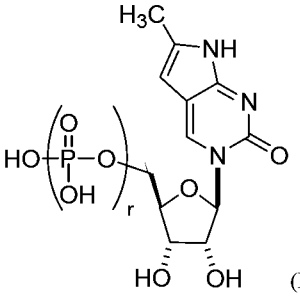

(BB- 242) or


(BB- 243) or


(BB- 244) or


(BB- 245), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5).

**[00186]** In yet a further embodiment, the major groove chemical modification can include a fused ring that is formed by the $NH_2$ at the C-4 position and the carbon atom at the C-5 position. For example, the building block molecule, which may be incorporated into a polynucleotide can be:


(BB- 246), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5).

*Modifications on the Sugar*

**[00187]** The modified nucleosides and nucleotides (e.g., building block molecules), which may be incorporated into a polynucleotide (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 substituents. Exemplary substitutions at the 2′-position include, but are not limited to, H, halo, optionally substituted $C_{1-6}$ alkyl; optionally substituted $C_{1-6}$ alkoxy; optionally substituted $C_{6-10}$ aryloxy; optionally substituted $C_{3-8}$ cycloalkyl; optionally substituted $C_{3-8}$ cycloalkoxy; optionally substituted $C_{6-10}$ aryloxy; optionally substituted $C_{6-10}$ aryl-$C_{1-6}$ alkoxy, optionally substituted $C_{1-12}$ (heterocyclyl)oxy; a sugar (e.g., ribose, pentose, or any described herein); a polyethyleneglycol (PEG), $-O(CH_2CH_2O)_nCH_2CH_2OR$, where R is H or optionally substituted alkyl, and n is an integer from 0 to 20 (e.g., from 0 to 4, from 0 to 8, from 0 to 10, from 0 to 16, from 1 to 4, from 1 to 8, from 1 to 10, from 1 to 16, from 1 to 20, from 2 to 4, from 2 to 8, from 2 to 10, from 2 to 16, from 2 to 20, from 4 to 8, from 4 to 10, from 4 to 16, and from 4 to 20); "locked" nucleic acids (LNA) in which the 2′-hydroxyl is connected by a $C_{1-6}$ alkylene or $C_{1-6}$ heteroalkylene bridge to the 4'-carbon of the same ribose sugar, where exemplary bridges included methylene, propylene, ether, or amino bridges; aminoalkyl, as defined herein; aminoalkoxy, as defined herein; amino as defined herein; and amino acid, as defined herein

**[00188]** Generally, RNA includes the sugar group ribose, which is a 5-membered ring having an oxygen. Exemplary, non-limiting modified nucleotides include replacement of the oxygen in ribose (e.g., with S, Se, or alkylene, such as methylene or ethylene); addition of a double bond (e.g., to replace ribose with cyclopentenyl or cyclohexenyl); ring contraction of ribose (e.g., to form a 4-membered ring of cyclobutane or oxetane); ring expansion of ribose (e.g., to form a 6- or 7-membered ring having an additional carbon or heteroatom, such as for anhydrohexitol, altritol, mannitol, cyclohexanyl, cyclohexenyl, and morpholino that also has a phosphoramidate backbone); multicyclic forms (e.g., tricyclo; and "unlocked" forms, such as glycol nucleic acid (GNA) (e.g., R-GNA or S-GNA, where ribose is replaced by glycol units attached to phosphodiester bonds), threose nucleic acid (TNA, where ribose is replace with α-L-threofuranosyl-(3′→2′)) , and peptide nucleic acid (PNA, where 2-amino-ethyl-glycine linkages replace the ribose and phosphodiester backbone). 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 polynucleotide molecule can include nucleotides containing, e.g., arabinose, as the sugar.

*Modifications on the Nucleobase*

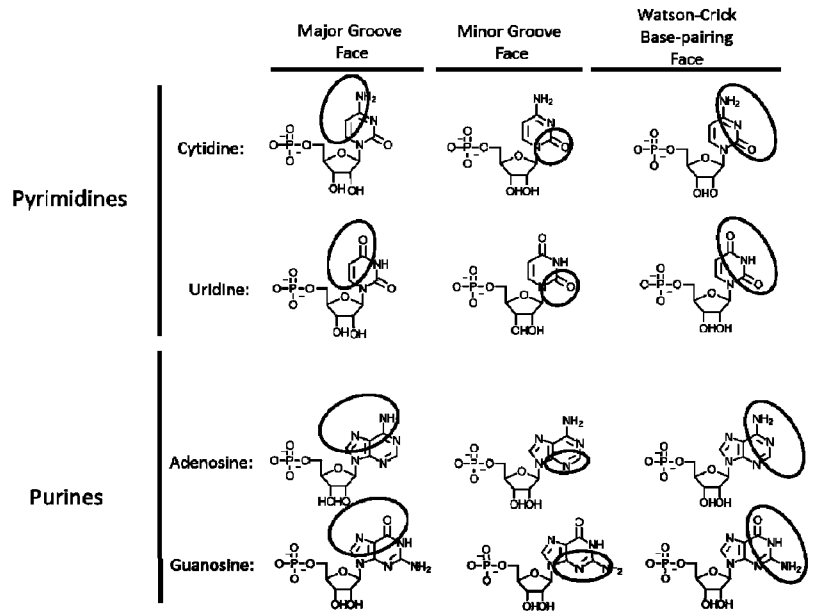
**[00189]** The present disclosure provides for modified nucleosides and nucleotides. As described herein "nucleoside" is defined as a compound containing a sugar molecule (e.g., a pentose or ribose) or derivative thereof in combination with an organic base (e.g., a purine or pyrimidine) or a derivative thereof (also referred to herein as "nucleobase"). As described herein, "nucleotide" is defined as a nucleoside including a phosphate group. In some embodiments, the nucleosides and nucleotides described herein are generally chemically modified on the major groove face. Exemplary non-limiting modifications include an amino group, a thiol group, an alkyl group, a halo group, or any described herein. The modified nucleotides may by synthesized by any useful method, as described herein (e.g., chemically, enzymatically, or recombinantly to include one or more modified or non-natural nucleosides).

**[00190]** The modified nucleotide base pairing encompasses not only the standard adenosine-thymine, adenosine-uracil, or guanosine-cytosine base pairs, but also base pairs formed between nucleotides and/or modified nucleotides comprising non-standard or modified bases, wherein the arrangement of hydrogen bond donors and hydrogen bond acceptors permits hydrogen bonding between a non-standard base and a standard base or between two complementary non-standard base structures. One example of such non-standard base pairing is the base pairing between the modified nucleotide inosine and adenine, cytosine or uracil.

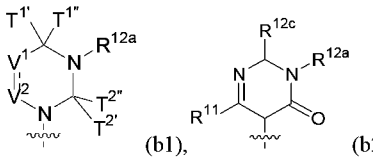
**[00191]** The modified nucleosides and nucleotides can include a modified nucleobase. Examples of nucleobases found in RNA include, but are not limited to, adenine, guanine, cytosine, and uracil. Examples of nucleobase found in DNA include, but are not limited to, adenine, guanine, cytosine, and thymine. These nucleobases can be modified or wholly replaced to provide polynucleotide molecules having enhanced properties, e.g., resistance to nucleases, stability, and these properties may manifest through disruption of the binding of a major groove binding partner. For example, the nucleosides and nucleotides described 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.

**[00192]** Table 1 below identifies the chemical faces of each canonical nucleotide. Circles identify the atoms comprising the respective chemical regions.

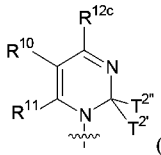
**Table 1**

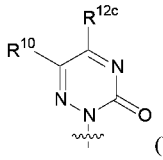


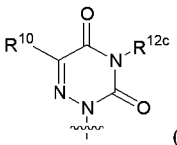

**[00193]** In some embodiments, B is a modified uracil. Exemplary modified uracils include those having Formula (b1)-(b5):


(b1),


(b2),


(b3),


(b4), or


(b5), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00194]** ⩬ is a single or double bond;

**[00195]** each of $T^{1'}$, $T^{1''}$, $T^{2'}$, and $T^{2''}$ is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of $T^{1'}$ and $T^{1''}$ or the combination of $T^{2'}$ and $T^{2''}$ join together (e.g., as in $T^2$) to form O (oxo), S (thio), or Se (seleno);

**[00196]** each of $V^1$ and $V^2$ is, independently, O, S, $N(R^{Vb})_{nv}$, or $C(R^{Vb})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vb}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl), optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted acylaminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl), optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, or optionally substituted alkoxycarbonylalkoxy (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl);

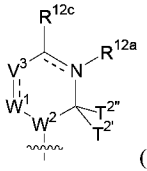
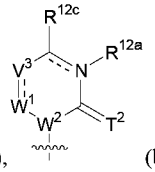
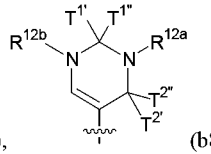
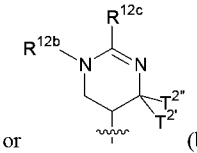
**[00197]** $R^{10}$ is H, halo, optionally substituted amino acid, hydroxy, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aminoalkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl;

**[00198]** $R^{11}$ is H or optionally substituted alkyl;

**[00199]** $R^{12a}$ is H, optionally substituted alkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl, optionally substituted carboxyalkyl (e.g., optionally substituted with hydroxy), optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, or optionally substituted carbamoylalkyl; and

**[00200]** $R^{12c}$ is H, halo, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted amino, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl.

**[00201]** Other exemplary modified uracils include those having Formula (b6)-(b9):


(b6),

(b7),

(b8), or

(b9), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00202]** ⁄ is a single or double bond;

**[00203]** each of $T^{1'}$, $T^{1''}$, $T^{2'}$, and $T^{2''}$ is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of $T^{1'}$ and $T^{1''}$ join together (e.g,. as in $T^1$) or the combination of $T^{2'}$ and $T^{2''}$ join together (e.g., as in $T^2$) to form O (oxo), S (thio), or Se (seleno), or each $T^1$ and $T^2$ is, independently, O (oxo), S (thio), or Se (seleno);

**[00204]** each of $W^1$ and $W^2$ is, independently, $N(R^{Wa})_{nw}$ or $C(R^{Wa})_{nw}$, wherein nw is an integer from 0 to 2 and each $R^{Wa}$ is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy;

**[00205]** each $V^3$ is, independently, O, S, $N(R^{Va})_{nv}$, or $C(R^{Va})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Va}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted alkoxy, optionally substituted alkenyloxy, or optionally substituted alkynyloxy, optionally substituted aminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl, or sulfoalkyl), optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted acylaminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl), optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylacyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkyl (e.g., optionally substituted with hydroxy and/or an *O*-protecting group), optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, or optionally substituted carbamoylalkyl (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl), and wherein $R^{Va}$ and $R^{12c}$ taken together with the carbon

atoms to which they are attached can form optionally substituted cycloalkyl, optionally substituted aryl, or optionally substituted heterocyclyl (e.g., a 5- or 6-membered ring);

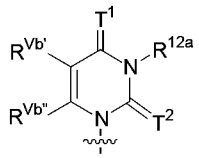
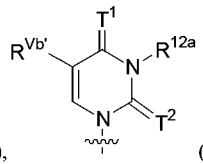
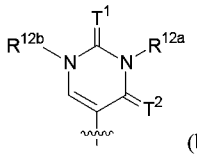
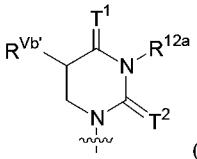
**[00206]** $R^{12a}$ is H, optionally substituted alkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted carboxyalkyl (e.g., optionally substituted with hydroxy and/or an *O*-protecting group), optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, optionally substituted carbamoylalkyl, or absent;

**[00207]** $R^{12b}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkaryl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted amino acid, optionally substituted alkoxycarbonylacyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkyl (e.g., optionally substituted with hydroxy and/or an *O*-protecting group), optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, or optionally substituted carbamoylalkyl,

**[00208]** wherein the combination of $R^{12b}$ and $T^{1'}$ or the combination of $R^{12b}$ and $R^{12c}$ can join together to form optionally substituted heterocyclyl; and

**[00209]** $R^{12c}$ is H, halo, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted amino, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl.

**[00210]** Further exemplary modified uracils include those having Formula (b28)-(b31):


(b28),

(b29),

(b30), or

(b31), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00211]** each of $T^1$ and $T^2$ is, independently, O (oxo), S (thio), or Se (seleno);

**[00212]** each $R^{Vb'}$ and $R^{Vb''}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl, or sulfoalkyl), optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted acylaminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl), optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylacyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkyl (e.g., optionally substituted with hydroxy and/or an *O*-protecting group), optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, or optionally substituted carbamoylalkyl (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl) (e.g., $R^{Vb'}$ is optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted aminoalkyl, e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl, or sulfoalkyl);

**[00213]** $R^{12a}$ is H, optionally substituted alkyl, optionally substituted carboxyaminoalkyl, optionally substituted aminoalkyl (e.g., e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl, or sulfoalkyl), optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl; and

**[00214]** $R^{12b}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl (e.g., e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl, or sulfoalkyl), optionally substituted alkoxycarbonylacyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl.

**[00215]** In particular embodiments, $T^1$ is O (oxo), and $T^2$ is S (thio) or Se (seleno). In other embodiments, $T^1$ is S (thio), and $T^2$ is O (oxo) or Se (seleno). In some embodiments, $R^{Vb'}$ is H, optionally substituted alkyl, or optionally substituted alkoxy.

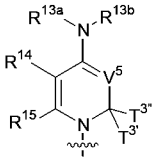
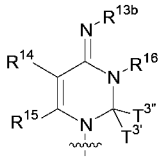
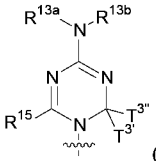
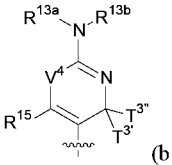
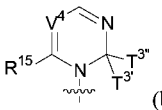
**[00216]** In other embodiments, each $R^{12a}$ and $R^{12b}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted hydroxyalkyl. In particular embodiments, $R^{12a}$ is H. In other embodiments, both $R^{12a}$ and $R^{12b}$ are H.

**[00217]** In some embodiments, each $R^{Vb'}$ of $R^{12b}$ is, independently, optionally substituted aminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl, or sulfoalkyl), optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or optionally substituted acylaminoalkyl (e.g., substituted with an N-protecting group, such as any described herein, e.g., trifluoroacetyl). In some embodiments, the amino and/or alkyl of the optionally substituted aminoalkyl is substituted with one or more of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted sulfoalkyl, optionally substituted carboxy (e.g., substituted with an *O*-protecting group), optionally substituted hydroxy (e.g., substituted with an *O*-protecting group), optionally substituted carboxyalkyl (e.g., substituted with an *O*-protecting group), optionally substituted alkoxycarbonylalkyl (e.g., substituted with an *O*-protecting group), or *N*-protecting group. In some embodiments, optionally substituted aminoalkyl is substituted with an optionally substituted sulfoalkyl or optionally substituted alkenyl. In particular embodiments, $R^{12a}$ and $R^{Vb''}$ are both H. In particular embodiments, $T^1$ is O (oxo), and $T^2$ is S (thio) or Se (seleno).

**[00218]** In some embodiments, $R^{Vb'}$ is optionally substituted alkoxycarbonylalkyl or optionally substituted carbamoylalkyl.

**[00219]** In particular embodiments, the optional substituent for $R^{12a}$, $R^{12b}$, $R^{12c}$, or $R^{Va}$ is a polyethylene glycol group (e.g., $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl); or an amino-polyethylene glycol group (e.g., $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl).

**[00220]** In some embodiments, B is a modified cytosine. Exemplary modified cytosines include compounds of Formula (b10)-(b14):


(b10),

(b11),

(b12),

(b13), or

(b14), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00221]** each of $T^{3'}$ and $T^{3''}$ is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of $T^{3'}$ and $T^{3''}$ join together (e.g., as in $T^3$) to form O (oxo), S (thio), or Se (seleno);

**[00222]** each $V^4$ is, independently, O, S, $N(R^{Vc})_{nv}$, or $C(R^{Vc})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vc}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, or optionally substituted alkynyloxy (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl), wherein the combination of $R^{13b}$ and $R^{Vc}$ can be taken together to form optionally substituted heterocyclyl;

**[00223]** each $V^5$ is, independently, $N(R^{Vd})_{nv}$, or $C(R^{Vd})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vd}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted heterocyclyl, optionally substituted

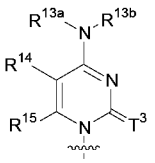
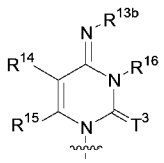
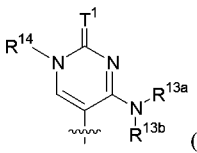
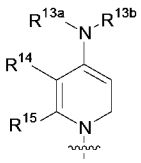
alkheterocyclyl, or optionally substituted alkynyloxy (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl) (e.g., $V^5$ is –CH or N);

**[00224]** each of $R^{13a}$ and $R^{13b}$ is, independently, H, optionally substituted acyl, optionally substituted acyloxyalkyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the combination of $R^{13b}$ and $R^{14}$ can be taken together to form optionally substituted heterocyclyl;

**[00225]** each $R^{14}$ is, independently, H, halo, hydroxy, thiol, optionally substituted acyl, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl (e.g., substituted with an *O*-protecting group), optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted acyloxyalkyl, optionally substituted amino (e.g., -NHR, wherein R is H, alkyl, aryl, or phosphoryl), azido, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl; and

**[00226]** each of $R^{15}$ and $R^{16}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl.

**[00227]** Further exemplary modified cytosines include those having Formula (b32)-(b35):


(b32),

(b33),

(b34), or


(b35), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00228]** each of $T^1$ and $T^3$ is, independently, O (oxo), S (thio), or Se (seleno);

**[00229]** each of $R^{13a}$ and $R^{13b}$ is, independently, H, optionally substituted acyl, optionally substituted acyloxyalkyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the combination of $R^{13b}$ and $R^{14}$ can be taken together to form optionally substituted heterocyclyl;

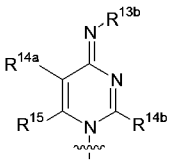
**[00230]** each $R^{14}$ is, independently, H, halo, hydroxy, thiol, optionally substituted acyl, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl (e.g., substituted with an *O*-protecting group), optionally substituted hydroxyalkenyl, optionally

substituted hydroxyalkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted acyloxyalkyl, optionally substituted amino (e.g., -NHR, wherein R is H, alkyl, aryl, or phosphoryl), azido, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted aminoalkyl (e.g., hydroxyalkyl, alkyl, alkenyl, or alkynyl), optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl; and

**[00231]** each of $R^{15}$ and $R^{16}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl (e.g., $R^{15}$ is H, and $R^{16}$ is H or optionally substituted alkyl).

**[00232]** In some embodiments, $R^{15}$ is H, and $R^{16}$ is H or optionally substituted alkyl. In particular embodiments, $R^{14}$ is H, acyl, or hydroxyalkyl. In some embodiments, $R^{14}$ is halo. In some embodiments, both $R^{14}$ and $R^{15}$ are H. In some embodiments, both $R^{15}$ and $R^{16}$ are H. In some embodiments, each of $R^{14}$ and $R^{15}$ and $R^{16}$ is H. In further embodiments, each of $R^{13a}$ and $R^{13b}$ is independently, H or optionally substituted alkyl.

**[00233]** Further non-limiting examples of modified cytosines include compounds of Formula (b36):



(b36) or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00234]** each $R^{13b}$ is, independently, H, optionally substituted acyl, optionally substituted acyloxyalkyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the combination of $R^{13b}$ and $R^{14b}$ can be taken together to form optionally substituted heterocyclyl;

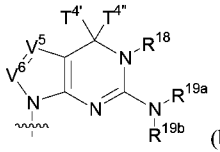
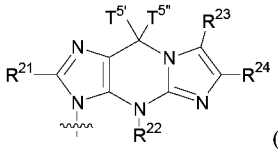
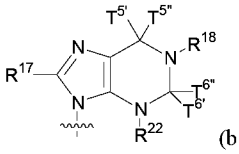
**[00235]** each $R^{14a}$ and $R^{14b}$ is, independently, H, halo, hydroxy, thiol, optionally substituted acyl, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl (e.g., substituted with an *O*-protecting group), optionally substituted hydroxyalkenyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted acyloxyalkyl, optionally substituted amino (e.g., -NHR, wherein R is H, alkyl, aryl, phosphoryl, optionally

substituted aminoalkyl, or optionally substituted carboxyaminoalkyl), azido, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl; and

**[00236]** each of $R^{15}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl.

**[00237]** In particular embodiments, $R^{14b}$ is an optionally substituted amino acid (e.g., optionally substituted lysine). In some embodiments, $R^{14a}$ is H.

**[00238]** In some embodiments, B is a modified guanine. Exemplary modified guanines include compounds of Formula (b15)-(b17):


(b15),

(b16), or

(b17),

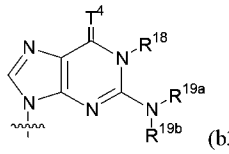
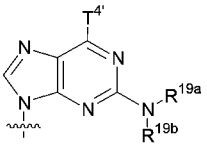
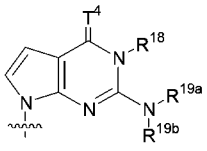
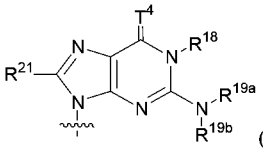
or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00239]** Each of $T^{4'}$, $T^{4''}$, $T^{5'}$, $T^{5''}$, $T^{6'}$, and $T^{6''}$ is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy, and wherein the combination of $T^{4'}$ and $T^{4''}$ (e.g., as in $T^4$) or the combination of $T^{5'}$ and $T^{5''}$ (e.g., as in $T^5$) or the combination of $T^{6'}$ and $T^{6''}$ join together (e.g., as in $T^6$) form O (oxo), S (thio), or Se (seleno);

**[00240]** each of $V^5$ and $V^6$ is, independently, O, S, $N(R^{Vd})_{nv}$, or $C(R^{Vd})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vd}$ is, independently, H, halo, thiol, optionally substituted amino acid, cyano, amidine, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl), optionally substituted thioalkoxy, or optionally substituted amino; and

**[00241]** each of $R^{17}$, $R^{18}$, $R^{19a}$, $R^{19b}$, $R^{21}$, $R^{22}$, $R^{23}$, and $R^{24}$ is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, optionally substituted amino, or optionally substituted amino acid.

**[00242]** Exemplary modified guanosines include compounds of Formula (b37)-(b40):

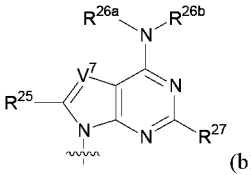
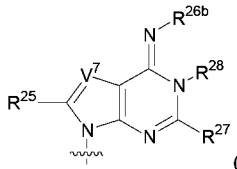
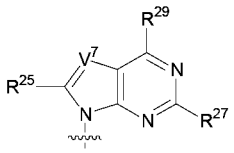

(b37),

(b38),

(b39), or

(b40), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00243]** each of $T^{4'}$ is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy, and each $T^4$ is, independently, O (oxo), S (thio), or Se (seleno);

**[00244]** each of $R^{18}$, $R^{19a}$, $R^{19b}$, and $R^{21}$ is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, optionally substituted amino, or optionally substituted amino acid.

**[00245]** In some embodiments, $R^{18}$ is H or optionally substituted alkyl. In further embodiments, $T^4$ is oxo. In some embodiments, each of $R^{19a}$ and $R^{19b}$ is, independently, H or optionally substituted alkyl.

**[00246]** In some embodiments, B is a modified adenine. Exemplary modified adenines include compounds of Formula (b18)-(b20):


(b18),

(b19), or

(b20),
or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00247]** each $V^7$ is, independently, O, S, $N(R^{Ve})_{nv}$, or $C(R^{Ve})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Ve}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, or optionally substituted alkynyloxy (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl);

**[00248]** each $R^{25}$ is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, or optionally substituted amino;

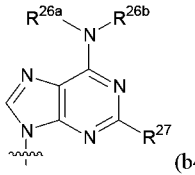
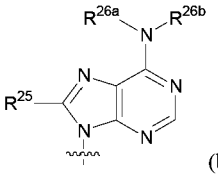
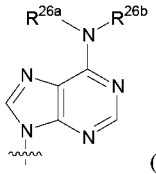
**[00249]** each of $R^{26a}$ and $R^{26b}$ is, independently, H, optionally substituted acyl, optionally substituted amino acid, optionally substituted carbamoylalkyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkoxy, or polyethylene glycol group (e.g., $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl); or an amino-polyethylene glycol group (e.g., $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl);

**[00250]** each $R^{27}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted thioalkoxy, or optionally substituted amino;

**[00251]** each $R^{28}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl; and

**[00252]** each $R^{29}$ is, independently, H, optionally substituted acyl, optionally substituted amino acid, optionally substituted carbamoylalkyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted alkoxy, or optionally substituted amino.

**[00253]** Exemplary modified adenines include compounds of Formula (b41)-(b43):


(b41),

(b42), or

(b43), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

**[00254]** each $R^{25}$ is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, or optionally substituted amino;

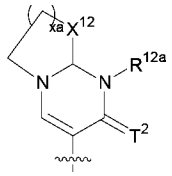
**[00255]** each of $R^{26a}$ and $R^{26b}$ is, independently, H, optionally substituted acyl, optionally substituted amino acid, optionally substituted carbamoylalkyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkoxy, or polyethylene glycol group (e.g., $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl); or an amino-polyethylene glycol group (e.g., $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl); and

**[00256]** each $R^{27}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted thioalkoxy, or optionally substituted amino.

**[00257]** In some embodiments, $R^{26a}$ is H, and $R^{26b}$ is optionally substituted alkyl. In some embodiments, each of $R^{26a}$ and $R^{26b}$ is, independently, optionally substituted alkyl. In particular embodiments, $R^{27}$ is optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy. In other embodiments, $R^{25}$ is optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy.

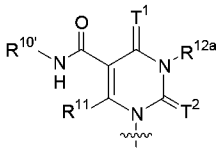
**[00258]** In particular embodiments, the optional substituent for $R^{26a}$, $R^{26b}$, or $R^{29}$ is a polyethylene glycol group (e.g., $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl); or an amino-polyethylene glycol group (e.g., $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl).

**[00259]** In some embodiments, B may have Formula (b21):



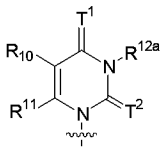
(b21), wherein $X^{12}$ is, independently, O, S, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene, xa is an integer from 0 to 3, and $R^{12a}$ and $T^2$ are as described herein.

**[00260]** In some embodiments, B may have Formula (b22):



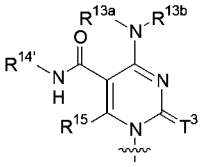
(b22), wherein $R^{10'}$ is, independently, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl, and $R^{11}$, $R^{12a}$, $T^1$, and $T^2$ are as described herein.

**[00261]** In some embodiments, B may have Formula (b23):



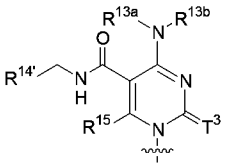
(b23), wherein $R^{10}$ is optionally substituted heterocyclyl (e.g., optionally substituted furyl, optionally substituted thienyl, or optionally substituted pyrrolyl), optionally substituted aryl (e.g., optionally substituted phenyl or optionally substituted naphthyl), or any substituent described herein (e.g., for $R^{10}$) ;and wherein $R^{11}$ (e.g., H or any substituent described herein), $R^{12a}$ (e.g., H or any substituent described herein), $T^1$ (e.g., oxo or any substituent described herein), and $T^2$ (e.g., oxo or any substituent described herein) are as described herein.

**[00262]** In some embodiments, B may have Formula (b24):



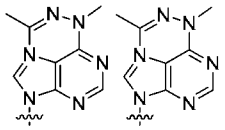
(b24), wherein $R^{14'}$ is, independently, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted alkaryl, optionally substituted alkheterocyclyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl, and $R^{13a}$, $R^{13b}$, $R^{15}$, and $T^3$ are as described herein.

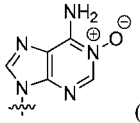
**[00263]** In some embodiments, B may have Formula (b25):



(b25), wherein $R^{14'}$ is optionally substituted heterocyclyl (e.g., optionally substituted furyl, optionally substituted thienyl, or optionally substituted pyrrolyl), optionally substituted aryl (e.g., optionally substituted phenyl or optionally substituted naphthyl), or any substituent described herein (e.g., for $R^{14}$ or $R^{14'}$); and wherein $R^{13a}$ (e.g., H or any substituent described herein), $R^{13b}$ (e.g., H or any substituent described herein), $R^{15}$ (e.g., H or any substituent described herein), and $T^3$ (e.g., oxo or any substituent described herein) are as described herein.

**[00264]** In some embodiments, B is a nucleobase selected from the group consisting of cytosine, guanine, adenine, and uracil. In some embodiments, B may be:



(b26) or  (b27).

**[00265]** In some embodiments, the modified nucleobase is a modified uracil. Exemplary nucleobases and nucleosides having a modified uracil include pseudouridine (ψ), pyridin-4-one ribonucleoside, 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine ($s^2U$), 4-thio-

uridine ($s^4U$), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine ($ho^5U$), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridineor 5-bromo-uridine), 3-methyl-uridine ($m^3U$), 5-methoxy-uridine ($mo^5U$), uridine 5-oxyacetic acid ($cmo^5U$), uridine 5-oxyacetic acid methyl ester ($mcmo^5U$), 5-carboxymethyl-uridine ($cm^5U$), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine ($chm^5U$), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm^5U$), 5-methoxycarbonylmethyl-uridine ($mcm^5U$), 5-methoxycarbonylmethyl-2-thio-uridine ($mcm^5s^2U$), 5-aminomethyl-2-thio-uridine ($nm^5s^2U$), 5-methylaminomethyl-uridine ($mnm^5U$), 5-methylaminomethyl-2-thio-uridine ($mnm^5s^2U$), 5-methylaminomethyl-2-seleno-uridine ($mnm^5se^2U$), 5-carbamoylmethyl-uridine ($ncm^5U$), 5-carboxymethylaminomethyl-uridine ($cmnm^5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s^2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine ($\tau m^5U$), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine($\tau m^5s^2U$), 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-uridine ($m^5U$, i.e., having the nucleobase deoxythymine), 1-methyl-pseudouridine ($m^1\psi$), 5-methyl-2-thio-uridine ($m^5s^2U$), 1-methyl-4-thio-pseudouridine ($m^1s^4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine ($m^5D$), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine ($acp^3U$), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine ($acp^3\ \psi$), 5-(isopentenylaminomethyl)uridine ($inm^5U$), 5-(isopentenylaminomethyl)-2-thio-uridine ($inm^5s^2U$), α-thio-uridine, 2′-O-methyl-uridine (Um), 5,2′-O-dimethyl-uridine ($m^5Um$), 2′-O-methyl-pseudouridine ($\psi m$), 2-thio-2′-O-methyl-uridine ($s^2Um$), 5-methoxycarbonylmethyl-2′-O-methyl-uridine ($mcm^5Um$), 5-carbamoylmethyl-2′-O-methyl-uridine ($ncm^5Um$), 5-carboxymethylaminomethyl-2′-O-methyl-uridine ($cmnm^5Um$), 3,2′-O-dimethyl-uridine ($m^3Um$), and 5-(isopentenylaminomethyl)-2′-O-methyl-uridine ($inm^5Um$), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, and 5-[3-(1-E-propenylamino)uridine.

**[00266]** In some embodiments, the modified nucleobase is a modified cytosine. Exemplary nucleobases and nucleosides having a modified cytosine include 5-aza-cytidine, 6-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine ($m^3C$), N4-acetyl-cytidine ($ac^4C$), 5-formyl-cytidine ($f^5C$), N4-methyl-cytidine ($m^4C$), 5-methyl-cytidine ($m^5C$), 5-halo-cytidine (e.g., 5-iodo-cytidine), 5-

hydroxymethyl-cytidine ($hm^5C$), 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine ($s^2C$), 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, 4-methoxy-1-methyl-pseudoisocytidine, lysidine ($k_2C$), α-thio-cytidine, 2′-O-methyl-cytidine (Cm), 5,2′-O-dimethyl-cytidine ($m^5Cm$), N4-acetyl-2′-O-methyl-cytidine ($ac^4Cm$), N4,2′-O-dimethyl-cytidine ($m^4Cm$), 5-formyl-2′-O-methyl-cytidine ($f^5Cm$), N4,N4,2′-O-trimethyl-cytidine ($m^4{}_2Cm$), 1-thio-cytidine, 2'-F-ara-cytidine, 2'-F-cytidine, and 2'-OH-ara-cytidine.

**[00267]** In some embodiments, the modified nucleobase is a modified adenine. Exemplary nucleobases and nucleosides having a modified adenine include 2-amino-purine, 2, 6-diaminopurine, 2-amino-6-halo-purine (e.g., 2-amino-6-chloro-purine), 6-halo-purine (e.g., 6-chloro-purine), 2-amino-6-methyl-purine, 8-azido-adenosine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-amino-purine, 7-deaza-8-aza-2-amino-purine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyl-adenosine ($m^1A$), 2-methyl-adenine ($m^2A$), N6-methyl-adenosine ($m^6A$), 2-methylthio-N6-methyl-adenosine ($ms^2m^6A$), N6-isopentenyl-adenosine ($i^6A$), 2-methylthio-N6-isopentenyl-adenosine ($ms^2i^6A$), N6-(cis-hydroxyisopentenyl)adenosine ($io^6A$), 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine ($ms^2io^6A$), N6-glycinylcarbamoyl-adenosine ($g^6A$), N6-threonylcarbamoyl-adenosine ($t^6A$), N6-methyl-N6-threonylcarbamoyl-adenosine ($m^6t^6A$), 2-methylthio-N6-threonylcarbamoyl-adenosine ($ms^2g^6A$), N6,N6-dimethyl-adenosine ($m^6{}_2A$), N6-hydroxynorvalylcarbamoyl-adenosine ($hn^6A$), 2-methylthio-N6-hydroxynorvalylcarbamoyl-adenosine ($ms^2hn^6A$), N6-acetyl-adenosine ($ac^6A$), 7-methyl-adenine, 2-methylthio-adenine, 2-methoxy-adenine, α-thio-adenosine, 2′-O-methyl-adenosine (Am), N6,2′-O-dimethyl-adenosine ($m^6Am$), N6,N6,2′-O-trimethyl-adenosine ($m^6{}_2Am$), 1,2′-O-dimethyl-adenosine ($m^1Am$), 2′-O-ribosyladenosine (phosphate) (Ar(p)), 2-amino-N6-methyl-purine, 1-thio-adenosine, 8-azido-adenosine, 2'-F-ara-adenosine, 2'-F-adenosine, 2'-OH-ara-adenosine, and N6-(19-amino-pentaoxanonadecyl)-adenosine.

**[00268]** In some embodiments, the modified nucleobase is a modified guanine. Exemplary nucleobases and nucleosides having a modified guanine include inosine (I), 1-methyl-inosine ($m^1I$), wyosine (imG), methylwyosine (mimG), 4-demethyl-wyosine (imG-14), isowyosine (imG2), wybutosine (yW), peroxywybutosine ($o_2yW$), hydroxywybutosine (OHyW), undermodified

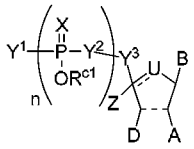
hydroxywybutosine (OHyW*), 7-deaza-guanosine, queuosine (Q), epoxyqueuosine (oQ), galactosyl-queuosine (galQ), mannosyl-queuosine (manQ), 7-cyano-7-deaza-guanosine ($preQ_0$), 7-aminomethyl-7-deaza-guanosine ($preQ_1$), archaeosine ($G^+$), 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine ($m^7G$), 6-thio-7-methyl-guanosine, 7-methyl-inosine, 6-methoxy-guanosine, 1-methyl-guanosine ($m^1G$), N2-methyl-guanosine ($m^2G$), N2,N2-dimethyl-guanosine ($m^2{}_2G$), N2,7-dimethyl-guanosine ($m^{2,7}G$), N2, N2,7-dimethyl-guanosine ($m^{2,2,7}G$), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, N2,N2-dimethyl-6-thio-guanosine, α-thio-guanosine, 2′-O-methyl-guanosine (Gm), N2-methyl-2′-O-methyl-guanosine ($m^2Gm$), N2,N2-dimethyl-2′-O-methyl-guanosine ($m^2{}_2Gm$), 1-methyl-2′-O-methyl-guanosine ($m^1Gm$), N2,7-dimethyl-2′-O-methyl-guanosine ($m^{2,7}Gm$), 2′-O-methyl-inosine (Im), 1,2′-O-dimethyl-inosine ($m^1Im$), 2′-O-ribosylguanosine (phosphate) (Gr(**p**)) , 1-thio-guanosine, O6-methyl-guanosine, 2'-F-ara-guanosine, and 2'-F-guanosine.

**[00269]** In some embodiments, the nucleotide can be modified on the major groove face. For example, such modifications include replacing hydrogen on C-5 of uracil or cytosine with alkyl (e.g., methyl) or halo.

**[00270]** The nucleobase of the nucleotide can be independently selected from a purine, a pyrimidine, a purine or pyrimidine analog. For example, the nucleobase can each be independently selected from adenine, cytosine, guanine, uracil, or hypoxanthine. In another embodiment, the nucleobase can also include, for example, naturally-occurring and synthetic derivatives of a base, including pyrazolo[3,4-d]pyrimidines, 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo (*e.g.,* 8-bromo), 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 8-azaguanine and 8-azaadenine, deazaguanine, 7-deazaguanine, 3-deazaguanine, deazaadenine, 7-deazaadenine, 3-deazaadenine, pyrazolo[3,4-d]pyrimidine, imidazo[1,5-a]1,3,5 triazinones, 9-deazapurines, imidazo[4,5-d]pyrazines, thiazolo[4,5-d]pyrimidines, pyrazin-2-ones, 1,2,4-triazine, pyridazine; and 1,3,5 triazine. When the nucleotides are depicted using the shorthand A, G, C, T or U, each letter

refers to the representative base and/or derivatives thereof, *e.g.*, A includes adenine or adenine analogs, *e.g.,* 7-deaza adenine).

**[00271]** In some embodiments, the modified nucleotide is a compound of Formula XI:



XI

**[00272]** wherein:

**[00273]** ⟋⟍ denotes a single or a double bond;

**[00274]** - - -denotes an optional single bond;

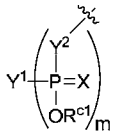
**[00275]** U is O, S, -$NR^a$-, or -$CR^aR^b$- when ⟋⟍ denotes a single bond, or U is -$CR^a$- when ⟋⟍ denotes a double bond;

**[00276]** Z is H, $C_{1-12}$ alkyl, or $C_{6-20}$ aryl, or Z is absent when ⟋⟍ denotes a double bond; and

**[00277]** Z can be -$CR^aR^b$- and form a bond with A;

**[00278]** A is H, OH, NHR wherein R= alkyl or aryl or phosphoryl, sulfate, -$NH_2$, $N_3$, azido, -SH, N an amino acid, or a peptide comprising 1 to 12 amino acids;

**[00279]** D is H, OH, NHR wherein R= alkyl or aryl or phosphoryl, -$NH_2$, -SH, an amino acid, a peptide comprising 1 to 12 amino acids, or a group of Formula XII:



XII

**[00280]** or A and D together with the carbon atoms to which they are attached form a 5-membered ring;

**[00281]** X is O or S;

**[00282]** each of $Y^1$ is independently selected from –$OR^{a1}$, -$NR^{a1}R^{b1}$, and –$SR^{a1}$;

**[00283]** each of $Y^2$ and $Y^3$ are independently selected from O, -$CR^aR^b$-, $NR^c$, S or a linker comprising one or more atoms selected from the group consisting of C, O, N, and S;

**[00284]** n is 0, 1, 2, or 3;

**[00285]** m is 0, 1, 2 or 3;

**[00286]** B is nucleobase;

**[00287]** $R^a$ and $R^b$ are each independently H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, $C_{2-12}$ alkynyl, or $C_{6-20}$ aryl;

**[00288]** $R^c$ is H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;

**[00289]** $R^{a1}$ and $R^{b1}$ are each independently H or a counterion; and

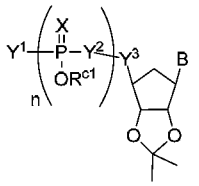
**[00290]** $-OR^{c1}$ is OH at a pH of about 1 or $-OR^{c1}$ is $O^-$ at physiological pH;

**[00291]** provided that the ring encompassing the variables A, B, D, U, Z, $Y^2$ and $Y^3$ cannot be ribose.

**[00292]** In some embodiments, B is a nucleobase selected from the group consisting of cytosine, guanine, adenine, and uracil.

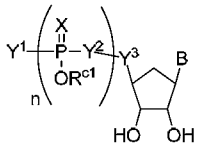
**[00293]** In some embodiments, the nucleobase is a pyrimidine or derivative thereof.

**[00294]** In some embodiments, the modified nucleotides are a compound of Formula XI-a:



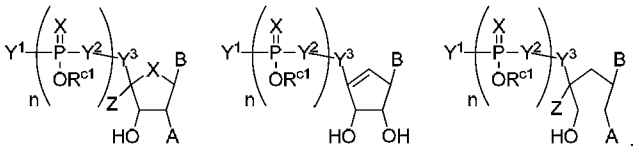
XI-a.

**[00295]** In some embodiments, the modified nucleotides are a compound of Formula XI-b:



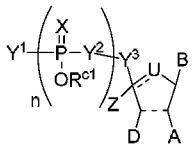
XI-b.

**[00296]** In some embodiments, the modified nucleotides are a compound of Formula XI-c1, XI-c2, or XI-c3:



XI-c1 XI-c2 XI-c3

**[00297]** In some embodiments, the modified nucleotides are a compound of Formula XI:



XI

**[00298]** wherein:

**[00299]** ≈ denotes a single or a double bond;

**[00300]** - - -denotes an optional single bond;

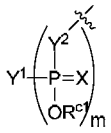
**[00301]** U is O, S, -$NR^a$-, or -$CR^aR^b$- when ≈ denotes a single bond, or U is -$CR^a$- when ≈ denotes a double bond;

**[00302]** Z is H, $C_{1-12}$ alkyl, or $C_{6-20}$ aryl, or Z is absent when ≈ denotes a double bond; and

**[00303]** Z can be -$CR^aR^b$- and form a bond with A;

**[00304]** A is H, OH, sulfate, -$NH_2$, -SH, an amino acid, or a peptide comprising 1 to 12 amino acids;

**[00305]** D is H, OH, -$NH_2$, -SH, an amino acid, a peptide comprising 1 to 12 amino acids, or a group of Formula XII:



XII

**[00306]** or A and D together with the carbon atoms to which they are attached form a 5-membered ring;

**[00307]** X is O or S;

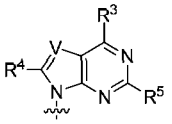
**[00308]** each of $Y^1$ is independently selected from –$OR^{a1}$, -$NR^{a1}R^{b1}$, and –$SR^{a1}$;

**[00309]** each of $Y^2$ and $Y^3$ are independently selected from O, -$CR^aR^b$-, $NR^c$, S or a linker comprising one or more atoms selected from the group consisting of C, O, N, and S;

**[00310]** n is 0, 1, 2, or 3;

**[00311]** m is 0, 1, 2 or 3;

**[00312]** B is a nucleobase of Formula XIII:



XIII

**[00313]** wherein:

**[00314]** V is N or positively charged $NR^c$;

**[00315]** $R^3$ is $NR^cR^d$, $-OR^a$, or $-SR^a$;

**[00316]** $R^4$ is H or can optionally form a bond with $Y^3$;

**[00317]** $R^5$ is H, $-NR^cR^d$, or $-OR^a$;

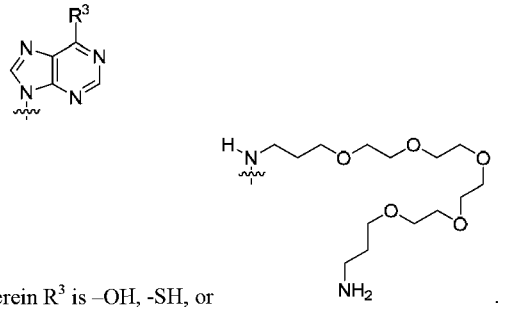
**[00318]** $R^a$ and $R^b$ are each independently H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, $C_{2-12}$ alkynyl, or $C_{6-20}$ aryl;

**[00319]** $R^c$ is H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;

**[00320]** $R^{a1}$ and $R^{b1}$ are each independently H or a counterion; and

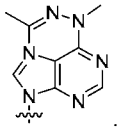
**[00321]** $-OR^{c1}$ is OH at a pH of about 1 or $-OR^{c1}$ is $O^-$ at physiological pH.

**[00322]** In some embodiments, B is:



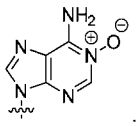
**[00323]** wherein $R^3$ is –OH, -SH, or .

**[00324]** In some embodiments, B is:

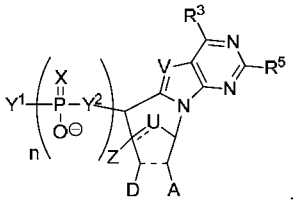


.

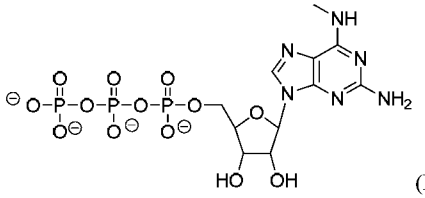
**[00325]** In some embodiments, B is:



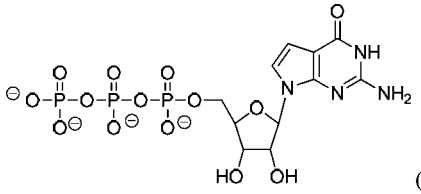
[00326] In some embodiments, the modified nucleotides are a compound of Formula I-d:



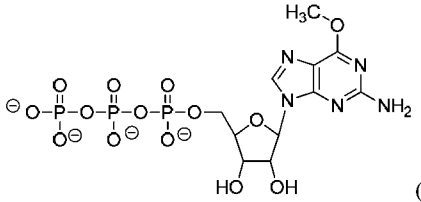
[00327] In some embodiments, the modified nucleotides are a compound selected from the group consisting of:



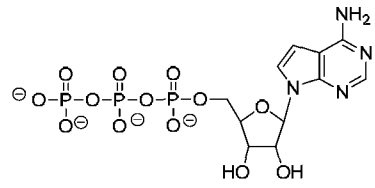
(BB- 247),

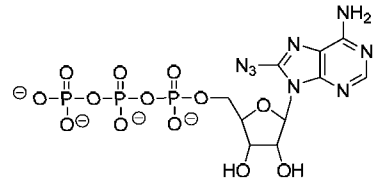


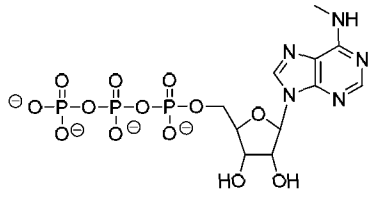
(BB- 248),

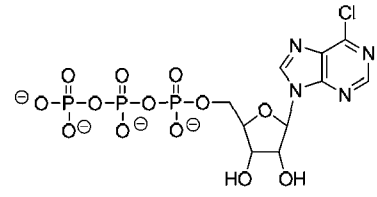


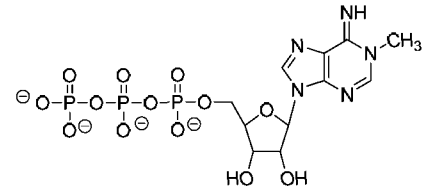
(BB- 249),

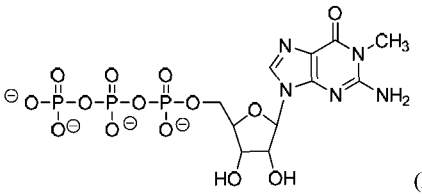

(BB- 250),


(BB- 251),

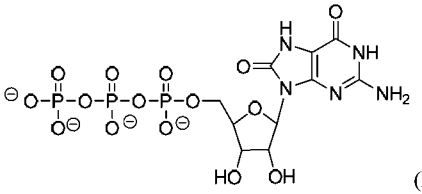

(BB- 252),


(BB- 253),

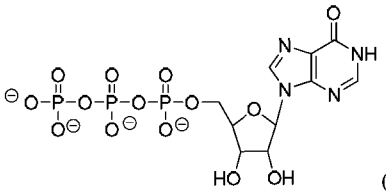

(BB- 254),



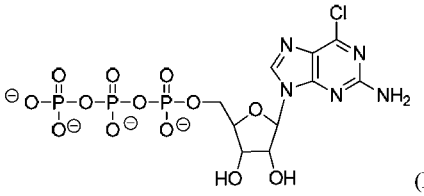
(BB- 255),



(BB- 256),

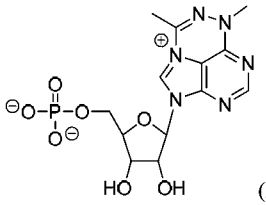


(BB- 257), and

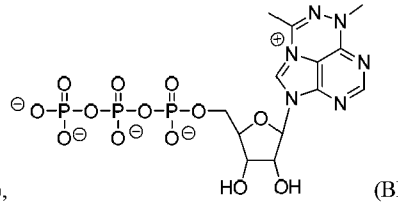


(BB- 258), or a pharmaceutically acceptable salt thereof.

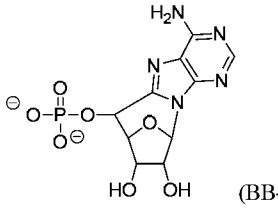
**[00328]** In some embodiments, the modified nucleotides are a compound selected from the group consisting of:

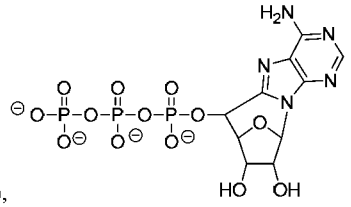


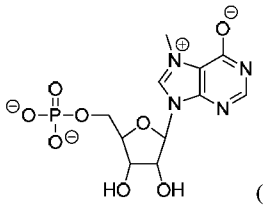
(BB- 259),

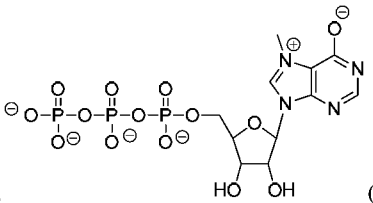


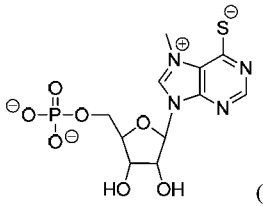
(BB- 260),

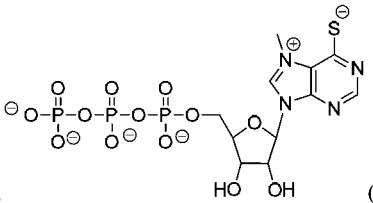

(BB- 261),

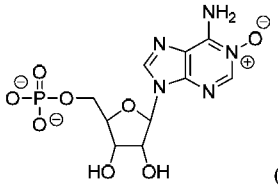

(BB- 262),

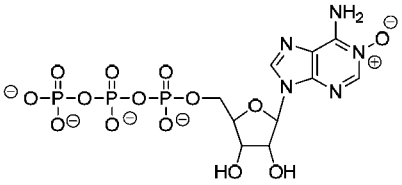

(BB- 263),

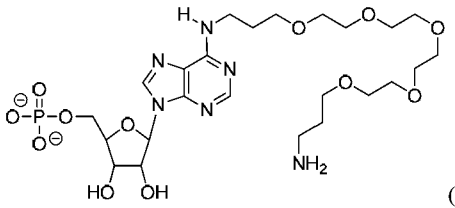

(BB- 264),


(BB- 265),

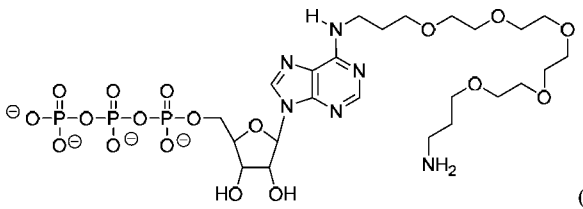

(BB- 266),


(BB- 267),

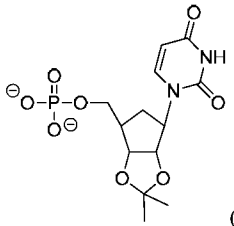

(BB- 268),



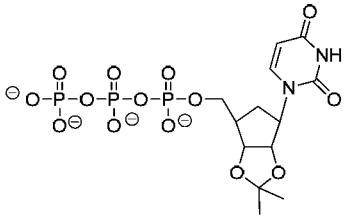
(BB- 269),



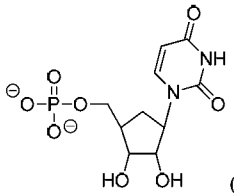
(BB- 270),



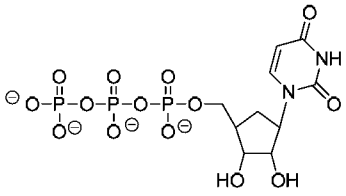
(BB- 271),



(BB- 272),



(BB- 273), and



(BB- 274), or a pharmaceutically acceptable salt thereof.

*Modifications on the Internucleoside Linkage*

**[00329]** The modified nucleotides, which may be incorporated into a polynucleotide molecule, can be modified on the internucleoside linkage (e.g., phosphate backbone). Herein, in the context of the polynucleotide backbone, the phrases "phosphate" and "phosphodiester" are used interchangeably. Backbone phosphate groups 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 another internucleoside linkage as described herein. Examples of modified phosphate groups include, but are not limited to, phosphorothioate,

phosphoroselenates, boranophosphates, boranophosphate esters, hydrogen phosphonates, phosphoramidates, phosphorodiamidates, 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 phosphoramidates), sulfur (bridged phosphorothioates), and carbon (bridged methylene-phosphonates).

**[00330]** 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. While not wishing to be bound by theory, phosphorothioate linked polynucleotide molecules are expected to also reduce the innate immune response through weaker binding/activation of cellular innate immune molecules.

**[00331]** In specific embodiments, a modified nucleoside includes an alpha-thio-nucleoside (e.g., 5′-O-(1-thiophosphate)-adenosine, 5′-O-(1-thiophosphate)-cytidine (α-thio-cytidine), 5′-O-(1-thiophosphate)-guanosine, 5′-O-(1-thiophosphate)-uridine, or 5′-O-(1-thiophosphate)-pseudouridine).

**[00332]** Other internucleoside linkages that may be employed according to the present invention, including internucleoside linkages which do not contain a phosphorous atom, are described herein below.

**Combinations of Modified Sugars, Nucleobases, and Internucleoside Linkages**

**[00333]** The polynucleotides of the invention can include a combination of modifications to the sugar, the nucleobase, and/or the internucleoside linkage. These combinations can include any one or more modifications described herein. For examples, any of the nucleotides described herein in Formulas (Ia), (Ia-1)-(Ia-3), (Ib)-(If), (IIa)-(IIp), (IIb-1), (IIb-2), (IIc-1)-(IIc-2), (IIn-1), (IIn-2), (IVa)-(IVl), and (IXa)-(IXr) can be combined with any of the nucleobases described herein (e.g., in Formulas (b1)-(b43) or any other described herein).

**Synthesis of Polynucleotide Molecules**

**[00334]** The polynucleotide molecules for use in accordance with the invention may be prepared according to any useful technique, as described herein. The modified nucleosides and nucleotides used in the synthesis of polynucleotide molecules disclosed herein can be prepared from readily available starting materials using the following general methods and procedures. Where typical or preferred process conditions (e.g., reaction temperatures, times, mole ratios of reactants, solvents,

pressures, etc.) are provided, a skilled artisan would be able to optimize and develop additional process conditions. 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.

**[00335]** 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., $^{1}H$ or $^{13}C$) infrared spectroscopy, spectrophotometry (e.g., UV-visible), or mass spectrometry, or by chromatography such as high performance liquid chromatography (HPLC) or thin layer chromatography.

**[00336]** Preparation of polynucleotide molecules 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.

**[00337]** 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.

**[00338]** Resolution of racemic mixtures of modified polynucleotides or nucleic acids (e.g., polynucleotides or modified mRNA molecules) 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.

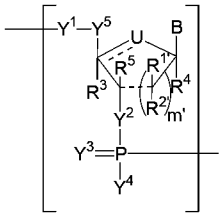
**[00339]** Modified nucleosides and nucleotides (e.g., building block molecules) can be prepared according to the synthetic methods described in Ogata et al., J. Org. Chem. 74:2585-2588 (2009); Purmal et al., Nucl. Acids Res. 22(1): 72-78, (1994); Fukuhara et al., Biochemistry, 1(4): 563-568 (1962); and Xu et al., Tetrahedron, 48(9): 1729-1740 (1992), each of which are incorporated by reference in their entirety.

**[00340]** The polynucleotides of the invention may or may not be uniformly modified along the entire length of the molecule. For example, one or more or all types of nucleotide (*e.g.*, purine or pyrimidine, or any one or more or all of A, G, U, C) may or may not be uniformly modified in a polynucleotide of the invention, or in a given predetermined sequence region thereof. In some embodiments, all nucleotides X in a polynucleotide of the invention (or in a given sequence region thereof) are modified, wherein X may any one of nucleotides A, G, U, C, or any one of the combinations A+G, A+U, A+C, G+U, G+C, U+C, A+G+U, A+G+C, G+U+C or A+G+C.

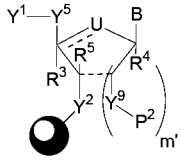
**[00341]** Different sugar modifications, nucleotide modifications, and/or internucleoside linkages (e.g., backbone structures) may exist at various positions in the polynucleotide. One of ordinary skill in the art will appreciate that the nucleotide analogs or other modification(s) may be located at any position(s) of a polynucleotide such that the function of the polynucleotide is not substantially decreased. A modification may also be a 5′ or 3′ terminal modification. The polynucleotide may contain from about 1% to about 100% modified nucleotides (either in relation to overall nucleotide content, or in relation to one or more types of nucleotide, i.e. any one or more of A, G, U or C) or any intervening percentage (e.g., from 1% to 20%, from 1% to 25%, from 1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 20% to 70%, from 20% to 80%, from 20% to 90%, from 20% to 95%, from 20% to 100%, from 50% to 60%, from 50% to 70%, from 50% to 80%, from 50% to 90%, from 50% to 95%, from 50% to 100%, from 70% to 80%, from 70% to 90%, from 70% to 95%, from 70% to 100%, from 80% to 90%, from 80% to 95%, from 80% to 100%, from 90% to 95%, from 90% to 100%, and from 95% to 100%).

**[00342]** In some embodiments, the polynucleotide includes a modified pyrimidine (e.g., a modified uracil/uridine/U or modified cytosine/cytidine/C). In some embodiments, the uracil or uridine (generally: U) in the polynucleotide molecule may be replaced with from about 1% to about 100% of a modified uracil or modified uridine (e.g., from 1% to 20%, from 1% to 25%, from 1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 20% to 70%, from 20% to 80%, from 20% to 90%, from 20% to 95%, from 20% to 100%, from 50% to 60%, from 50% to 70%, from 50% to 80%, from 50% to 90%, from 50% to 95%, from 50% to 100%, from 70% to 80%, from 70% to 90%, from 70% to 95%, from 70% to 100%, from 80% to 90%, from 80% to 95%, from 80% to 100%, from 90% to 95%, from 90% to 100%, and from 95% to 100% of a modified uracil or modified uridine). The modified uracil or uridine can be replaced by a compound having a single unique structure or by a plurality of compounds having different structures (e.g., 2, 3, 4 or more unique structures, as described herein). In some embodiments, the cytosine or cytidine (generally: C) in the polynucleotide molecule may be replaced with from about 1% to about 100% of a modified cytosine or modified cytidine (e.g., from 1% to 20%, from 1% to 25%, from 1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 20% to 70%, from 20% to 80%, from 20% to 90%, from 20% to 95%, from 20% to 100%, from 50% to 60%, from 50% to 70%, from 50% to 80%, from 50% to 90%, from 50% to 95%, from 50% to 100%, from 70% to 80%, from 70% to 90%, from 70% to 95%, from 70% to 100%, from 80% to 90%, from 80% to 95%, from 80% to 100%, from 90% to 95%, from 90% to 100%, and from 95% to 100% of a modified cytosine or modified cytidine). The modified cytosine or cytidine can be replaced by a compound having a single unique structure or by a plurality of compounds having different structures (e.g., 2, 3, 4 or more unique structures, as described herein).

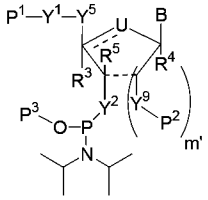
**[00343]** In some embodiments, the present disclosure provides methods of synthesizing a polynucleotide (e.g., the first region, first flanking region, or second flanking region) including n number of linked nucleosides having Formula (Ia-1):


(Ia-1), comprising:

**[00344]** a) reacting a nucleotide of Formula (IV-1):

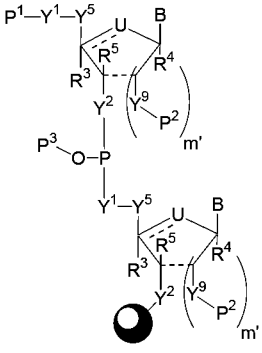

(IV-1),

**[00345]** with a phosphoramidite compound of Formula (V-1):

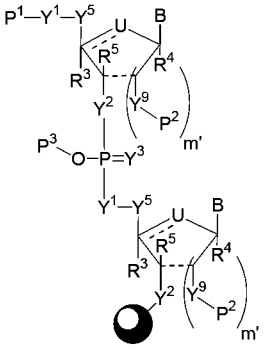

(V-1),

**[00346]** wherein $Y^9$ is H, hydroxy, phosphoryl, pyrophosphate, sulfate, amino, thiol, optionally substituted amino acid, or a peptide (e.g., including from 2 to 12 amino acids); and each $P^1$, $P^2$, and $P^3$ is, independently, a suitable protecting group; and ● denotes a solid support;

**[00347]** to provide a polynucleotide of Formula (VI-1):


(VI-1), and

**[00348]** b) oxidizing or sulfurizing the polynucleotide of Formula (V) to yield a polynucleotide of Formula (VII-1):



(VII-1), and

**[00349]** c) removing the protecting groups to yield the polynucleotide of Formula (Ia).

**[00350]** In some embodiments, steps a) and b) are repeated from 1 to about 10,000 times. In some embodiments, the methods further comprise a nucleotide selected from the group consisting of A, C, G and U adenosine, cytosine, guanosine, and uracil. In some embodiments, the nucleobase may be a pyrimidine or derivative thereof. In some embodiments, the polynucleotide is translatable.

**[00351]** Other components of polynucleotides 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 nucleotide modifications. In such embodiments, nucleotide modifications may also be present in the translatable region. Also provided are polynucleotides containing a Kozak sequence.

**Combinations of Nucleotides**

**[00352]** Further examples of modified nucleotides and modified nucleotide combinations are provided below in Table 2. These combinations of modified nucleotides can be used to form the polynucleotides of the invention. Unless otherwise noted, the modified nucleotides may be completely substituted for the natural nucleotides of the polynucleotides of the invention. As a non-limiting example, the natural nucleotide uridine may be substituted with a modified nucleoside described herein. In another non-limiting example, the natural nucleotide uridine may be partially substituted (e.g., about 0.1%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 99.9%) with at least one of the modified nucleoside disclosed herein.

**Table 2**

| Modified Nucleotide | Modified Nucleotide Combination |
|---|---|
| α-thio-cytidine | α-thio-cytidine/5-iodo-uridine |
| | α-thio-cytidine/N1-methyl-pseudo-uridine |
| | α-thio-cytidine/α-thio-uridine |
| | α-thio-cytidine/5-methyl-uridine |
| | α-thio-cytidine/pseudo-uridine |
| | about 50% of the cytosines are α-thio-cytidine |
| pseudoisocytidine | pseudoisocytidine/5-iodo-uridine |
| | pseudoisocytidine/N1-methyl-pseudouridine |
| | pseudoisocytidine/α-thio-uridine |
| | pseudoisocytidine/5-methyl-uridine |
| | pseudoisocytidine/pseudouridine |
| | about 25% of cytosines are pseudoisocytidine |
| | pseudoisocytidine/about 50% of uridines are N1-methyl-pseudouridine and about 50% of uridines are pseudouridine |
| | pseudoisocytidine/about 25% of uridines are N1-methyl-pseudouridine and about 25% of uridines are pseudouridine (e.g., 25% N1-methyl-pseudouridine/75% pseudouridine) |
| pyrrolo-cytidine | pyrrolo-cytidine/5-iodo-uridine |
| | pyrrolo-cytidine/N1-methyl-pseudouridine |
| | pyrrolo-cytidine/α-thio-uridine |
| | pyrrolo-cytidine/5-methyl-uridine |
| | pyrrolo-cytidine/pseudouridine |
| | about 50% of the cytosines are pyrrolo-cytidine |
| 5-methyl-cytidine | 5-methyl-cytidine/5-iodo-uridine |
| | 5-methyl-cytidine/N1-methyl-pseudouridine |
| | 5-methyl-cytidine/α-thio-uridine |
| | 5-methyl-cytidine/5-methyl-uridine |
| | 5-methyl-cytidine/pseudouridine |
| | about 25% of cytosines are 5-methyl-cytidine |
| | about 50% of cytosines are 5-methyl-cytidine |
| | 5-methyl-cytidine/5-methoxy-uridine |
| | 5-methyl-cytidine/5-bromo-uridine |
| | 5-methyl-cytidine/2-thio-uridine |
| | 5-methyl-cytidine/about 50% of uridines are 2-thio-uridine |
| | about 50% of uridines are 5-methyl-cytidine/ about 50% of uridines are 2-thio-uridine |
| N4-acetyl-cytidine | N4-acetyl-cytidine /5-iodo-uridine |
| | N4-acetyl-cytidine /N1-methyl-pseudouridine |
| | N4-acetyl-cytidine /α-thio-uridine |
| | N4-acetyl-cytidine /5-methyl-uridine |
| | N4-acetyl-cytidine /pseudouridine |

| | |
|---|---|
| | about 50% of cytosines are N4-acetyl-cytidine |
| | about 25% of cytosines are N4-acetyl-cytidine |
| | N4-acetyl-cytidine /5-methoxy-uridine |
| | N4-acetyl-cytidine /5-bromo-uridine |
| | N4-acetyl-cytidine /2-thio-uridine |
| | about 50% of cytosines are N4-acetyl-cytidine/ about 50% of uridines are 2-thio-uridine |

**[00353]** Certain modified nucleotides and nucleotide combinations have been explored by the current inventors. These findings are described in U.S. Provisional Application No 61/404,413, filed on October 1, 2010, entitled Engineered Nucleic Acids and Methods of Use Thereof, U.S. Patent Application No 13/251,840, filed on October 3, 2011, entitled Modified Nucleotides, and Nucleic Acids, and Uses Thereof, now abandoned, U.S. Patent Application No 13/481,127, filed on May 25, 2012, entitled Modified Nucleotides, and Nucleic Acids, and Uses Thereof, International Patent Publication No WO2012045075, filed on October 3, 2011, entitled Modified Nucleosides, Nucleotides, And Nucleic Acids, and Uses Thereof, U.S. Patent Publication No US20120237975 filed on October 3, 2011, entitled Engineered Nucleic Acids and Method of Use Thereof, and International Patent Publication No WO2012045082, which are incorporated by reference in their entireties.

**[00354]** Further examples of modified nucleotide combinations are provided below in Table 3. These combinations of modified nucleotides can be used to form the polynucleotides of the invention.

**Table 3**

| **Modified Nucleotide** | **Modified Nucleotide Combination** |
|---|---|
| modified cytidine having one or more nucleobases of Formula (b10) | modified cytidine with (b10)/pseudouridine |
| | modified cytidine with (b10)/N1-methyl-pseudouridine |
| | modified cytidine with (b10)/5-methoxy-uridine |
| | modified cytidine with (b10)/5-methyl-uridine |
| | modified cytidine with (b10)/5-bromo-uridine |
| | modified cytidine with (b10)/2-thio-uridine |
| | about 50% of cytidine substituted with modified cytidine (b10)/ about 50% of uridines are 2-thio-uridine |
| modified cytidine having one or more nucleobases of Formula (b32) | modified cytidine with (b32)/pseudouridine |
| | modified cytidine with (b32)/N1-methyl-pseudouridine |
| | modified cytidine with (b32)/5-methoxy-uridine |
| | modified cytidine with (b32)/5-methyl-uridine |
| | modified cytidine with (b32)/5-bromo-uridine |

| | |
|---|---|
| | modified cytidine with (b32)/2-thio-uridine |
| | about 50% of cytidine substituted with modified cytidine (b32)/ about 50% of uridines are 2-thio-uridine |
| modified uridine having one or more nucleobases of Formula (b1) | modified uridine with (b1)/ N4-acetyl-cytidine |
| | modified uridine with (b1)/ 5-methyl-cytidine |
| modified uridine having one or more nucleobases of Formula (b8) | modified uridine with (b8)/ N4-acetyl-cytidine |
| | modified uridine with (b8)/ 5-methyl-cytidine |
| modified uridine having one or more nucleobases of Formula (b28) | modified uridine with (b28)/ N4-acetyl-cytidine |
| | modified uridine with (b28)/ 5-methyl-cytidine |
| modified uridine having one or more nucleobases of Formula (b29) | modified uridine with (b29)/ N4-acetyl-cytidine |
| | modified uridine with (b29)/ 5-methyl-cytidine |
| modified uridine having one or more nucleobases of Formula (b30) | modified uridine with (b30)/ N4-acetyl-cytidine |
| | modified uridine with (b30)/ 5-methyl-cytidine |

**[00355]** In some embodiments, at least 25% of the cytosines are replaced by a compound of Formula (b10)-(b14), (b24), (b25), or (b32)-(b35) (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% of, e.g., a compound of Formula (b10) or (b32)).

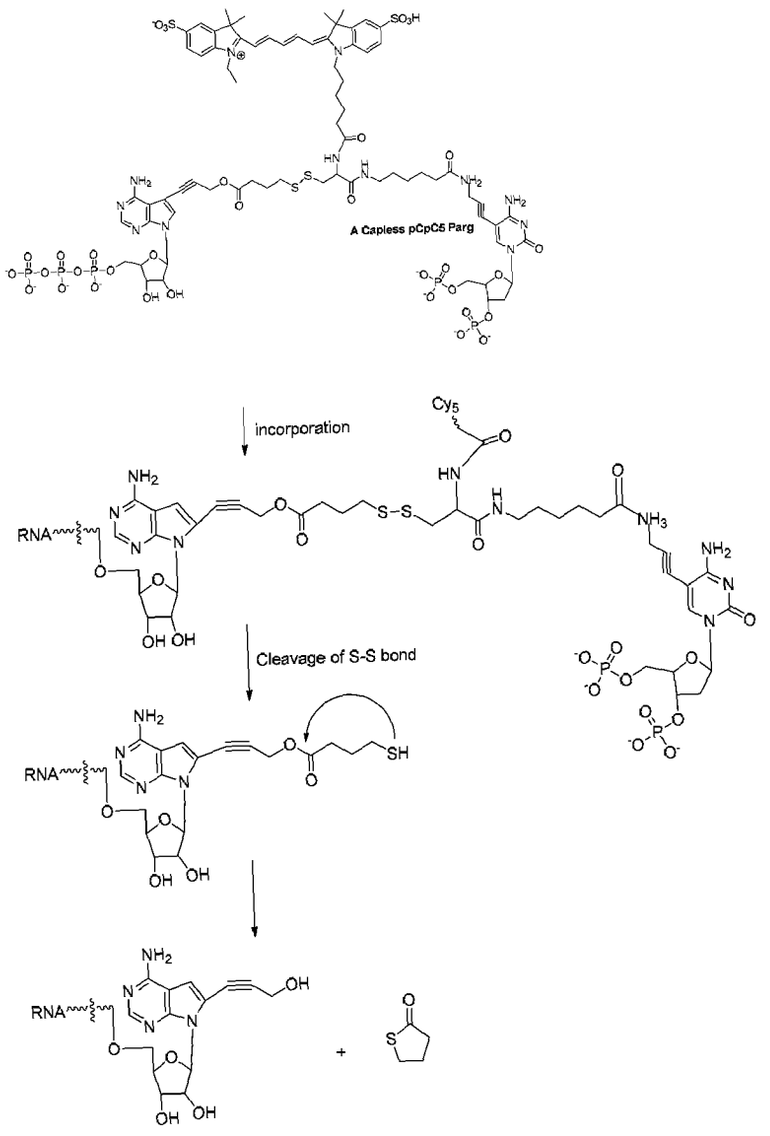
**[00356]** In some embodiments, at least 25% of the uracils are replaced by a compound of Formula (b1)-(b9), (b21)-(b23), or (b28)-(b31) (e.g., at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, 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% of, e.g., a compound of Formula (b1), (b8), (b28), (b29), or (b30)).

**[00357]** In some embodiments, at least 25% of the cytosines are replaced by a compound of Formula (b10)-(b14), (b24), (b25), or (b32)-(b35) (e.g. Formula (b10) or (b32)), and at least 25% of the uracils are replaced by a compound of Formula (b1)-(b9), (b21)-(b23), or (b28)-(b31) (e.g. Formula (b1), (b8), (b28), (b29), or (b30)) (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%).

**Modifications including Linker and a Payload**

**[00358]** The nucleobase of the nucleotide 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 an exemplary 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 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 rationale for the structure of the modified nucleotides is that 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 this function and that the inhibiter be in a stereochemical orientation that inhibits or prohibits second and follow on nucleotides into the growing polynucleotide strand.

**Scheme 1**


A Capless pCpC5 Parg
incorporation
Cy5
RNA
Cleavage of S-S bond
SH
OH
+

**Linker**

**[00359]** 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, alkylamino, oxygen, sulfur, sulfoxide, sulfonyl, 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 is of sufficient length as to not interfere with incorporation into a nucleic acid sequence.

**[00360]** Examples of chemical groups that can be incorporated into the linker include, but are not limited to, an alkyl, alkene, an alkyne, an amido, an ether, a thioether, an 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 is an aryl group containing from 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.

**[00361]** For example, the linker can include 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 a divalent alkyl, alkenyl, and/or alkynyl moiety. The linker can include an ester, amide, or ether moiety.

**[00362]** Other examples include 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. A cleavable bond incorporated into the linker and attached to a modified nucleotide, when cleaved, results in, for example, a short "scar" or chemical modification on the nucleotide. 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**

**[00363]** 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).

*Payload: Therapeutic Agents*

**[00364]** In some embodiments the payload is a therapeutic agent such as a cytotoxin, radioactive ion, chemotherapeutic, or other therapeutic agent. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin 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), CC-1065 (see U.S. Pat. Nos. 5,475,092, 5,585,499, 5,846,545) 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 decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, CC-1065, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, 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).

*Payload:Detectable Agents*

**[00365]** Examples of detectable substances include 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'-isothiocyanatostilbene-2,2'disulfonic acid; acridine and derivatives: acridine, acridine isothiocyanate; 5-(*2'*-aminoethyl)aminonaphthalene-1 -sulfonic acid (EDANS); 4-amino-N-[3-vinylsulfonyl)phenyl]naphthalimide-3,5 disulfonate; N-(4-anilino-l-naphthyl)maleimide; anthranilamide; BODIPY; Brilliant Yellow; coumarin and derivatives; coumarin, 7-amino-4-

methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcouluarin (Coumaran 151); cyanine dyes; cyanosine; 4′,6-diaminidino-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-sulfonyl chloride (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 cresolphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-phthaldialdehyde; 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 is a fluorescent dye, such as Cy5 and Cy3.

**[00366]** Examples luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin.

**[00367]** Examples of suitable radioactive material include $^{18}F$, $^{67}Ga$, $^{81m}Kr$, $^{82}Rb$, $^{111}In$, $^{123}I$, $^{133}Xe$, $^{201}Tl$, $^{125}I$, $^{35}S$, $^{14}C$, or $^{3}H$, $^{99m}Tc$ (e.g., as pertechnetate (technetate(VII), $TcO_4^-$) either directly or indirectly, or other radioisotope detectable by direct counting of radioemission or by scintillation counting.

**[00368]** In addition, contrast agents, *e.g.*, contrast agents for MRI or NMR, for X-ray CT, Raman imaging, optical coherence tomography, absorption imaging, ultrasound imaging, or thermal imaging can be used. Exemplary contrast agents include 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 can also be used.

**[00369]** In some embodiments, the detectable agent is a non-detectable pre-cursor that becomes detectable upon activation. Examples include 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)).

**[00370]** When the compounds are enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, the enzymatic label is detected by determination of conversion of an appropriate substrate to product.

**[00371]** *In vitro* assays in which these compositions can be used include enzyme linked immunosorbent assays (ELISAs), immunoprecipitations, immunofluorescence, enzyme immunoassay (EIA), radioimmunoassay (RIA), and Western blot analysis.

**[00372]** Labels other than those described herein are contemplated by the present disclosure, including 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.

**[00373]** *Payload:Cell Penetrating Payloads*

**[00374]** In some embodiments, the modified nucleotides and modified nucleic acids 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 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):310-8; Langel, Cell-Penetrating Peptides: Processes and Applications (CRC Press, Boca Raton FL 2002); El-Andaloussi et al., (2005) Curr Pharm Des. 11(28):3597-611; and Deshayes et al., (2005) Cell Mol Life Sci. 62(16):1839-49. 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.

*Payload:Biological Targets*

**[00375]** The modified nucleotides and modified nucleic acids described herein 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.

**[00376]** Exemplary biological targets include biopolymers, *e.g.*, antibodies, nucleic acids such as RNA and DNA, proteins, enzymes; exemplary proteins include enzymes, receptors, and ion channels. In some embodiments the target is a tissue- or cell-type specific marker, *e.g.*, a protein that is expressed specifically on a selected tissue or cell type. In some embodiments, the target is a receptor, such as, but not limited to, plasma membrane receptors and nuclear receptors; more specific examples include G-protein-coupled receptors, cell pore proteins, transporter proteins, surface-expressed antibodies, HLA proteins, MHC proteins and growth factor receptors.

**Synthesis of Modified Nucleotides**

**[00377]** The modified nucleosides and nucleotides 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.

**[00378]** 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.*, $^{1}H$ or $^{13}C$) infrared spectroscopy, spectrophotometry (*e.g.*, UV-visible), or mass spectrometry, or by chromatography such as high performance liquid chromatography (HPLC) or thin layer chromatography.

**[00379]** 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.

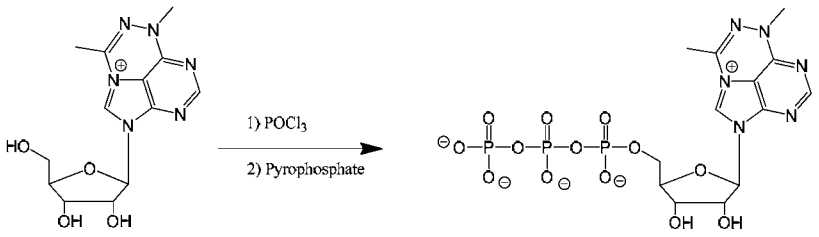
**[00380]** The reactions of the processes described herein can be carried out in suitable solvents, which can be readily selected by one of skill in the art of organic synthesis. Suitable solvents can be substantially nonreactive with the starting materials (reactants), the intermediates, or products 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.

**[00381]** 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.

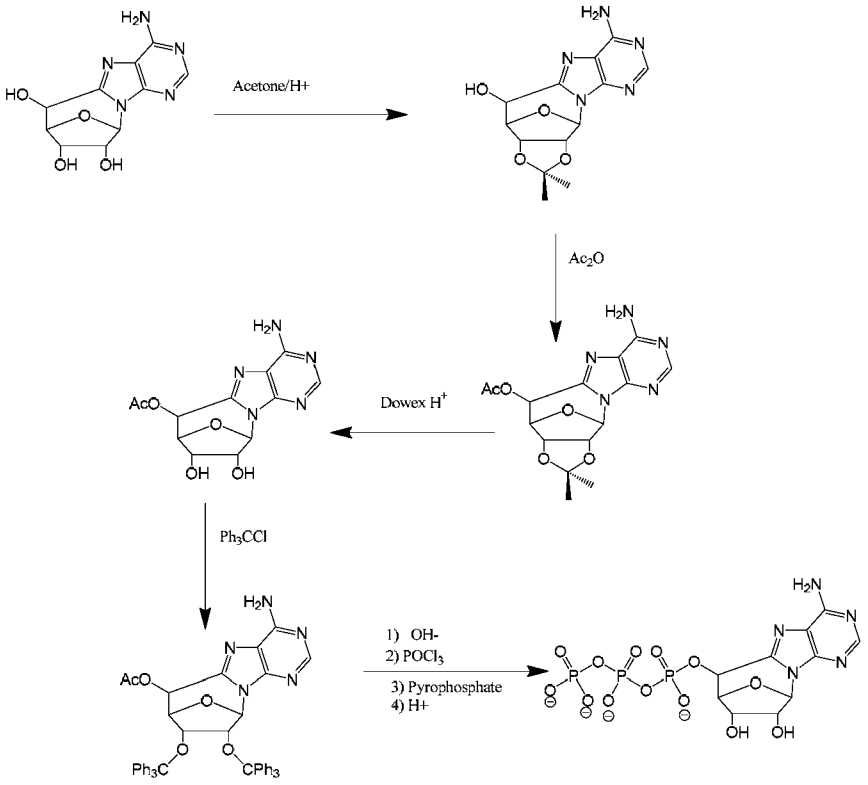
**[00382]** Exemplary syntheses of modified nucleotides, which are incorporated into a polynucleotides, e.g., RNA or mRNA, are provided below in Scheme 2 through Scheme 12. Scheme 2 provides a general method for phosphorylation of nucleosides, including modified nucleosides.

**Scheme 2**



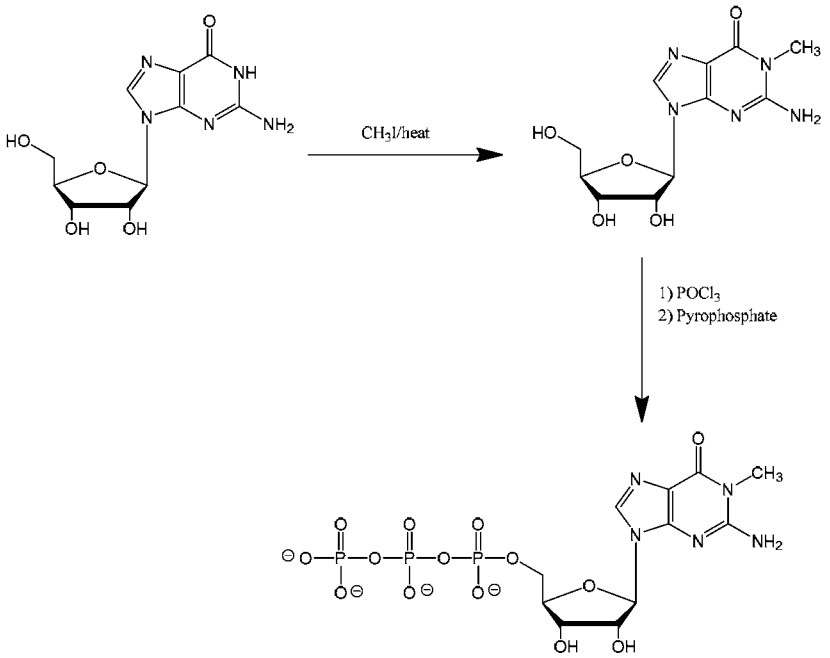

**[00383]** Various protecting groups may be used to control the reaction. For example, Scheme 3 provides the use of multiple protecting and deprotecting steps to promote phosphorylation at the 5′ position of the sugar, rather than the 2′ and 3′ hydroxyl groups.

**Scheme 3**

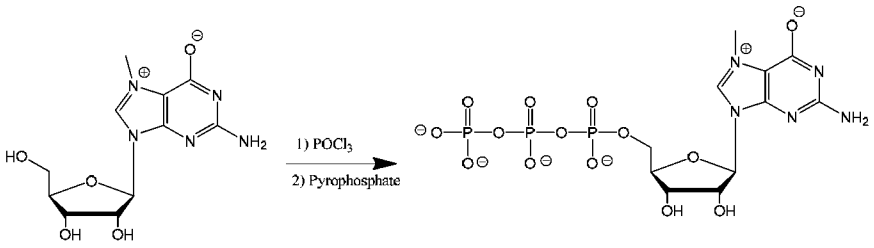



**[00384]** Modified nucleotides can be synthesized in any useful manner. Schemes 4, 5, and 8 provide exemplary methods for synthesizing modified nucleotides having a modified purine nucleobase; and Schemes 6 and 7 provide exemplary methods for synthesizing modified nucleotides having a modified pseudouridine or pseudoisocytidine, respectively.

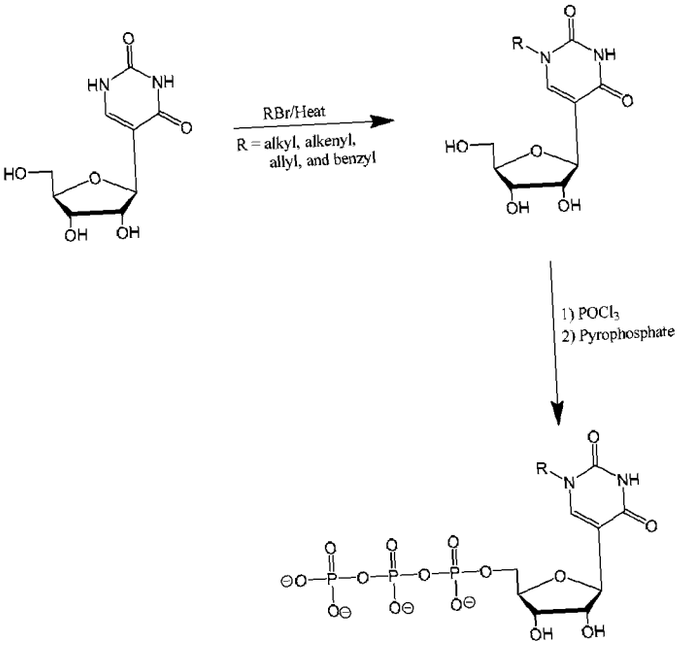
**Scheme 4**



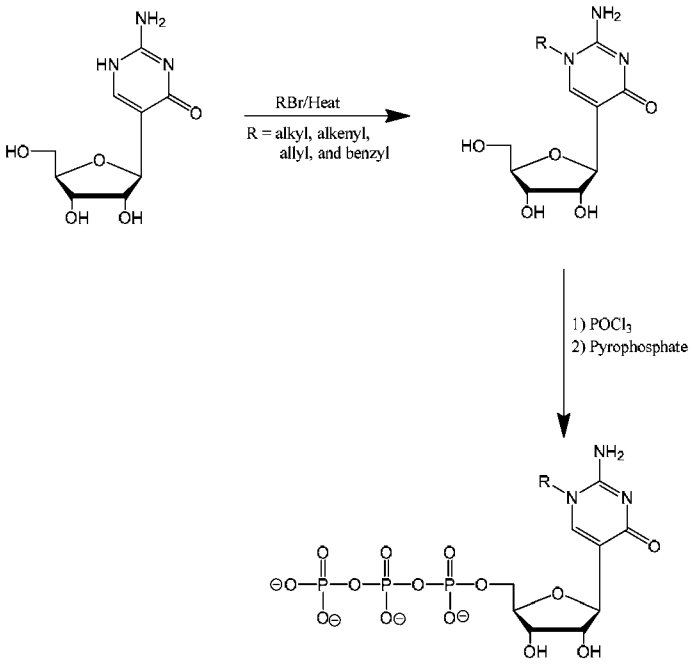

**Scheme 5**



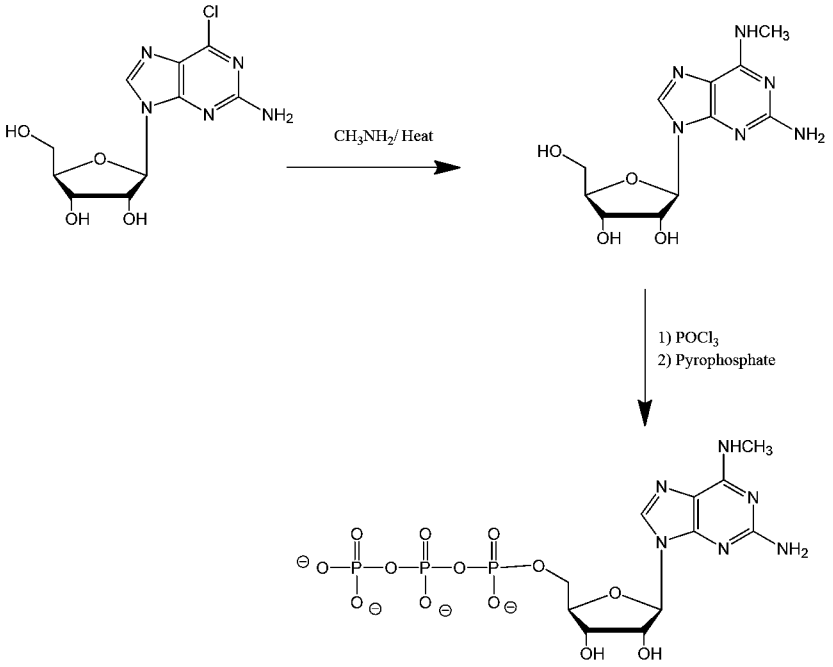
**Scheme 6**


RBr/Heat
R = alkyl, alkenyl, allyl, and benzyl
1) POCl3
2) Pyrophosphate

**Scheme 7**

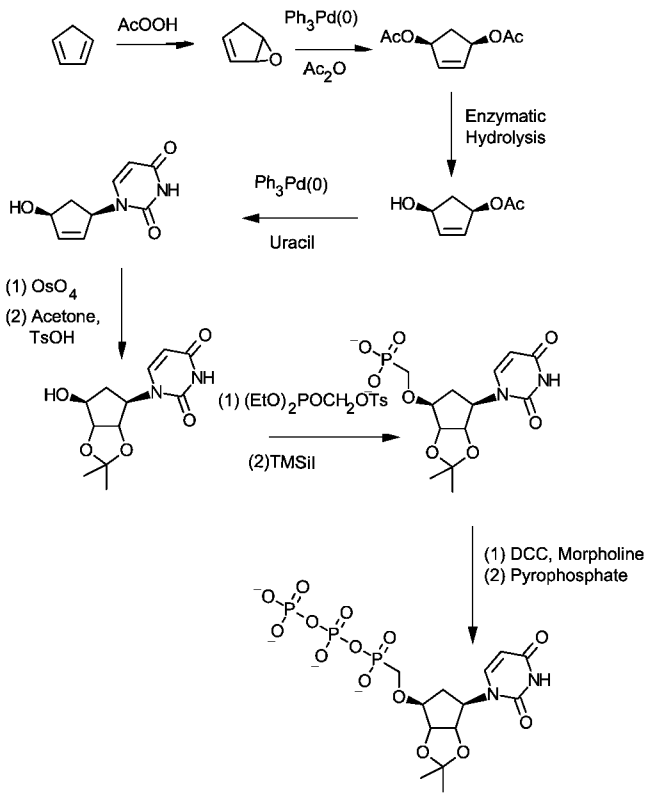

NH2
HN
N
O
HO
O
OH
OH
RBr/Heat
R = alkyl, alkenyl,
allyl, and benzyl
R
1) POCl3
2) Pyrophosphate
⊖O
P
O⊖

**Scheme 8**

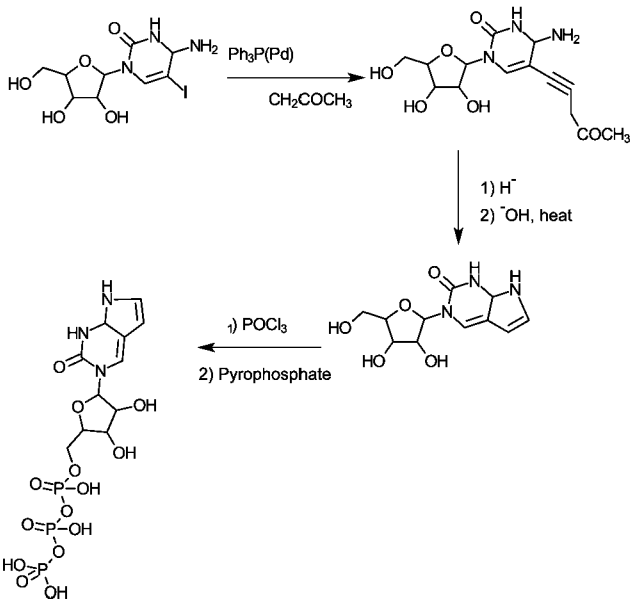



**[00385]** Schemes 9 and 10 provide exemplary syntheses of modified nucleotides. Scheme 11 provides a non-limiting biocatalytic method for producing nucleotides.

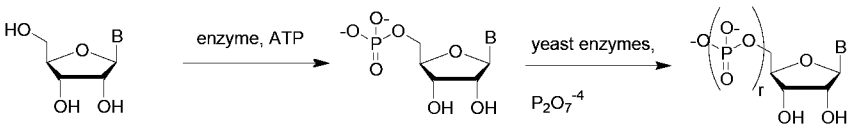
**Scheme 9**


AcOOH
Ph3Pd(0)
Ac2O
AcO
OAc
Enzymatic
Hydrolysis
HO
OAc
Ph3Pd(0)
Uracil
HO
NH
(1) OsO4
(2) Acetone,
TsOH
(1) (EtO)2POCH2OTs
(2)TMSiI
(1) DCC, Morpholine
(2) Pyrophosphate

**Scheme 10**



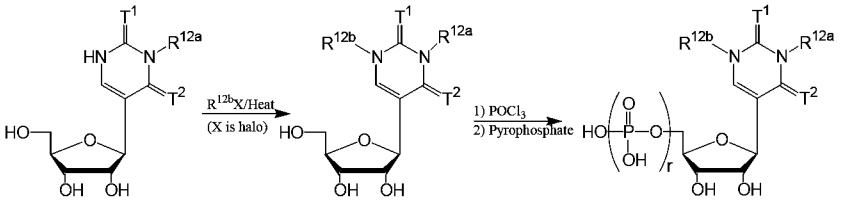

**Scheme 11**




**[00386]** Scheme 12 provides an exemplary synthesis of a modified uracil, where the N1 position on the major groove face is modified with $R^{12b}$, as provided elsewhere, and the 5′-position of ribose is phosphorylated. $T^1$, $T^2$, $R^{12a}$, $R^{12b}$, and r are as provided herein. This synthesis, as well as optimized versions thereof, can be used to modify the major groove face of other pyrimidine nucleobases and purine nucleobases (see e.g., Formulas (b1)-(b43)) and/or to install one or more phosphate groups (e.g., at the 5′ position of the sugar). This alkylating reaction can also be used to include one or more optionally substituted alkyl group at any reactive group (e.g., amino group) in

any nucleobase described herein (e.g., the amino groups in the Watson-Crick base-pairing face for cytosine, uracil, adenine, and guanine).

**Scheme 12**




**[00387]** Modified nucleosides and nucleotides can also be prepared according to the synthetic methods described in Ogata et al. *Journal of Organic Chemistry* 74:2585-2588, 2009; Purmal et al. *Nucleic Acids Research* 22(1): 72-78, 1994; Fukuhara et al. *Biochemistry* 1(4): 563-568, 1962; and Xu et al. *Tetrahedron* 48(9): 1729-1740, 1992, each of which are incorporated by reference in their entirety.

**Modified Nucleic Acids**

**[00388]** The present disclosure provides nucleic acids (or polynucleotides), including RNAs such as mRNAs that contain one or more modified nucleosides (termed "modified nucleic acids") or nucleotides as described herein, 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 also termed "enhanced nucleic acids" herein.

**[00389]** In addition, the present disclosure provides nucleic acids, which have decreased binding affinity to a major groove interacting, *e.g.* binding, partner. For example, the nucleic acids are comprised of at least one nucleotide that has been chemically modified on the major groove face as described herein.

**[00390]** The term "nucleic acid," in its broadest sense, includes any compound and/or substance that is or can be incorporated into an oligonucleotide chain. In this context, the term nucleic acid is used synonymously with polynucleotide. Exemplary nucleic acids for use in accordance with the present disclosure include, but are not limited to, one or more of DNA, RNA including messenger mRNA (mRNA), hybrids thereof, RNAi-inducing agents, RNAi agents, siRNAs, shRNAs, miRNAs,

antisense RNAs, ribozymes, catalytic DNA, RNAs that induce triple helix formation, aptamers, vectors, *etc.*, described in detail herein.

**[00391]** 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. Exemplary nucleic acids include ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), threose nucleic acids (TNAs), glycol nucleic acids (GNAs), or a hybrid thereof. In preferred embodiments, the modified nucleic acid includes messenger RNAs (mRNAs). As described herein, the nucleic acids of the present disclosure do not substantially induce an innate immune response of a cell into which the mRNA is introduced.

**[00392]** 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 present disclosure provides a modified nucleic acid containing a degradation domain, which is capable of being acted on in a directed manner within a cell.

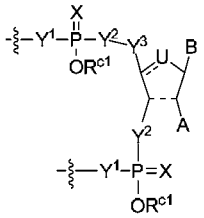
**[00393]** Other components of nucleic acid are optional, and are beneficial in some embodiments. For example, a 5' untranslated region (UTR) and/or a 3'UTR are provided, wherein either or both may independently contain one or more different nucleoside modifications. In such embodiments, nucleoside modifications may also be present in the translatable region. Also provided are nucleic acids containing a Kozak sequence.

**[00394]** Additionally, provided are nucleic acids containing one or more intronic nucleotide sequences capable of being excised from the nucleic acid.

**[00395]** 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 ("multicistronic 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 present disclosure 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).

**[00396]** In another aspect, the present disclosure provides for nucleic acid sequences comprising at least two nucleotides, the nucleic acid sequence comprising a nucleotide that disrupts binding of a major groove binding partner with the nucleic acid sequence, wherein the nucleotide has decreased binding affinity to the major groove binding partner.

**[00397]** In some embodiments, the nucleic acid is a compound of Formula XI-a:



XI-a

**[00398]** wherein:

**[00399]** ⩬ denotes an optional double bond;

**[00400]** - - -denotes an optional single bond;

**[00401]** U is O, S, -$NR^a$-, or -$CR^aR^b$- when ⩬ denotes a single bond, or U is -$CR^a$- when ⩬ denotes a double bond;

**[00402]** A is H, OH, phosphoryl, pyrophosphate, sulfate, -$NH_2$, -SH, an amino acid, a peptide comprising 2 to 12 amino acids;

**[00403]** X is O or S;

**[00404]** each of $Y^1$ is independently selected from –$OR^{a1}$, -$NR^{a1}R^{b1}$, and –$SR^{a1}$;

**[00405]** each of $Y^2$ and $Y^3$ are independently selected from O, -$CR^aR^b$-, $NR^c$, S or a linker comprising one or more atoms selected from the group consisting of C, O, N, and S;

**[00406]** $R^a$ and $R^b$ are each independently H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, $C_{2-12}$ alkynyl, or $C_{6-20}$ aryl;

**[00407]** $R^c$ is H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;

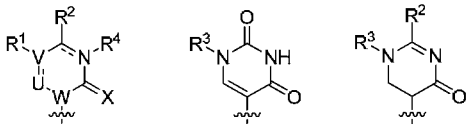
**[00408]** $R^{a1}$ and $R^{b1}$ are each independently H or a counterion;

**[00409]** –$OR^{c1}$ is OH at a pH of about 1 or –$OR^{c1}$ is $O^-$ at physiological pH; and

**[00410]** B is nucleobase;

**[00411]** provided that the ring encompassing the variables A, B, D, U, Z, $Y^2$ and $Y^3$ cannot be ribose.

**[00412]** In some embodiments, B is a nucleobase of Formula XII-a, XII-b, or XII-c:



XII-a XII-b XII-c

**[00413]** wherein:

**[00414]** ⩬ denotes a single or double bond;

**[00415]** X is O or S;

**[00416]** U and W are each independently C or N;

**[00417]** V is O, S, C or N;

**[00418]** wherein when V is C then $R^1$ is H, $C_{1-6}$ alkyl, $C_{1-6}$ alkenyl, $C_{1-6}$ alkynyl, halo, or $-OR^c$, wherein $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl are each optionally substituted with –OH, $-NR^aR^b$, -SH, $-C(O)R^c$, $-C(O)OR^c$, $-NHC(O)R^c$, or $-NHC(O)OR^c$;

**[00419]** and wherein when V is O, S, or N then $R^1$ is absent;

**[00420]** $R^2$ is H, $-OR^c$, $-SR^c$, $-NR^aR^b$, or halo;

**[00421]** 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, $-NR^aR^b$, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl, $C_{1-20}$ alkoxy, or $C_{1-20}$ thioalkyl;

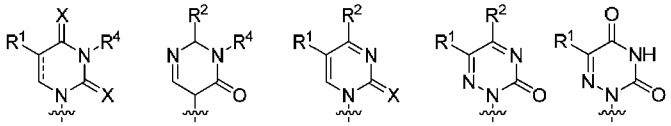
**[00422]** $R^3$ is H or $C_{1-20}$ alkyl;

**[00423]** $R^4$ is H or $C_{1-20}$ alkyl; wherein when ⩬ denotes a double bond then $R^4$ is absent, or N-$R^4$, taken together, forms a positively charged N substituted with $C_{1-20}$ alkyl;

**[00424]** $R^a$ 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

**[00425]** $R^c$ is H, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

**[00426]** In some embodiments, B is a nucleobase of Formula XII-a1, XII-a2, XII-a3, XII-a4, or XII-a5:



XII-a1 XII-a2 XII-a3 XII-a4 XII-a5.

**[00427]** In some embodiments, the nucleobase is a pyrimidine or derivative thereof.

**[00428]** In some embodiments, the nucleic acid contains a plurality of structurally unique compounds of Formula XI-a.

**[00429]** In some embodiments, at least 25% of the cytosines are replaced by a compound of Formula XI-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%).

**[00430]** In some embodiments, at least 25% of the uracils are replaced by a compound of Formula XI-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%).

**[00431]** In some embodiments, at least 25% of the cytosines and 25% of the uracils are replaced by a compound of Formula XI-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%).

**[00432]** In some embodiments, the nucleic acid is translatable.

**[00433]** In some embodiments, when the nucleic acid includes a nucleotide modified with a linker and payload, for example, as described herein, the nucleotide modified with a linker and payload is on the 3' end of the nucleic acid.

**Major Groove Interacting Partners**

**[00434]** As described herein, the phrase "major groove interacting partner" refers 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 as described herein decrease interactions with major groove binding partners, and therefore decrease an innate immune response, or expression and secretion of pro-inflammatory cytokines, or both.

**[00435]** Example major groove interacting, *e.g.* binding, partners include, but are not limited to the following nucleases 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 DE*X*(D/H) helicases and ATPases can sense RNAs to initiate antiviral responses. These helicases include the RIG-I (retinoic acid-inducible gene I) 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.

**Prevention or reduction of innate cellular immune response**

**[00436]** 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 which is triggered by introduction of exogenous nucleic acids, the present disclosure provides modified nucleic acids such as 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 or lack of induction 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.

**[00437]** In some embodiments, the modified nucleic acids, including polynucleotides and/or mRNA molecules are modified in such a way as to not induce, or induce only minimally, an immune response by the recipient cell or organism. Such evasion or avoidance of an immune response trigger or activation is a novel feature of the modified polynucleotides of the present invention.

**[00438]** The present disclosure 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 *in vitro* and/or *in vivo*.

**Polypeptide variants**

**[00439]** 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; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M. Stockton Press, New York, 1991; and Carillo et al., SIAM J. Applied Math. 48, 1073 (1988).

**[00440]** 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 present disclosure 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.

**[00441]** 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 present disclosure. 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 present disclosure. In certain embodiments, a protein sequence to be utilized in accordance with the present disclosure 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**

**[00442]** 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.

**[00443]** 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**

**[00444]** Proper protein translation involves the physical aggregation of a number of polypeptides and nucleic acids associated with the mRNA. Provided by the present disclosure are protein-nucleic acid complexes, 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.

**Untranslatable modified nucleic acids**

**[00445]** 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.

**[00446]** 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).

**Synthesis of Modified Nucleic Acids**

**[00447]** Nucleic acids for use in accordance with the present disclosure 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 Herdewijn, 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).

**[00448]** 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 5% modified nucleotides, at least 10% modified nucleotides, at least 25% modified nucleotides, at least 50% modified nucleotides, at least 80% modified nucleotides, or at least 90% modified nucleotides. 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 is 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 is 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).

**[00449]** Generally, the shortest length of a modified mRNA of the present disclosure can be the length of an mRNA sequence that is sufficient to encode for a dipeptide. In another embodiment, the length of the mRNA sequence is sufficient to encode for a tripeptide. In another embodiment, the length of an mRNA sequence is sufficient to encode for a tetrapeptide. In another embodiment, the length of an mRNA sequence is sufficient to encode for a pentapeptide. In another embodiment, the length of an mRNA sequence is sufficient to encode for a hexapeptide. In another embodiment, the length of an mRNA sequence is sufficient to encode for a heptapeptide. In another embodiment, the length of an mRNA sequence is sufficient to encode for an octapeptide. In another embodiment, the length of an mRNA sequence is sufficient to encode for a nonapeptide. In another embodiment, the length of an mRNA sequence is sufficient to encode for a decapeptide.

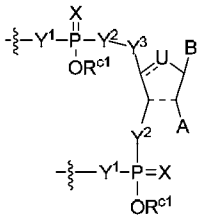
**[00450]** Examples of dipeptides that the modified nucleic acid sequences can encode for include, but are not limited to, carnosine and anserine.

**[00451]** In a further embodiment, the mRNA 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 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.

**[00452]** For example, the modified nucleic acids described herein can be prepared using methods that are known to those skilled in the art of nucleic acid synthesis.

**[00453]** In some embodiments, the present disclosure provides methods, e.g., enzymatic, of preparing a nucleic acid sequence comprising a nucleotide that disrupts binding of a major groove binding partner with the nucleic acid sequence, wherein the nucleic acid sequence comprises a compound of Formula XI-a:



XI-a

**[00454]** wherein:

**[00455]** the nucleotide has decreased binding affinity to the major groove binding partner;

**[00456]** ⩬ denotes an optional double bond;

**[00457]** - - -denotes an optional single bond;

**[00458]** U is O, S, -$NR^a$-, or -$CR^aR^b$- when ⩬ denotes a single bond, or U is -$CR^a$- when ⩬ denotes a double bond;

**[00459]** A is H, OH, phosphoryl, pyrophosphate, sulfate, -$NH_2$, -SH, an amino acid, a peptide comprising 2 to 12 amino acids;

**[00460]** X is O or S;

**[00461]** each of $Y^1$ is independently selected from $-OR^{a1}$, $-NR^{a1}R^{b1}$, and $-SR^{a1}$;

**[00462]** each of $Y^2$ and $Y^3$ are independently selected from O, $-CR^aR^b-$, $NR^c$, S or a linker comprising one or more atoms selected from the group consisting of C, O, N, and S;

**[00463]** $R^a$ and $R^b$ are each independently H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, $C_{2-12}$ alkynyl, or $C_{6-20}$ aryl;

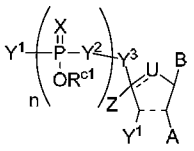
**[00464]** $R^c$ is H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;

**[00465]** $R^{a1}$ and $R^{b1}$ are each independently H or a counterion;

**[00466]** $-OR^{c1}$ is OH at a pH of about 1 or $-OR^{c1}$ is $O^-$ at physiological pH; and

**[00467]** B is nucleobase;

**[00468]** provided that the ring encompassing the variables A, B, D, U, Z, $Y^2$ and $Y^3$ cannot be ribose the method comprising reacting a compound of Formula XIII:

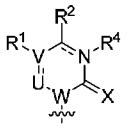


XIII

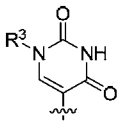
**[00469]** with an RNA polymerase, and a cDNA template.

**[00470]** In some embodiments, the reaction is repeated from 1 to about 7,000 times.

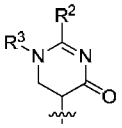
**[00471]** In some embodiments, B is a nucleobase of Formula XII-a, XII-b, or XII-c:



XII-a



XII-b



XII-c

**[00472]** wherein:

**[00473]** ⋰ denotes a single or double bond;

**[00474]** X is O or S;

**[00475]** U and W are each independently C or N;

**[00476]** V is O, S, C or N;

**[00477]** wherein when V is C then $R^1$ is H, $C_{1-6}$ alkyl, $C_{1-6}$ alkenyl, $C_{1-6}$ alkynyl, halo, or $-OR^c$, wherein $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl are each optionally substituted with –OH, $-NR^aR^b$, -SH, $-C(O)R^c$, $-C(O)OR^c$, $-NHC(O)R^c$, or $-NHC(O)OR^c$;

**[00478]** and wherein when V is O, S, or N then $R^1$ is absent;

**[00479]** $R^2$ is H, $-OR^c$, $-SR^c$, $-NR^aR^b$, or halo;

**[00480]** 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, $-NR^aR^b$, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl, $C_{1-20}$ alkoxy, or $C_{1-20}$ thioalkyl;

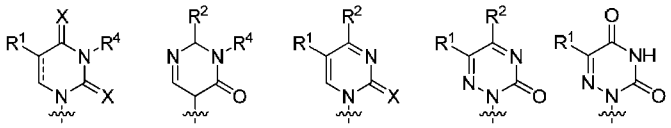
**[00481]** $R^3$ is H or $C_{1-20}$ alkyl;

**[00482]** $R^4$ is H or $C_{1-20}$ alkyl; wherein when $\backsim$ denotes a double bond then $R^4$ is absent, or N-$R^4$, taken together, forms a positively charged N substituted with $C_{1-20}$ alkyl;

**[00483]** $R^a$ 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

**[00484]** $R^c$ is H, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

**[00485]** In some embodiments, B is a nucleobase of Formula XII-a1, XII-a2, XII-a3, XII-a4, or XII-a5:



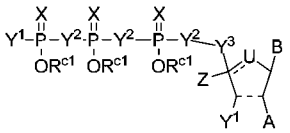
XII-a XII-a2 XII-a3 XII-a4 XII-a5.

**[00486]** In some embodiments, the methods further comprise a nucleotide selected from the group consisting of adenosine, cytosine, guanosine, and uracil.

**[00487]** In some embodiments, the nucleobase is a pyrimidine or derivative thereof.

**[00488]** In another aspect, the present disclosure provides for methods of amplifying a nucleic acid sequence comprising a nucleotide that disrupts binding of a major groove binding partner with the nucleic acid sequence, the method comprising:

**[00489]** reacting a compound of Formula XI-d:



XI-d

**[00490]** wherein:

**[00491]** the nucleotide has decreased binding affinity to the major groove binding partner;

**[00492]** ⟋⟍ denotes a single or a double bond;

**[00493]** - - -denotes an optional single bond;

**[00494]** U is O, S, -$NR^a$-, or -$CR^aR^b$- when ⟋⟍ denotes a single bond, or U is -$CR^a$- when ⟋⟍ denotes a double bond;

**[00495]** Z is H, $C_{1-12}$ alkyl, or $C_{6-20}$ aryl, or Z is absent when ⟋⟍ denotes a double bond; and

**[00496]** Z can be -$CR^aR^b$- and form a bond with A;

**[00497]** A is H, OH, phosphoryl, pyrophosphate, sulfate, -$NH_2$, -SH, an amino acid, or a peptide comprising 1 to 12 amino acids;

**[00498]** X is O or S;

**[00499]** each of $Y^1$ is independently selected from –$OR^{a1}$, -$NR^{a1}R^{b1}$, and –$SR^{a1}$;

**[00500]** each of $Y^2$ and $Y^3$ are independently selected from O, -$CR^aR^b$-, $NR^c$, S or a linker comprising one or more atoms selected from the group consisting of C, O, N, and S;

**[00501]** n is 0, 1, 2, or 3;

**[00502]** m is 0, 1, 2 or 3;

**[00503]** B is nucleobase;

**[00504]** $R^a$ and $R^b$ are each independently H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, $C_{2-12}$ alkynyl, or $C_{6-20}$ aryl;

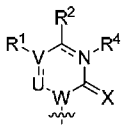
**[00505]** $R^c$ is H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;

**[00506]** $R^{a1}$ and $R^{b1}$ are each independently H or a counterion; and

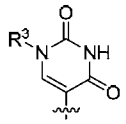
**[00507]** –$OR^{c1}$ is OH at a pH of about 1 or –$OR^{c1}$ is $O^-$ at physiological pH;

**[00508]** provided that the ring encompassing the variables A, B, D, U, Z, $Y^2$ and $Y^3$ cannot be ribose with a primer, a cDNA template, and an RNA polymerase.

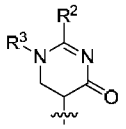
**[00509]** In some embodiments, B is a nucleobase of Formula XII-a, XII-b, or XII-c:




XII-a




XII-b




XII-c

**[00510]** wherein:

**[00511]** ⟋⟍ denotes a single or double bond;

**[00512]** X is O or S;

**[00513]** U and W are each independently C or N;

**[00514]** V is O, S, C or N;

**[00515]** wherein when V is C then $R^1$ is H, $C_{1-6}$ alkyl, $C_{1-6}$ alkenyl, $C_{1-6}$ alkynyl, halo, or $-OR^c$, wherein $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl are each optionally substituted with –OH, $-NR^aR^b$, -SH, $-C(O)R^c$, $-C(O)OR^c$, $-NHC(O)R^c$, or $-NHC(O)OR^c$;

**[00516]** and wherein when V is O, S, or N then $R^1$ is absent;

**[00517]** $R^2$ is H, $-OR^c$, $-SR^c$, $-NR^aR^b$, or halo;

**[00518]** 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, $-NR^aR^b$, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl, $C_{1-20}$ alkoxy, or $C_{1-20}$ thioalkyl;

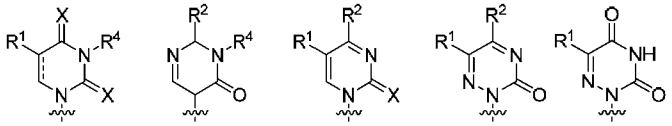
**[00519]** $R^3$ is H or $C_{1-20}$ alkyl;

**[00520]** $R^4$ is H or $C_{1-20}$ alkyl; wherein when $\backslash\!\!=\!\!/$ denotes a double bond then $R^4$ is absent, or N-$R^4$, taken together, forms a positively charged N substituted with $C_{1-20}$ alkyl;

**[00521]** $R^a$ 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

**[00522]** $R^c$ is H, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

**[00523]** In some embodiments, B is a nucleobase of Formula XII-a1, XII-a2, XII-a3, XII-a4, or XII-a5:



XII-a1 XII-a2 XII-a3 XII-a4 XII-a5.

**[00524]** In some embodiments, the methods further comprise a nucleotide selected from the group consisting of adenosine, cytosine, guanosine, and uracil.

**[00525]** In some embodiments, the nucleobase is a pyrimidine or derivative thereof.

**[00526]** In some embodiments, the present disclosure provides for methods of synthesizing a pharmaceutical nucleic acid, comprising the steps of:

**[00527]** a) providing a complementary deoxyribonucleic acid (cDNA) that encodes a pharmaceutical protein of interest;

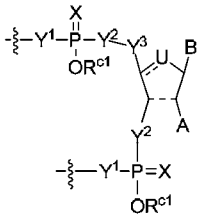
**[00528]** b) selecting a nucleotide that is known to disrupt a binding of a major groove binding partner with a nucleic acid, wherein the nucleotide has decreased binding affinity to the major groove binding partner; and

**[00529]** c) contacting the provided cDNA and the selected nucleotide with an RNA polymerase, under conditions such that the pharmaceutical nucleic acid is synthesized.

**[00530]** In further embodiments, the pharmaceutical nucleic acid is a ribonucleic acid (RNA).

**[00531]** In still a further aspect of the present disclosure, the modified nucleic acids can be prepared using solid phase synthesis methods.

**[00532]** In some embodiments, the present disclosure provides methods of synthesizing a nucleic acid comprising a compound of Formula XI-a:



XI-a

**[00533]** wherein:

**[00534]** ⟋⟍ denotes an optional double bond;

**[00535]** ---denotes an optional single bond;

**[00536]** U is O, S, $-NR^a-$, or $-CR^aR^b-$ when ⟋⟍ denotes a single bond, or U is $-CR^a-$ when ⟋⟍ denotes a double bond;

**[00537]** A is H, OH, phosphoryl, pyrophosphate, sulfate, $-NH_2$, -SH, an amino acid, a peptide comprising 2 to 12 amino acids;

**[00538]** X is O or S;

**[00539]** each of $Y^1$ is independently selected from $-OR^{a1}$, $-NR^{a1}R^{b1}$, and $-SR^{a1}$;

**[00540]** each of $Y^2$ and $Y^3$ are independently selected from O, $-CR^aR^b-$, $NR^c$, S or a linker comprising one or more atoms selected from the group consisting of C, O, N, and S;

**[00541]** $R^a$ and $R^b$ are each independently H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, $C_{2-12}$ alkynyl, or $C_{6-20}$ aryl;

**[00542]** $R^c$ is H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group;

**[00543]** $R^{a1}$ and $R^{b1}$ are each independently H or a counterion;

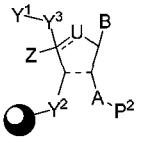
**[00544]** $-OR^{c1}$ is OH at a pH of about 1 or $-OR^{c1}$ is $O^-$ at physiological pH; and

**[00545]** B is nucleobase;

**[00546]** provided that the ring encompassing the variables A, B, U, Z, $Y^2$ and $Y^3$ cannot be ribose;

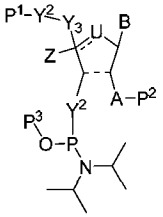
**[00547]** comprising:

**[00548]** a) reacting a nucleotide of Formula XIII-a:



XIII-a

**[00549]** with a phosphoramidite compound of Formula XIII-b:

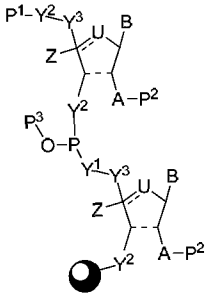


XXIII-b

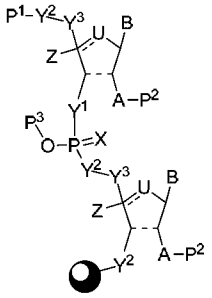
**[00550]** wherein: ● denotes a solid support; and

**[00551]** $P^1$, $P^2$ and $P^3$ are each independently suitable protecting groups;

**[00552]** to provide a nucleic acid of Formula XIV-a:



XIV-a and b) oxidizing or sulfurizing the nucleic acid of Formula XIV-a to yield a nucleic acid of Formula XIVb:

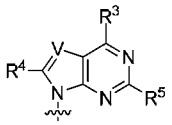


XIV-b

**[00553]** and c) removing the protecting groups to yield the nucleic acid of Formula XI-a.

**[00554]** In some embodiments, the methods further comprise a nucleotide selected from the group consisting of adenosine, cytosine, guanosine, and uracil.

**[00555]** In some embodiments, B is a nucleobase of Formula XIII:



XIII

**[00556]** wherein:

**[00557]** V is N or positively charged $NR^c$;

**[00558]** $R^3$ is $NR^cR^d$, $-OR^a$, or $-SR^a$;

**[00559]** $R^4$ is H or can optionally form a bond with $Y^3$;

**[00560]** $R^5$ is H, $-NR^cR^d$, or $-OR^a$;

**[00561]** $R^a$ and $R^b$ are each independently H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, $C_{2-12}$ alkynyl, or $C_{6-20}$ aryl; and

**[00562]** $R^c$ is H, $C_{1-12}$ alkyl, $C_{2-12}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

**[00563]** In some embodiments, steps a) and b) are repeated from 1 to about 10,000 times.

**Uses of Modified Nucleic Acids**

*Therapeutic Agents*

**[00564]** The modified nucleic acids described herein can be used as therapeutic agents. For example, a modified nucleic acid described herein can be administered to an animal or subject, wherein the modified nucleic acid is translated in vivo to produce a therapeutic peptide in the animal or subject. Accordingly, provided herein 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 present disclosure include modified nucleic acids, cells containing modified nucleic acids or polypeptides translated from the modified nucleic acids, polypeptides translated from modified nucleic acids, cells contacted with cells containing modified nucleic acids or polypeptides translated from the modified nucleic acids, tissues containing cells containing modified nucleic acids and organs containing tissues containing cells containing modified nucleic acids.

**[00565]** Provided are methods of inducing translation of a synthetic or recombinant polynucleotide to produce a 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 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.

**[00566]** 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 or improve therapeutic utility.

**[00567]** Aspects of the present disclosure 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 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 or cells 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.

**[00568]** Other aspects of the present disclosure 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, intraperitoneally, intravenously, intranasally, subcutaneously, endoscopically, transdermally, or intrathecally. In some embodiments, the composition is formulated for extended release.

**[00569]** 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.

**[00570]** 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.

**[00571]** 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 other embodiments, the administered modified nucleic acid directs production of one or more recombinant polypeptides to supplement the amount of polypeptide (or multiple polypeptides) that is present in the cell in which the recombinant polypeptide is translated. 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, due 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, or a small molecule toxin.

**[00572]** 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.

**[00573]** As described herein, a useful feature of the modified nucleic acids of the present disclosure is the capacity to reduce, evade, avoid or eliminate 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.

*Therapeutics for diseases and conditions*

**[00574]** 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 disclosure 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 present disclosure is advantageous in that accurate titration of protein production is achievable. Multiple diseases are

characterized by missing (or substantially diminished such that proper protein function does not occur) protein activity. Such proteins may not be present, are present in very low quantities or are essentially non-functional. The present disclosure 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.

**[00575]** Diseases characterized by dysfunctional or aberrant protein activity include, but not limited to, cancer and proliferative diseases, genetic diseases (*e.g.*, cystic fibrosis), autoimmune diseases, diabetes, neurodegenerative diseases, cardiovascular diseases, and metabolic diseases. The present disclosure 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.

**[00576]** Specific examples of a dysfunctional protein are the missense or nonsense mutation variants of the cystic fibrosis transmembrane conductance regulator (CFTR) gene, which produce a dysfunctional or nonfunctional, respectively, protein variant of CFTR protein, which causes cystic fibrosis.

**[00577]** 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 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.

**[00578]** In another embodiment, the present disclosure 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 1p13 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 1p13 cholesterol locus. *Nature* 2010; 466: 714-721).

*Methods of cellular nucleic acid delivery*

**[00579]** Methods of the present disclosure 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.

**[00580]** 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.

*Targeting Moieties*

**[00581]** In embodiments of the present disclosure, 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.

*Permanent Gene Expression Silencing*

**[00582]** A method for epigenetically silencing gene expression in a mammalian subject, comprising a nucleic acid where the translatable region encodes a polypeptide or polypeptides capable of directing sequence-specific histone H3 methylation to initiate heterochromatin formation and reduce gene transcription around specific genes for the purpose of silencing the gene. For example, a gain-of-function mutation in the Janus Kinase 2 gene is responsible for the family of Myeloproliferative Diseases.

*Delivery of a Detectable or Therapeutic Agent to a Biological Target*

**[00583]** The modified nucleosides, modified nucleotides, and modified nucleic acids described herein can be used in a number of different scenarios in which delivery of a substance (the "payload") to a biological target is desired, for example delivery of detectable substances for detection of the target, or delivery of a therapeutic agent. Detection methods can include both imaging *in vitro* and *in vivo* imaging methods, *e.g.*, immunohistochemistry, bioluminescence imaging (BLI), Magnetic Resonance Imaging (MRI), positron emission tomography (PET), electron microscopy, X-ray computed tomography, Raman imaging, optical coherence tomography, absorption imaging, thermal imaging, fluorescence reflectance imaging, fluorescence microscopy, fluorescence molecular tomographic imaging, nuclear magnetic resonance imaging, X-ray imaging, ultrasound imaging, photoacoustic imaging, lab assays, or in any situation where tagging/staining/imaging is required.

**[00584]** For example, the modified nucleosides, modified nucleotides, and modified nucleic acids described herein can be used in reprogramming induced pluripotent stem cells (iPS cells), which can then be used to directly track cells that are transfected compared to total cells in the cluster. In another example, a drug that is attached to the modified nucleic acid via a linker and is fluorescently labeled can be used to track the drug *in vivo*, *e.g.* intracellularly. Other examples include the use of a modified nucleic acid in reversible drug delivery into cells.

**[00585]** The modified nucleosides, modified nucleotides, and modified nucleic acids described herein can be used in intracellular targeting of a payload, *e.g.*, detectable or therapeutic agent, to specific organelle. Exemplary intracellular targets can include the nuclear localization for advanced mRNA processing, or a nuclear localization sequence (NLS) linked to the mRNA containing an inhibitor.

**[00586]** In addition, the modified nucleosides, modified nucleotides, and modified nucleic acids described herein can be used to deliver therapeutic agents to cells or tissues, *e.g.*, in living animals. For example, the modified nucleosides, modified nucleotides, and modified nucleic acids described herein can be used to deliver highly polar chemotherapeutics agents to kill cancer cells. The modified nucleic acids attached to the therapeutic agent through a linker can facilitate member permeation allowing the therapeutic agent to travel into a cell to reach an intracellular target.

**[00587]** In another example, the modified nucleosides, modified nucleotides, and modified nucleic acids can be attached to a viral inhibitory peptide (VIP) through a cleavable linker. The cleavable linker will release the VIP and dye into the cell. In another example, the modified nucleosides, modified nucleotides, and modified nucleic acids can be attached through the linker to a ADP-ribosylate, which is responsible for the actions of some bacterial toxins, such as cholera toxin, diphtheria toxin, and pertussis toxin. These toxin proteins are ADP-ribosyltransferases that modify target proteins in human cells. For example, cholera toxin ADP-ribosylates G proteins, causing massive fluid secretion from the lining of the small intestine, resulting in life-threatening diarrhea.

**Pharmaceutical Compositions**

**[00588]** The present disclosure provides proteins 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 a modified nucleic acide encoding one or more proteins to be delivered to a subject in need thereof is provided. In some embodiments, compositions are administered to humans. For the purposes of the present disclosure, the phrase "active ingredient" generally refers to a protein, protein encoding or protein-containing complex as described herein.

**[00589]** 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.

**[00590]** 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.

**[00591]** A pharmaceutical composition in accordance with the present disclosure 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.

**[00592]** Relative amounts of the active ingredient, the pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the present disclosure 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.

**[00593]** 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 present disclosure.

**[00594]** 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.

**[00595]** 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.

**[00596]** 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.

**[00597]** 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 (croscarmellose), 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.

**[00598]** 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, glyceryl 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.* 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], glyceryl monooleate, sorbitan monooleate [Span®80]), polyoxyethylene esters (*e.g.* polyoxyethylene monostearate [Myrj®45], polyoxyethylene hydrogenated castor oil, polyethoxylated castor oil, polyoxymethylene 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.

**[00599]** 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); alginates; polyethylene oxide; polyethylene glycol; inorganic calcium salts; silicic acid; polymethacrylates; waxes; water; alcohol; *etc.*; and combinations thereof.

**[00600]** 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, acorbyl 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, deteroxime mesylate, cetrimide, butylated hydroxyanisol (BHA), butylated hydroxytoluened (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®.

**[00601]** 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 glubionate, calcium gluceptate, calcium gluconate, d-gluconic acid, calcium glycerophosphate, 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.

**[00602]** Exemplary lubricating agents include, but are not limited to, magnesium stearate, calcium stearate, stearic acid, silica, talc, malt, glyceryl behanate, hydrogenated vegetable oils, polyethylene glycol, sodium benzoate, sodium acetate, sodium chloride, leucine, magnesium lauryl sulfate, sodium lauryl sulfate, *etc.*, and combinations thereof.

**[00603]** 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, macademia 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, diethyl sebacate, dimethicone 360, isopropyl myristate, mineral oil, octyldodecanol, oleyl alcohol, silicone oil, and/or combinations thereof.

**[00604]** 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-butylene 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.

**[00605]** 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-butanediol. 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.

**[00606]** 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.

**[00607]** 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.

**[00608]** 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.

**[00609]** 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.

**[00610]** 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.

**[00611]** 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 disclosure 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.

**[00612]** 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; 5,704,911; 5,383,851; 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.

**[00613]** 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.

**[00614]** 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 conveniently 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.

**[00615]** 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).

**[00616]** 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.

**[00617]** 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.

**[00618]** 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.

**[00619]** 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 opthalmically-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 present disclosure.

**[00620]** 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).

*Administration*

**[00621]** The present disclosure provides methods comprising administering proteins or complexes in accordance with the present disclosure to a subject in need thereof. 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 present disclosure 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 disclosure 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.

**[00622]** Proteins to be delivered and/or pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof may be administered to animals, such as mammals (*e.g.*, humans, domesticated animals, cats, dogs, mice, rats, *etc.*). In some embodiments, pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof are administered to humans.

**[00623]** Proteins to be delivered and/or pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof in accordance with the present disclosure may be administered by any route. In some embodiments, proteins and/or pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof, are administered by one or more of a variety of routes, including oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, subcutaneous, intraventricular, transdermal, interdermal, rectal, intravaginal, intraperitoneal, topical (*e.g.* by powders, ointments, creams, gels, lotions, and/or drops), mucosal, nasal, buccal, enteral, vitreal, intratumoral, sublingual; by intratracheal instillation, bronchial instillation, and/or inhalation; as an oral spray, nasal spray, and/or aerosol, and/or through a portal vein catheter. In some embodiments, proteins or complexes, and/or pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof, are administered by systemic intravenous injection. In specific embodiments, proteins or complexes and/or pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof may be administered intravenously and/or orally. In specific embodiments, proteins or complexes, and/or pharmaceutical, prophylactic, diagnostic, or imaging compositions thereof, may be administered in a way which allows the protein or complex to cross the blood-brain barrier, vascular barrier, or other epithelial barrier.

**[00624]** However, the present disclosure encompasses the delivery of proteins 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.

**[00625]** In general the most appropriate route of administration will depend upon a variety of factors including the nature of the protein or complex comprising proteins associated with at least one agent to be delivered (*e.g.*, its stability in the environment of the gastrointestinal tract, bloodstream, *etc.*), the condition of the patient (*e.g.*, whether the patient is able to tolerate particular routes of administration), *etc.* The present disclosure encompasses the delivery of the pharmaceutical, prophylactic, diagnostic, or imaging compositions by any appropriate route taking into consideration likely advances in the sciences of drug delivery.

**[00626]** In certain embodiments, compositions in accordance with the present disclosure 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).

**[00627]** Proteins 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 improve their bioavailability, reduce and/or modify their metabolism, inhibit their excretion, and/or modify their distribution within the body.

**[00628]** It will further be appreciated that therapeutically, prophylactically, diagnostically, or imaging active agents utilized in combination may be administered together in a single composition or administered separately in different compositions. In general, it is expected that agents utilized in combination with be utilized at levels that do not exceed the levels at which they are utilized individually. In some embodiments, the levels utilized in combination will be lower than those utilized individually.

**[00629]** The particular combination of therapies (therapeutics or procedures) to employ in a combination regimen will take into account compatibility of the desired therapeutics and/or procedures and the desired therapeutic effect to be achieved. It will also be appreciated that the therapies employed may achieve a desired effect for the same disorder (for example, a composition useful for treating cancer in accordance with the present disclosure may be administered

concurrently with a chemotherapeutic agent), or they may achieve different effects (*e.g.*, control of any adverse effects).

**Kits**

**[00630]** The present disclosure provides a variety of kits for conveniently and/or effectively carrying out methods of the present disclosure. 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.

**[00631]** In one aspect, the disclosure provides kits for protein production, comprising a first isolated nucleic acid comprising a translatable region and a nucleic acid modification, wherein the nucleic acid is capable of evading or avoiding induction of an innate immune response of a cell into which the first isolated nucleic acid is introduced, and packaging and instructions.

**[00632]** In one aspect, the disclosure provides kits for protein production, comprising: a first isolated modified 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.

**[00633]** In one aspect, the disclosure 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.

**[00634]** In one aspect, the disclosure 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.

**[00635]** In one aspect, the disclosure 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.

**[00636]** 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-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, 4-methoxy-2-thio-pseudouridine or any disclosed herein.

**[00637]** 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, 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, 4-methoxy-1-methyl-pseudoisocytidine or any disclosed herein.

**[00638]** 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, 2-methoxy-adenine or any disclosed herein.

**[00639]** 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, N2,N2-dimethyl-6-thio-guanosine or any disclosed herein.

**[00640]** 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**

**[00641]** 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 "$C_{1-6}$ alkyl" is specifically intended to individually disclose methyl, ethyl, $C_3$ alkyl, $C_4$ alkyl, $C_5$ alkyl, and $C_6$ alkyl.

**[00642]** *About:* As used herein, the term "about" means +/- 10% of the recited value.

**[00643]** *Administered in combination:* As used herein, the term "administered in combination" or "combined administration" means that two or more agents are administered to a subject at the same time or within an interval such that there may be an overlap of an effect of each agent on the patient. In some embodiments, they are administered within about 60, 30, 15, 10, 5, or 1 minute of one another. In some embodiments, the administrations of the agents are spaced sufficiently closely together such that a combinatorial (*e.g.*, a synergistic) effect is achieved.

**[00644]** *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.

**[00645]** *Antigens of interest or desired antigens*: As used herein, the terms "antigens of interest" or "desired antigens" include those proteins and other biomolecules provided herein that are immunospecifically bound by the antibodies and fragments, mutants, variants, and alterations thereof described herein. Examples of antigens of interest include, but are not limited to, insulin, insulin-like growth factor, hGH, tPA, cytokines, such as interleukins (IL), e.g., IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, interferon

(IFN) alpha, IFN beta, IFN gamma, IFN omega or IFN tau, tumor necrosis factor (TNF), such as TNF alpha and TNF beta, TNF gamma, TRAIL; G-CSF, GM-CSF, M-CSF, MCP-1 and VEGF.

**[00646]** *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).

**[00647]** *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.

**[00648]** *Biocompatible*: As used herein, the term "biocompatible" means compatible with living cells, tissues, organs or systems posing little to no risk of injury, toxicity or rejection by the immune system.

**[00649]** *Biodegradable*: As used herein, the term "biodegradable" means capable of being broken down into innocuous products by the action of living things.

**[00650]** *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 effect on that organism, is considered to be biologically active. In particular embodiments, a polynucleotide of the present invention may be considered biologically active if even a portion of the polynucleotide is biologically active or mimics an activity considered biologically relevant.

**[00651]** *Chemical terms*: The following provides the definition of various chemical terms from "acyl" to "thiol."

**[00652]** The term "acyl," as used herein, represents a hydrogen or an alkyl group (e.g., a haloalkyl group), as defined herein, that is attached to the parent molecular group through a carbonyl group, as defined herein, and is exemplified by formyl (i.e., a carboxyaldehyde group), acetyl, trifluoroacetyl, propionyl, butanoyl and the like. Exemplary unsubstituted acyl groups include from 1 to 7, from 1 to 11, or from 1 to 21 carbons. In some embodiments, the alkyl group is further substituted with 1, 2, 3, or 4 substituents as described herein.

**[00653]** The term "acylamino," as used herein, represents an acyl group, as defined herein, attached to the parent molecular group though an amino group, as defined herein (i.e., $-N(R^{N1})-C(O)-R$, where R is H or an optionally substituted $C_{1-6}$, $C_{1-10}$, or $C_{1-20}$ alkyl group (e.g., haloalkyl) and $R^{N1}$ is as defined herein). Exemplary unsubstituted acylamino groups include from 1 to 41 carbons (e.g., from 1 to 7, from 1 to 13, from 1 to 21, from 2 to 7, from 2 to 13, from 2 to 21, or from 2 to 41 carbons). In some embodiments, the alkyl group is further substituted with 1, 2, 3, or 4 substituents as described herein, and/or the amino group is $-NH_2$ or $-NHR^{N1}$, wherein $R^{N1}$ is, independently, OH, $NO_2$, $NH_2$, $NR^{N2}{}_2$, $SO_2OR^{N2}$, $SO_2R^{N2}$, $SOR^{N2}$, alkyl, aryl, acyl (e.g., acetyl, trifluoroacetyl, or others described herein), or alkoxycarbonylalkyl, and each $R^{N2}$ can be H, alkyl, or aryl.

**[00654]** The term "acylaminoalkyl," as used herein, represents an acyl group, as defined herein, attached to an amino group that is in turn attached to the parent molecular group though an alkyl group, as defined herein (i.e., $-alkyl-N(R^{N1})-C(O)-R$, where R is H or an optionally substituted $C_{1-6}$, $C_{1-10}$, or $C_{1-20}$ alkyl group (e.g., haloalkyl) and $R^{N1}$ is as defined herein). Exemplary unsubstituted acylamino groups include from 1 to 41 carbons (e.g., from 1 to 7, from 1 to 13, from 1 to 21, from 2 to 7, from 2 to 13, from 2 to 21, or from 2 to 41 carbons). In some embodiments, the alkyl group is further substituted with 1, 2, 3, or 4 substituents as described herein, and/or the amino group is $-NH_2$ or $-NHR^{N1}$, wherein $R^{N1}$ is, independently, OH, $NO_2$, $NH_2$, $NR^{N2}{}_2$, $SO_2OR^{N2}$, $SO_2R^{N2}$, $SOR^{N2}$, alkyl, aryl, acyl (e.g., acetyl, trifluoroacetyl, or others described herein), or alkoxycarbonylalkyl, and each $R^{N2}$ can be H, alkyl, or aryl.

**[00655]** The term "acyloxy," as used herein, represents an acyl group, as defined herein, attached to the parent molecular group though an oxygen atom (i.e., $-O-C(O)-R$, where R is H or an optionally substituted $C_{1-6}$, $C_{1-10}$, or $C_{1-20}$ alkyl group). Exemplary unsubstituted acyloxy groups include from 1 to 21 carbons (e.g., from 1 to 7 or from 1 to 11 carbons). In some embodiments, the alkyl group is further substituted with 1, 2, 3, or 4 substituents as described herein.

**[00656]** The term "acyloxyalkyl," as used herein, represents an acyl group, as defined herein, attached to an oxygen atom that in turn is attached to the parent molecular group though an alkyl group (i.e., –alkyl-O-C(O)-R, where R is H or an optionally substituted $C_{1-6}$, $C_{1-10}$, or $C_{1-20}$ alkyl group). Exemplary unsubstituted acyloxyalkyl groups include from 1 to 21 carbons (e.g., from 1 to 7 or from 1 to 11 carbons). In some embodiments, the alkyl group is, independently, further substituted with 1, 2, 3, or 4 substituents as described herein.

**[00657]** The term "alkaryl," as used herein, represents an aryl group, as defined herein, attached to the parent molecular group through an alkylene group, as defined herein. Exemplary unsubstituted alkaryl groups are from 7 to 30 carbons (e.g., from 7 to 16 or from 7 to 20 carbons, such as $C_{1-6}$ alk-$C_{6-10}$ aryl, $C_{1-10}$ alk-$C_{6-10}$ aryl, or $C_{1-20}$ alk-$C_{6-10}$ aryl). In some embodiments, the alkylene and the aryl each can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein for the respective groups. Other groups preceded by the prefix "alk-" are defined in the same manner, where "alk" refers to a $C_{1-6}$ alkylene, unless otherwise noted, and the attached chemical structure is as defined herein.

**[00658]** The term "alkcycloalkyl" represents a cycloalkyl group, as defined herein, attached to the parent molecular group through an alkylene group, as defined herein (e.g., an alkylene group of from 1 to 4, from 1 to 6, from 1 to 10, or form 1 to 20 carbons). In some embodiments, the alkylene and the cycloalkyl each can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein for the respective group.

**[00659]** The term "alkenyl," as used herein, represents monovalent straight or branched chain groups of, unless otherwise specified, from 2 to 20 carbons (e.g., from 2 to 6 or from 2 to 10 carbons) containing one or more carbon-carbon double bonds and is exemplified by ethenyl, 1-propenyl, 2-propenyl, 2-methyl-1-propenyl, 1-butenyl, 2-butenyl, and the like. Alkenyls include both cis and trans isomers. Alkenyl groups may be optionally substituted with 1, 2, 3, or 4 substituent groups that are selected, independently, from amino, aryl, cycloalkyl, or heterocyclyl (e.g., heteroaryl), as defined herein, or any of the exemplary alkyl substituent groups described herein.

**[00660]** The term "alkenyloxy" represents a chemical substituent of formula –OR, where R is a $C_{2-20}$ alkenyl group (e.g., $C_{2-6}$ or $C_{2-10}$ alkenyl), unless otherwise specified. Exemplary alkenyloxy groups include ethenyloxy, propenyloxy, and the like. In some embodiments, the alkenyl group can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein (e.g., a hydroxy group).

**[00661]** The term "alkheteroaryl" refers to a heteroaryl group, as defined herein, attached to the parent molecular group through an alkylene group, as defined herein. Exemplary unsubstituted alkheteroaryl groups are from 2 to 32 carbons (e.g., from 2 to 22, from 2 to 18, from 2 to 17, from 2 to 16, from 3 to 15, from 2 to 14, from 2 to 13, or from 2 to 12 carbons, such as $C_{1-6}$ alk-$C_{1-12}$ heteroaryl, $C_{1-10}$ alk-$C_{1-12}$ heteroaryl, or $C_{1-20}$ alk-$C_{1-12}$ heteroaryl). In some embodiments, the alkylene and the heteroaryl each can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein for the respective group. Alkheteroaryl groups are a subset of alkheterocyclyl groups.

**[00662]** The term "alkheterocyclyl" represents a heterocyclyl group, as defined herein, attached to the parent molecular group through an alkylene group, as defined herein. Exemplary unsubstituted alkheterocyclyl groups are from 2 to 32 carbons (e.g., from 2 to 22, from 2 to 18, from 2 to 17, from 2 to 16, from 3 to 15, from 2 to 14, from 2 to 13, or from 2 to 12 carbons, such as $C_{1-6}$ alk-$C_{1-12}$ heterocyclyl, $C_{1-10}$ alk-$C_{1-12}$ heterocyclyl, or $C_{1-20}$ alk-$C_{1-12}$ heterocyclyl). In some embodiments, the alkylene and the heterocyclyl each can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein for the respective group.

**[00663]** The term "alkoxy" represents a chemical substituent of formula –OR, where R is a $C_{1-20}$ alkyl group (e.g., $C_{1-6}$ or $C_{1-10}$ alkyl), unless otherwise specified. Exemplary alkoxy groups include methoxy, ethoxy, propoxy (e.g., n-propoxy and isopropoxy), t-butoxy, and the like. In some embodiments, the alkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein (e.g., hydroxy or alkoxy).

**[00664]** The term "alkoxyalkoxy" represents an alkoxy group that is substituted with an alkoxy group. Exemplary unsubstituted alkoxyalkoxy groups include between 2 to 40 carbons (e.g., from 2 to 12 or from 2 to 20 carbons, such as $C_{1-6}$ alkoxy-$C_{1-6}$ alkoxy, $C_{1-10}$ alkoxy-$C_{1-10}$ alkoxy, or $C_{1-20}$ alkoxy-$C_{1-20}$ alkoxy). In some embodiments, the each alkoxy group can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein.

**[00665]** The term "alkoxyalkyl" represents an alkyl group that is substituted with an alkoxy group. Exemplary unsubstituted alkoxyalkyl groups include between 2 to 40 carbons (e.g., from 2 to 12 or from 2 to 20 carbons, such as $C_{1-6}$ alkoxy-$C_{1-6}$ alkyl, $C_{1-10}$ alkoxy-$C_{1-10}$ alkyl, or $C_{1-20}$ alkoxy-$C_{1-20}$ alkyl). In some embodiments, the alkyl and the alkoxy each can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein for the respective group.

**[00666]** The term "alkoxycarbonyl," as used herein, represents an alkoxy, as defined herein, attached to the parent molecular group through a carbonyl atom (e.g., -C(O)-OR, where R is H or an

optionally substituted $C_{1-6}$, $C_{1-10}$, or $C_{1-20}$ alkyl group). Exemplary unsubstituted alkoxycarbonyl include from 1 to 21 carbons (e.g., from 1 to 11 or from 1 to 7 carbons). In some embodiments, the alkoxy group is further substituted with 1, 2, 3, or 4 substituents as described herein.

**[00667]** The term "alkoxycarbonylacyl," as used herein, represents an acyl group, as defined herein, that is substituted with an alkoxycarbonyl group, as defined herein (e.g., -C(O) -alkyl-C(O)-OR, where R is an optionally substituted $C_{1-6}$, $C_{1-10}$, or $C_{1-20}$ alkyl group). Exemplary unsubstituted alkoxycarbonylacyl include from 3 to 41 carbons (e.g., from 3 to 10, from 3 to 13, from 3 to 17, from 3 to 21, or from 3 to 31 carbons, such as $C_{1-6}$ alkoxycarbonyl-$C_{1-6}$ acyl, $C_{1-10}$ alkoxycarbonyl-$C_{1-10}$ acyl, or $C_{1-20}$ alkoxycarbonyl-$C_{1-20}$ acyl). In some embodiments, each alkoxy and alkyl group is further independently substituted with 1, 2, 3, or 4 substituents, as described herein (e.g., a hydroxy group) for each group.

**[00668]** The term "alkoxycarbonylalkoxy," as used herein, represents an alkoxy group, as defined herein, that is substituted with an alkoxycarbonyl group, as defined herein (e.g., -O-alkyl-C(O)-OR, where R is an optionally substituted $C_{1-6}$, $C_{1-10}$, or $C_{1-20}$ alkyl group). Exemplary unsubstituted alkoxycarbonylalkoxy include from 3 to 41 carbons (e.g., from 3 to 10, from 3 to 13, from 3 to 17, from 3 to 21, or from 3 to 31 carbons, such as $C_{1-6}$ alkoxycarbonyl-$C_{1-6}$ alkoxy, $C_{1-10}$ alkoxycarbonyl-$C_{1-10}$ alkoxy, or $C_{1-20}$ alkoxycarbonyl-$C_{1-20}$ alkoxy). In some embodiments, each alkoxy group is further independently substituted with 1, 2, 3, or 4 substituents, as described herein (e.g., a hydroxy group).

**[00669]** The term "alkoxycarbonylalkyl," as used herein, represents an alkyl group, as defined herein, that is substituted with an alkoxycarbonyl group, as defined herein (e.g., -alkyl-C(O)-OR, where R is an optionally substituted $C_{1-20}$, $C_{1-10}$, or $C_{1-6}$ alkyl group). Exemplary unsubstituted alkoxycarbonylalkyl include from 3 to 41 carbons (e.g., from 3 to 10, from 3 to 13, from 3 to 17, from 3 to 21, or from 3 to 31 carbons, such as $C_{1-6}$ alkoxycarbonyl-$C_{1-6}$ alkyl, $C_{1-10}$ alkoxycarbonyl-$C_{1-10}$ alkyl, or $C_{1-20}$ alkoxycarbonyl-$C_{1-20}$ alkyl). In some embodiments, each alkyl and alkoxy group is further independently substituted with 1, 2, 3, or 4 substituents as described herein (e.g., a hydroxy group).

**[00670]** The term "alkoxycarbonylalkenyl," as used herein, represents an alkenyl group, as defined herein, that is substituted with an alkoxycarbonyl group, as defined herein (e.g., -alkenyl-C(O)-OR, where R is an optionally substituted $C_{1-20}$, $C_{1-10}$, or $C_{1-6}$ alkyl group). Exemplary unsubstituted alkoxycarbonylalkenyl include from 4 to 41 carbons (e.g., from 4 to 10, from 4 to 13, from 4 to 17,

from 4 to 21, or from 4 to 31 carbons, such as $C_{1-6}$ alkoxycarbonyl-$C_{2-6}$ alkenyl, $C_{1-10}$ alkoxycarbonyl-$C_{2-10}$ alkenyl, or $C_{1-20}$ alkoxycarbonyl-$C_{2-20}$ alkenyl). In some embodiments, each alkyl, alkenyl, and alkoxy group is further independently substituted with 1, 2, 3, or 4 substituents as described herein (e.g., a hydroxy group).

**[00671]** The term "alkoxycarbonylalkynyl," as used herein, represents an alkynyl group, as defined herein, that is substituted with an alkoxycarbonyl group, as defined herein (e.g., -alkynyl-C(O)-OR, where R is an optionally substituted $C_{1-20}$, $C_{1-10}$, or $C_{1-6}$ alkyl group). Exemplary unsubstituted alkoxycarbonylalkynyl include from 4 to 41 carbons (e.g., from 4 to 10, from 4 to 13, from 4 to 17, from 4 to 21, or from 4 to 31 carbons, such as $C_{1-6}$ alkoxycarbonyl-$C_{2-6}$ alkynyl, $C_{1-10}$ alkoxycarbonyl-$C_{2-10}$ alkynyl, or $C_{1-20}$ alkoxycarbonyl-$C_{2-20}$ alkynyl). In some embodiments, each alkyl, alkynyl, and alkoxy group is further independently substituted with 1, 2, 3, or 4 substituents as described herein (e.g., a hydroxy group).

**[00672]** The term "alkyl," as used herein, is inclusive of both straight chain and branched chain saturated groups from 1 to 20 carbons (e.g., from 1 to 10 or from 1 to 6), unless otherwise specified. Alkyl groups are exemplified by methyl, ethyl, n- and iso-propyl, n-, sec-, iso- and tert-butyl, neopentyl, and the like, and may be optionally substituted with one, two, three, or, in the case of alkyl groups of two carbons or more, four substituents independently selected from the group consisting of: (1) $C_{1-6}$ alkoxy; (2) $C_{1-6}$ alkylsulfinyl; (3) amino, as defined herein (e.g., unsubstituted amino (i.e., $-NH_2$) or a substituted amino (i.e., $-N(R^{N1})_2$, where $R^{N1}$ is as defined for amino); (4) $C_{6-10}$ aryl-$C_{1-6}$ alkoxy; (5) azido; (6) halo; (7) ($C_{2-9}$ heterocyclyl)oxy; (8) hydroxy, optionally substituted with an *O*-protecting group; (9) nitro; (10) oxo (e.g., carboxyaldehyde or acyl); (11) $C_{1-7}$ spirocyclyl; (12) thioalkoxy; (13) thiol; (14) $-CO_2R^{A'}$, optionally substituted with an *O*-protecting group and where $R^{A'}$ is selected from the group consisting of (a) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c) $C_{6-10}$ aryl, (d) hydrogen, (e) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f) amino-$C_{1-20}$ alkyl, (g) polyethylene glycol of $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h) amino-polyethylene glycol of $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; (15) $-C(O)NR^{B'}R^{C'}$, where each of $R^{B'}$

and $R^{C'}$ is, independently, selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (16) -$SO_2R^{D'}$, where $R^{D'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) $C_{1-6}$ alk-$C_{6-10}$ aryl, and (d) hydroxy; (17) -$SO_2NR^{E'}R^{F'}$, where each of $R^{E'}$ and $R^{F'}$ is, independently, selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (18) -$C(O)R^{G'}$, where $R^{G'}$ is selected from the group consisting of (a) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c) $C_{6-10}$ aryl, (d) hydrogen, (e) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f) amino-$C_{1-20}$ alkyl, (g) polyethylene glycol of -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h) amino-polyethylene glycol of -$NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; (19) -$NR^{H'}C(O)R^{I'}$, wherein $R^{H'}$ is selected from the group consisting of (a1) hydrogen and (b1) $C_{1-6}$ alkyl, and $R^{I'}$ is selected from the group consisting of (a2) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b2) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c2) $C_{6-10}$ aryl, (d2) hydrogen, (e2) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f2) amino-$C_{1-20}$ alkyl, (g2) polyethylene glycol of -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h2) amino-polyethylene glycol of -$NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; (20) -$NR^{J'}C(O)OR^{K'}$, wherein $R^{J'}$ is selected from the group consisting of (a1) hydrogen and (b1) $C_{1-6}$ alkyl, and $R^{K'}$ is selected from the group consisting of (a2) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b2) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c2) $C_{6-10}$ aryl, (d2) hydrogen, (e2) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f2) amino-$C_{1-20}$ alkyl, (g2) polyethylene glycol of -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h2) amino-polyethylene glycol of -$NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer

from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; and (21) amidine. In some embodiments, each of these groups can be further substituted as described herein. For example, the alkylene group of a $C_1$-alkaryl can be further substituted with an oxo group to afford the respective aryloyl substituent.

**[00673]** The term "alkylene" and the prefix "alk-," as used herein, represent a saturated divalent hydrocarbon group derived from a straight or branched chain saturated hydrocarbon by the removal of two hydrogen atoms, and is exemplified by methylene, ethylene, isopropylene, and the like. The term "$C_{x-y}$ alkylene" and the prefix "$C_{x-y}$ alk-" represent alkylene groups having between x and y carbons. Exemplary values for x are 1, 2, 3, 4, 5, and 6, and exemplary values for y are 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, or 20 (e.g., $C_{1-6}$, $C_{1-10}$, $C_{2-20}$, $C_{2-6}$, $C_{2-10}$, or $C_{2-20}$ alkylene). In some embodiments, the alkylene can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein for an alkyl group.

**[00674]** The term "alkylsulfinyl," as used herein, represents an alkyl group attached to the parent molecular group through an -S(O)- group. Exemplary unsubstituted alkylsulfinyl groups are from 1 to 6, from 1 to 10, or from 1 to 20 carbons. In some embodiments, the alkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein.

**[00675]** The term "alkylsulfinylalkyl," as used herein, represents an alkyl group, as defined herein, substituted by an alkylsulfinyl group. Exemplary unsubstituted alkylsulfinylalkyl groups are from 2 to 12, from 2 to 20, or from 2 to 40 carbons. In some embodiments, each alkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein.

**[00676]** The term "alkynyl," as used herein, represents monovalent straight or branched chain groups from 2 to 20 carbon atoms (e.g., from 2 to 4, from 2 to 6, or from 2 to 10 carbons) containing a carbon-carbon triple bond and is exemplified by ethynyl, 1-propynyl, and the like. Alkynyl groups may be optionally substituted with 1, 2, 3, or 4 substituent groups that are selected, independently, from aryl, cycloalkyl, or heterocyclyl (e.g., heteroaryl), as defined herein, or any of the exemplary alkyl substituent groups described herein.

**[00677]** The term "alkynyloxy" represents a chemical substituent of formula –OR, where R is a $C_{2-20}$ alkynyl group (e.g., $C_{2-6}$ or $C_{2-10}$ alkynyl), unless otherwise specified. Exemplary alkynyloxy groups include ethynyloxy, propynyloxy, and the like. In some embodiments, the alkynyl group can be further substituted with 1, 2, 3, or 4 substituent groups as defined herein (e.g., a hydroxy group).

**[00678]** The term "amidine," as used herein, represents a $-C(=NH)NH_2$ group.

**[00679]** The term "amino," as used herein, represents $-N(R^{N1})_2$, wherein each $R^{N1}$ is, independently, H, OH, $NO_2$, $N(R^{N2})_2$, $SO_2OR^{N2}$, $SO_2R^{N2}$, $SOR^{N2}$, an *N*-protecting group, alkyl, alkenyl, alkynyl, alkoxy, aryl, alkaryl, cycloalkyl, alkcycloalkyl, carboxyalkyl (e.g., optionally substituted with an *O*-protecting group, such as optionally substituted arylalkoxycarbonyl groups or any described herein), sulfoalkyl, acyl (e.g., acetyl, trifluoroacetyl, or others described herein), alkoxycarbonylalkyl (e.g., optionally substituted with an *O*-protecting group, such as optionally substituted arylalkoxycarbonyl groups or any described herein), heterocyclyl (e.g., heteroaryl), or alkheterocyclyl (e.g., alkheteroaryl), wherein each of these recited $R^{N1}$ groups can be optionally substituted, as defined herein for each group; or two $R^{N1}$ combine to form a heterocyclyl or an *N*-protecting group, and wherein each $R^{N2}$ is, independently, H, alkyl, or aryl. The amino groups of the invention can be an unsubstituted amino (i.e., $-NH_2$) or a substituted amino (i.e., $-N(R^{N1})_2$). In a preferred embodiment, amino is $-NH_2$ or $-NHR^{N1}$, wherein $R^{N1}$ is, independently, OH, $NO_2$, $NH_2$, $NR^{N2}{}_2$, $SO_2OR^{N2}$, $SO_2R^{N2}$, $SOR^{N2}$, alkyl, carboxyalkyl, sulfoalkyl, acyl (e.g., acetyl, trifluoroacetyl, or others described herein), alkoxycarbonylalkyl (e.g., t-butoxycarbonylalkyl) or aryl, and each $R^{N2}$ can be H, $C_{1\text{-}20}$ alkyl (e.g., $C_{1\text{-}6}$ alkyl), or $C_{6\text{-}10}$ aryl.

**[00680]** The term "amino acid," as described herein, refers to a molecule having a side chain, an amino group, and an acid group (e.g., a carboxy group of $-CO_2H$ or a sulfo group of $-SO_3H$), wherein the amino acid is attached to the parent molecular group by the side chain, amino group, or acid group (e.g., the side chain). In some embodiments, the amino acid is attached to the parent molecular group by a carbonyl group, where the side chain or amino group is attached to the carbonyl group. Exemplary side chains include an optionally substituted alkyl, aryl, heterocyclyl, alkaryl, alkheterocyclyl, aminoalkyl, carbamoylalkyl, and carboxyalkyl. Exemplary amino acids include alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, hydroxynorvaline, isoleucine, leucine, lysine, methionine, norvaline, ornithine, phenylalanine, proline, pyrrolysine, selenocysteine, serine, taurine, threonine, tryptophan, tyrosine, and valine. Amino acid groups may be optionally substituted with one, two, three, or, in the case of amino acid groups of two carbons or more, four substituents independently selected from the group consisting of: (1) $C_{1\text{-}6}$ alkoxy; (2) $C_{1\text{-}6}$ alkylsulfinyl; (3) amino, as defined herein (e.g., unsubstituted amino (i.e., $-NH_2$) or a substituted amino (i.e., $-N(R^{N1})_2$, where $R^{N1}$ is as defined for amino); (4) $C_{6\text{-}10}$ aryl-$C_{1\text{-}6}$ alkoxy; (5) azido; (6) halo; (7) ($C_{2\text{-}9}$ heterocyclyl)oxy; (8) hydroxy; (9) nitro; (10) oxo (e.g.,

carboxyaldehyde or acyl); (11) $C_{1-7}$ spirocyclyl; (12) thioalkoxy; (13) thiol; (14) $-CO_2R^{A'}$, where $R^{A'}$ is selected from the group consisting of (a) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c) $C_{6-10}$ aryl, (d) hydrogen, (e) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f) amino-$C_{1-20}$ alkyl, (g) polyethylene glycol of $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h) amino-polyethylene glycol of $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; (15) $-C(O)NR^{B'}R^{C'}$, where each of $R^{B'}$ and $R^{C'}$ is, independently, selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (16) $-SO_2R^{D'}$, where $R^{D'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) $C_{1-6}$ alk-$C_{6-10}$ aryl, and (d) hydroxy; (17) $-SO_2NR^{E'}R^{F'}$, where each of $R^{E'}$ and $R^{F'}$ is, independently, selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (18) $-C(O)R^{G'}$, where $R^{G'}$ is selected from the group consisting of (a) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c) $C_{6-10}$ aryl, (d) hydrogen, (e) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f) amino-$C_{1-20}$ alkyl, (g) polyethylene glycol of $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h) amino-polyethylene glycol of $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; (19) $-NR^{H'}C(O)R^{I'}$, wherein $R^{H'}$ is selected from the group consisting of (a1) hydrogen and (b1) $C_{1-6}$ alkyl, and $R^{I'}$ is selected from the group consisting of (a2) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b2) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c2) $C_{6-10}$ aryl, (d2) hydrogen, (e2) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f2) amino-$C_{1-20}$ alkyl, (g2) polyethylene glycol of $-(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h2) amino-polyethylene glycol of $-NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3,

independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; (20) -$NR^{J'}C(O)OR^{K'}$, wherein $R^{J'}$ is selected from the group consisting of (a1) hydrogen and (b1) $C_{1-6}$ alkyl, and $R^{K'}$ is selected from the group consisting of (a2) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl), (b2) $C_{2-20}$ alkenyl (e.g., $C_{2-6}$ alkenyl), (c2) $C_{6-10}$ aryl, (d2) hydrogen, (e2) $C_{1-6}$ alk-$C_{6-10}$ aryl, (f2) amino-$C_{1-20}$ alkyl, (g2) polyethylene glycol of -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and R' is H or $C_{1-20}$ alkyl, and (h2) amino-polyethylene glycol of -$NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10 (e.g., from 1 to 6 or from 1 to 4), each of s2 and s3, independently, is an integer from 0 to 10 (e.g., from 0 to 4, from 0 to 6, from 1 to 4, from 1 to 6, or from 1 to 10), and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl; and (21) amidine. In some embodiments, each of these groups can be further substituted as described herein.

**[00681]** The term "aminoalkoxy," as used herein, represents an alkoxy group, as defined herein, substituted by an amino group, as defined herein. The alkyl and amino each can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for the respective group (e.g., $CO_2R^{A'}$, where $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl, e.g., carboxy).

**[00682]** The term "aminoalkyl," as used herein, represents an alkyl group, as defined herein, substituted by an amino group, as defined herein. The alkyl and amino each can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for the respective group (e.g., $CO_2R^{A'}$, where $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl, e.g., carboxy, and/or an *N*-protecting group).

**[00683]** The term "aminoalkenyl," as used herein, represents an alkenyl group, as defined herein, substituted by an amino group, as defined herein. The alkenyl and amino each can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for the respective group (e.g., $CO_2R^{A'}$, where $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl, e.g., carboxy, and/or an *N*-protecting group).

**[00684]** The term "aminoalkynyl," as used herein, represents an alkynyl group, as defined herein, substituted by an amino group, as defined herein. The alkynyl and amino each can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for the respective group (e.g.,

$CO_2R^{A'}$, where $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl, e.g., carboxy, and/or an *N*-protecting group).

**[00685]** The term "aryl," as used herein, represents a mono-, bicyclic, or multicyclic carbocyclic ring system having one or two aromatic rings and is exemplified by phenyl, naphthyl, 1,2-dihydronaphthyl, 1,2,3,4-tetrahydronaphthyl, anthracenyl, phenanthrenyl, fluorenyl, indanyl, indenyl, and the like, and may be optionally substituted with 1, 2, 3, 4, or 5 substituents independently selected from the group consisting of: (1) $C_{1-7}$ acyl (e.g., carboxyaldehyde); (2) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl, $C_{1-6}$ alkoxy-$C_{1-6}$ alkyl, $C_{1-6}$ alkylsulfinyl-$C_{1-6}$ alkyl, amino-$C_{1-6}$ alkyl, azido-$C_{1-6}$ alkyl, (carboxyaldehyde)-$C_{1-6}$ alkyl, halo-$C_{1-6}$ alkyl (e.g., perfluoroalkyl), hydroxy-$C_{1-6}$ alkyl, nitro-$C_{1-6}$ alkyl, or $C_{1-6}$ thioalkoxy-$C_{1-6}$ alkyl); (3) $C_{1-20}$ alkoxy (e.g., $C_{1-6}$ alkoxy, such as perfluoroalkoxy); (4) $C_{1-6}$ alkylsulfinyl; (5) $C_{6-10}$ aryl; (6) amino; (7) $C_{1-6}$ alk-$C_{6-10}$ aryl; (8) azido; (9) $C_{3-8}$ cycloalkyl; (10) $C_{1-6}$ alk-$C_{3-8}$ cycloalkyl; (11) halo; (12) $C_{1-12}$ heterocyclyl (e.g., $C_{1-12}$ heteroaryl); (13) ($C_{1-12}$ heterocyclyl)oxy; (14) hydroxy; (15) nitro; (16) $C_{1-20}$ thioalkoxy (e.g., $C_{1-6}$ thioalkoxy); (17) $-(CH_2)_qCO_2R^{A'}$, where q is an integer from zero to four, and $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (18) $-(CH_2)_qCONR^{B'}R^{C'}$, where q is an integer from zero to four and where $R^{B'}$ and $R^{C'}$ are independently selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (19) $-(CH_2)_qSO_2R^{D'}$, where q is an integer from zero to four and where $R^{D'}$ is selected from the group consisting of (a) alkyl, (b) $C_{6-10}$ aryl, and (c) alk-$C_{6-10}$ aryl; (20) $-(CH_2)_qSO_2NR^{E'}R^{F'}$, where q is an integer from zero to four and where each of $R^{E'}$ and $R^{F'}$ is, independently, selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (21) thiol; (22) $C_{6-10}$ aryloxy; (23) $C_{3-8}$ cycloalkoxy; (24) $C_{6-10}$ aryl-$C_{1-6}$ alkoxy; (25) $C_{1-6}$ alk-$C_{1-12}$ heterocyclyl (e.g., $C_{1-6}$ alk-$C_{1-12}$ heteroaryl); (26) $C_{2-20}$ alkenyl; and (27) $C_{2-20}$ alkynyl. In some embodiments, each of these groups can be further substituted as described herein. For example, the alkylene group of a $C_1$-alkaryl or a $C_1$-alkheterocyclyl can be further substituted with an oxo group to afford the respective aryloyl and (heterocyclyl)oyl substituent group.

**[00686]** The term "arylalkoxy," as used herein, represents an alkaryl group, as defined herein, attached to the parent molecular group through an oxygen atom. Exemplary unsubstituted arylalkoxy groups include from 7 to 30 carbons (e.g., from 7 to 16 or from 7 to 20 carbons, such as $C_{6-10}$ aryl-$C_{1-6}$ alkoxy, $C_{6-10}$ aryl-$C_{1-10}$ alkoxy, or $C_{6-10}$ aryl-$C_{1-20}$ alkoxy). In some embodiments, the arylalkoxy group can be substituted with 1, 2, 3, or 4 substituents as defined herein

**[00687]** The term "arylalkoxycarbonyl," as used herein, represents an arylalkoxy group, as defined herein, attached to the parent molecular group through a carbonyl (e.g., -C(O)-O-alkyl-aryl). Exemplary unsubstituted arylalkoxy groups include from 8 to 31 carbons (e.g., from 8 to 17 or from 8 to 21 carbons, such as $C_{6-10}$ aryl-$C_{1-6}$ alkoxy-carbonyl, $C_{6-10}$ aryl-$C_{1-10}$ alkoxy-carbonyl, or $C_{6-10}$ aryl-$C_{1-20}$ alkoxy-carbonyl). In some embodiments, the arylalkoxycarbonyl group can be substituted with 1, 2, 3, or 4 substituents as defined herein.

**[00688]** The term "aryloxy" represents a chemical substituent of formula –OR′, where R′ is an aryl group of 6 to 18 carbons, unless otherwise specified. In some embodiments, the aryl group can be substituted with 1, 2, 3, or 4 substituents as defined herein.

**[00689]** The term "aryloyl," as used herein, represents an aryl group, as defined herein, that is attached to the parent molecular group through a carbonyl group. Exemplary unsubstituted aryloyl groups are of 7 to 11 carbons. In some embodiments, the aryl group can be substituted with 1, 2, 3, or 4 substituents as defined herein.

**[00690]** The term "azido" represents an $–N_3$ group, which can also be represented as –N=N=N.

**[00691]** The term "bicyclic," as used herein, refer to a structure having two rings, which may be aromatic or non-aromatic. Bicyclic structures include spirocyclyl groups, as defined herein, and two rings that share one or more bridges, where such bridges can include one atom or a chain including two, three, or more atoms. Exemplary bicyclic groups include a bicyclic carbocyclyl group, where the first and second rings are carbocyclyl groups, as defined herein; a bicyclic aryl groups, where the first and second rings are aryl groups, as defined herein; bicyclic heterocyclyl groups, where the first ring is a heterocyclyl group and the second ring is a carbocyclyl (e.g., aryl) or heterocyclyl (e.g., heteroaryl) group; and bicyclic heteroaryl groups, where the first ring is a heteroaryl group and the second ring is a carbocyclyl (e.g., aryl) or heterocyclyl (e.g., heteroaryl) group. In some embodiments, the bicyclic group can be substituted with 1, 2, 3, or 4 substituents as defined herein for cycloalkyl, heterocyclyl, and aryl groups.

**[00692]** The term "boranyl," as used herein, represents $–B(R^{B1})_3$, where each $R^{B1}$ is, independently, selected from the group consisting of H and optionally substituted alkyl. In some embodiments, the boranyl group can be substituted with 1, 2, 3, or 4 substituents as defined herein for alkyl.

**[00693]** The terms "carbocyclic" and "carbocyclyl," as used herein, refer to an optionally substituted $C_{3-12}$ monocyclic, bicyclic, or tricyclic structure in which the rings, which may be

aromatic or non-aromatic, are formed by carbon atoms. Carbocyclic structures include cycloalkyl, cycloalkenyl, and aryl groups.

**[00694]** The term "carbamoyl," as used herein, represents $-C(O)-N(R^{N1})_2$, where the meaning of each $R^{N1}$ is found in the definition of "amino" provided herein.

**[00695]** The term "carbamoylalkyl," as used herein, represents an alkyl group, as defined herein, substituted by a carbamoyl group, as defined herein. The alkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein.

**[00696]** The term "carbamyl," as used herein, refers to a carbamate group having the structure $-NR^{N1}C(=O)OR$ or $-OC(=O)N(R^{N1})_2$, where the meaning of each $R^{N1}$ is found in the definition of "amino" provided herein, and R is alkyl, cycloalkyl , alkcycloalkyl, aryl, alkaryl, heterocyclyl (e.g., heteroaryl), or alkheterocyclyl (e.g., alkheteroaryl), as defined herein.

**[00697]** The term "carbonyl," as used herein, represents a C(O) group, which can also be represented as C=O.

**[00698]** The term "carboxyaldehyde" represents an acyl group having the structure –CHO.

**[00699]** The term "carboxy," as used herein, means $-CO_2H$.

**[00700]** The term "carboxyalkoxy," as used herein, represents an alkoxy group, as defined herein, substituted by a carboxy group, as defined herein. The alkoxy group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for the alkyl group, and the carboxy group can be optionally substituted with one or more *O*-protecting groups.

**[00701]** The term "carboxyalkyl," as used herein, represents an alkyl group, as defined herein, substituted by a carboxy group, as defined herein. The alkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein, and the carboxy group can be optionally substituted with one or more *O*-protecting groups.

**[00702]** The term "carboxyaminoalkyl," as used herein, represents an aminoalkyl group, as defined herein, substituted by a carboxy, as defined herein. The carboxy, alkyl, and amino each can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for the respective group (e.g., $CO_2R^{A'}$, where $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl, e.g., carboxy, and/or an *N*-protecting group, and/or an *O*-protecting group).

**[00703]** The term "cyano," as used herein, represents an –CN group.

**[00704]** The term "cycloalkoxy" represents a chemical substituent of formula –OR, where R is a $C_{3-8}$ cycloalkyl group, as defined herein, unless otherwise specified. The cycloalkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein. Exemplary unsubstituted cycloalkoxy groups are from 3 to 8 carbons. In some embodiment, the cycloalkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein.

**[00705]** The term "cycloalkyl," as used herein represents a monovalent saturated or unsaturated non-aromatic cyclic hydrocarbon group from three to eight carbons, unless otherwise specified, and is exemplified by cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, bicycle heptyl, and the like. When the cycloalkyl group includes one carbon-carbon double bond, the cycloalkyl group can be referred to as a "cycloalkenyl" group. Exemplary cycloalkenyl groups include cyclopentenyl, cyclohexenyl, and the like. The cycloalkyl groups of this invention can be optionally substituted with: (1) $C_{1-7}$ acyl (e.g., carboxyaldehyde); (2) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl, $C_{1-6}$ alkoxy-$C_{1-6}$ alkyl, $C_{1-6}$ alkylsulfinyl-$C_{1-6}$ alkyl, amino-$C_{1-6}$ alkyl, azido-$C_{1-6}$ alkyl, (carboxyaldehyde)-$C_{1-6}$ alkyl, halo-$C_{1-6}$ alkyl (e.g., perfluoroalkyl), hydroxy-$C_{1-6}$ alkyl, nitro-$C_{1-6}$ alkyl, or $C_{1-6}$ thioalkoxy-$C_{1-6}$ alkyl); (3) $C_{1-20}$ alkoxy (e.g., $C_{1-6}$ alkoxy, such as perfluoroalkoxy); (4) $C_{1-6}$ alkylsulfinyl; (5) $C_{6-10}$ aryl; (6) amino; (7) $C_{1-6}$ alk-$C_{6-10}$ aryl; (8) azido; (9) $C_{3-8}$ cycloalkyl; (10) $C_{1-6}$ alk-$C_{3-8}$ cycloalkyl; (11) halo; (12) $C_{1-12}$ heterocyclyl (e.g., $C_{1-12}$ heteroaryl); (13) ($C_{1-12}$ heterocyclyl)oxy; (14) hydroxy; (15) nitro; (16) $C_{1-20}$ thioalkoxy (e.g., $C_{1-6}$ thioalkoxy); (17) $-(CH_2)_qCO_2R^{A'}$, where q is an integer from zero to four, and $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (18) $-(CH_2)_qCONR^{B'}R^{C'}$, where q is an integer from zero to four and where $R^{B'}$ and $R^{C'}$ are independently selected from the group consisting of (a) hydrogen, (b) $C_{6-10}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (19) $-(CH_2)_qSO_2R^{D'}$, where q is an integer from zero to four and where $R^{D'}$ is selected from the group consisting of (a) $C_{6-10}$ alkyl, (b) $C_{6-10}$ aryl, and (c) $C_{1-6}$ alk-$C_{6-10}$ aryl; (20) $-(CH_2)_qSO_2NR^{E'}R^{F'}$, where q is an integer from zero to four and where each of $R^{E'}$ and $R^{F'}$ is, independently, selected from the group consisting of (a) hydrogen, (b) $C_{6-10}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (21) thiol; (22) $C_{6-10}$ aryloxy; (23) $C_{3-8}$ cycloalkoxy; (24) $C_{6-10}$ aryl-$C_{1-6}$ alkoxy; (25) $C_{1-6}$ alk-$C_{1-12}$ heterocyclyl (e.g., $C_{1-6}$ alk-$C_{1-12}$ heteroaryl); (26) oxo; (27) $C_{2-20}$ alkenyl; and (28) $C_{2-20}$ alkynyl. In some embodiments, each of these groups can be further substituted as described herein. For example, the alkylene group of a $C_1$-alkaryl or a $C_1$-alkheterocyclyl can be further substituted with an oxo group to afford the respective aryloyl and (heterocyclyl)oyl substituent group.

**[00706]** The term "diastereomer," as used herein means stereoisomers that are not mirror images of one another and are non-superimposable on one another.

**[00707]** The term "effective amount" of an agent, as used herein, is that amount sufficient to effect beneficial or desired results, for example, clinical results, and, as such, an "effective amount" depends upon the context in which it is being applied. For example, in the context of administering an agent that treats cancer, an effective amount of an agent is, for example, an amount sufficient to achieve treatment, as defined herein, of cancer, as compared to the response obtained without administration of the agent.

**[00708]** The term "enantiomer," as used herein, means each individual optically active form of a compound of the invention, having an optical purity or enantiomeric excess (as determined by methods standard in the art) of at least 80% (i.e., at least 90% of one enantiomer and at most 10% of the other enantiomer), preferably at least 90% and more preferably at least 98%.

**[00709]** The term "halo," as used herein, represents a halogen selected from bromine, chlorine, iodine, or fluorine.

**[00710]** The term "haloalkoxy," as used herein, represents an alkoxy group, as defined herein, substituted by a halogen group (i.e., F, Cl, Br, or I). A haloalkoxy may be substituted with one, two, three, or, in the case of alkyl groups of two carbons or more, four halogens. Haloalkoxy groups include perfluoroalkoxys (e.g., $-OCF_3$), $-OCHF_2$, $-OCH_2F$, $-OCCl_3$, $-OCH_2CH_2Br$, $-OCH_2CH(CH_2CH_2Br)CH_3$, and $-OCHICH_3$. In some embodiments, the haloalkoxy group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for alkyl groups.

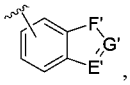
**[00711]** The term "haloalkyl," as used herein, represents an alkyl group, as defined herein, substituted by a halogen group (i.e., F, Cl, Br, or I). A haloalkyl may be substituted with one, two, three, or, in the case of alkyl groups of two carbons or more, four halogens. Haloalkyl groups include perfluoroalkyls (e.g., $-CF_3$), $-CHF_2$, $-CH_2F$, $-CCl_3$, $-CH_2CH_2Br$, $-CH_2CH(CH_2CH_2Br)CH_3$, and $-CHICH_3$. In some embodiments, the haloalkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for alkyl groups.

**[00712]** The term "heteroalkylene," as used herein, refers to an alkylene group, as defined herein, in which one or two of the constituent carbon atoms have each been replaced by nitrogen, oxygen, or sulfur. In some embodiments, the heteroalkylene group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein for alkylene groups.

**[00713]** The term "heteroaryl," as used herein, represents that subset of heterocyclyls, as defined herein, which are aromatic: i.e., they contain $4n+2$ pi electrons within the mono- or multicyclic ring system. Exemplary unsubstituted heteroaryl groups are of 1 to 12 (e.g., 1 to 11, 1 to 10, 1 to 9, 2 to 12, 2 to 11, 2 to 10, or 2 to 9) carbons. In some embodiment, the heteroaryl is substituted with 1, 2, 3, or 4 substituents groups as defined for a heterocyclyl group.

**[00714]** The term "heterocyclyl," as used herein represents a 5-, 6- or 7-membered ring, unless otherwise specified, containing one, two, three, or four heteroatoms independently selected from the group consisting of nitrogen, oxygen, and sulfur. The 5-membered ring has zero to two double bonds, and the 6- and 7-membered rings have zero to three double bonds. Exemplary unsubstituted heterocyclyl groups are of 1 to 12 (e.g., 1 to 11, 1 to 10, 1 to 9, 2 to 12, 2 to 11, 2 to 10, or 2 to 9) carbons. The term "heterocyclyl" also represents a heterocyclic compound having a bridged multicyclic structure in which one or more carbons and/or heteroatoms bridges two non-adjacent members of a monocyclic ring, e.g., a quinuclidinyl group. The term "heterocyclyl" includes bicyclic, tricyclic, and tetracyclic groups in which any of the above heterocyclic rings is fused to one, two, or three carbocyclic rings, e.g., an aryl ring, a cyclohexane ring, a cyclohexene ring, a cyclopentane ring, a cyclopentene ring, or another monocyclic heterocyclic ring, such as indolyl, quinolyl, isoquinolyl, tetrahydroquinolyl, benzofuryl, benzothienyl and the like. Examples of fused heterocyclyls include tropanes and 1,2,3,5,8,8a-hexahydroindolizine. Heterocyclics include pyrrolyl, pyrrolinyl, pyrrolidinyl, pyrazolyl, pyrazolinyl, pyrazolidinyl, imidazolyl, imidazolinyl, imidazolidinyl, pyridyl, piperidinyl, homopiperidinyl, pyrazinyl, piperazinyl, pyrimidinyl, pyridazinyl, oxazolyl, oxazolidinyl, isoxazolyl, isoxazolidiniyl, morpholinyl, thiomorpholinyl, thiazolyl, thiazolidinyl, isothiazolyl, isothiazolidinyl, indolyl, indazolyl, quinolyl, isoquinolyl, quinoxalinyl, dihydroquinoxalinyl, quinazolinyl, cinnolinyl, phthalazinyl, benzimidazolyl, benzothiazolyl, benzoxazolyl, benzothiadiazolyl, furyl, thienyl, thiazolidinyl, isothiazolyl, triazolyl, tetrazolyl, oxadiazolyl (e.g., 1,2,3-oxadiazolyl), purinyl, thiadiazolyl (e.g., 1,2,3-thiadiazolyl), tetrahydrofuranyl, dihydrofuranyl, tetrahydrothienyl, dihydrothienyl, dihydroindolyl, dihydroquinolyl, tetrahydroquinolyl, tetrahydroisoquinolyl, dihydroisoquinolyl, pyranyl, dihydropyranyl, dithiazolyl, benzofuranyl, isobenzofuranyl, benzothienyl, and the like, including dihydro and tetrahydro forms thereof, where one or more double bonds are reduced and replaced with hydrogens. Still other exemplary heterocyclyls include: 2,3,4,5-tetrahydro-2-oxo-oxazolyl; 2,3-dihydro-2-oxo-1H-imidazolyl; 2,3,4,5-tetrahydro-5-oxo-1H-pyrazolyl (e.g., 2,3,4,5-tetrahydro-2-

phenyl-5-oxo-1H-pyrazolyl); 2,3,4,5-tetrahydro-2,4-dioxo-1H-imidazolyl (e.g., 2,3,4,5-tetrahydro-2,4-dioxo-5-methyl-5-phenyl-1H-imidazolyl); 2,3-dihydro-2-thioxo-1,3,4-oxadiazolyl (e.g., 2,3-dihydro-2-thioxo-5-phenyl-1,3,4-oxadiazolyl); 4,5-dihydro-5-oxo-1*H*-triazolyl (e.g., 4,5-dihydro-3-methyl-4-amino 5-oxo-1*H*-triazolyl); 1,2,3,4-tetrahydro-2,4-dioxopyridinyl (e.g., 1,2,3,4-tetrahydro-2,4-dioxo-3,3-diethylpyridinyl); 2,6-dioxo-piperidinyl (e.g., 2,6-dioxo-3-ethyl-3-phenylpiperidinyl); 1,6-dihydro-6-oxopyridiminyl; 1,6-dihydro-4-oxopyrimidinyl (e.g., 2-(methylthio)-1,6-dihydro-4-oxo-5-methylpyrimidin-1-yl); 1,2,3,4-tetrahydro-2,4-dioxopyrimidinyl (e.g., 1,2,3,4-tetrahydro-2,4-dioxo-3-ethylpyrimidinyl); 1,6-dihydro-6-oxo-pyridazinyl (e.g., 1,6-dihydro-6-oxo-3-ethylpyridazinyl); 1,6-dihydro-6-oxo-1,2,4-triazinyl (e.g., 1,6-dihydro-5-isopropyl-6-oxo-1,2,4-triazinyl); 2,3-dihydro-2-oxo-1*H*-indolyl (e.g., 3,3-dimethyl-2,3-dihydro-2-oxo-1*H*-indolyl and 2,3-dihydro-2-oxo-3,3′-spiropropane-1*H*-indol-1-yl); 1,3-dihydro-1-oxo-2*H*-iso-indolyl; 1,3-dihydro-1,3-dioxo-2*H*-iso-indolyl; 1*H*-benzopyrazolyl (e.g., 1-(ethoxycarbonyl)- 1*H*-benzopyrazolyl); 2,3-dihydro-2-oxo-1*H*-benzimidazolyl (e.g., 3-ethyl-2,3-dihydro-2-oxo-1*H*-benzimidazolyl); 2,3-dihydro-2-oxo-benzoxazolyl (e.g., 5-chloro-2,3-dihydro-2-oxo-benzoxazolyl); 2,3-dihydro-2-oxo-benzoxazolyl; 2-oxo-2H-benzopyranyl; 1,4-benzodioxanyl; 1,3-benzodioxanyl; 2,3-dihydro-3-oxo,4*H*-1,3-benzothiazinyl; 3,4-dihydro-4-oxo-3*H*-quinazolinyl (e.g., 2-methyl-3,4-dihydro-4-oxo-3*H*-quinazolinyl); 1,2,3,4-tetrahydro-2,4-dioxo-3*H*-quinazolyl (e.g., 1-ethyl-1,2,3,4-tetrahydro-2,4-dioxo-3*H*-quinazolyl); 1,2,3,6-tetrahydro-2,6-dioxo-7*H*-purinyl (e.g., 1,2,3,6-tetrahydro-1,3-dimethyl-2,6-dioxo-7 *H* -purinyl); 1,2,3,6-tetrahydro-2,6-dioxo-1 *H* –purinyl (e.g., 1,2,3,6-tetrahydro-3,7-dimethyl-2,6-dioxo-1 *H* -purinyl); 2-oxobenz[*c,d*]indolyl; 1,1-dioxo-2H-naphth[1,8-*c,d*]isothiazolyl; and 1,8-naphthylenedicarboxamido. Additional heterocyclics include 3,3a,4,5,6,6a-hexahydro-pyrrolo[3,4-b]pyrrol-(2H)-yl, and 2,5-diazabicyclo[2.2.1]heptan-2-yl, homopiperazinyl (or diazepanyl), tetrahydropyranyl, dithiazolyl, benzofuranyl, benzothienyl, oxepanyl, thiepanyl, azocanyl, oxecanyl, and thiocanyl. Heterocyclic groups also include groups of the formula



, where

**[00715]** E′ is selected from the group consisting of -N- and -CH-; F′ is selected from the group consisting of -N=CH-, -NH-$CH_2$-, -NH-C(O)-, -NH-, -CH=N-, -$CH_2$-NH-, -C(O)-NH-, -CH=CH-, -$CH_2$-, -$CH_2CH_2$-, -$CH_2O$-, -$OCH_2$-, -O-, and -S-; and G′ is selected from the group consisting of -CH- and -N-. Any of the heterocyclyl groups mentioned herein may be optionally substituted with one, two, three, four or five substituents independently selected from the group consisting of: (1) $C_1$-

$_7$ acyl (e.g., carboxyaldehyde ); (2) $C_{1-20}$ alkyl (e.g., $C_{1-6}$ alkyl, $C_{1-6}$ alkoxy-$C_{1-6}$ alkyl, $C_{1-6}$ alkylsulfinyl-$C_{1-6}$ alkyl, amino-$C_{1-6}$ alkyl, azido-$C_{1-6}$ alkyl, (carboxyaldehyde)-$C_{1-6}$ alkyl, halo-$C_{1-6}$ alkyl (e.g., perfluoroalkyl), hydroxy-$C_{1-6}$ alkyl, nitro-$C_{1-6}$ alkyl, or $C_{1-6}$ thioalkoxy-$C_{1-6}$ alkyl); (3) $C_{1-20}$ alkoxy (e.g., $C_{1-6}$ alkoxy, such as perfluoroalkoxy); (4) $C_{1-6}$ alkylsulfinyl; (5) $C_{6-10}$ aryl; (6) amino; (7) $C_{1-6}$ alk-$C_{6-10}$ aryl; (8) azido; (9) $C_{3-8}$ cycloalkyl; (10) $C_{1-6}$ alk-$C_{3-8}$ cycloalkyl; (11) halo; (12) $C_{1-12}$ heterocyclyl (e.g., $C_{2-12}$ heteroaryl); (13) ($C_{1-12}$ heterocyclyl)oxy; (14) hydroxy; (15) nitro; (16) $C_{1-20}$ thioalkoxy (e.g., $C_{1-6}$ thioalkoxy); (17) $-(CH_2)_qCO_2R^{A'}$, where q is an integer from zero to four, and $R^{A'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, (c) hydrogen, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (18) $-(CH_2)_qCONR^{B'}R^{C'}$, where q is an integer from zero to four and where $R^{B'}$ and $R^{C'}$ are independently selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (19) $-(CH_2)_qSO_2R^{D'}$, where q is an integer from zero to four and where $R^{D'}$ is selected from the group consisting of (a) $C_{1-6}$ alkyl, (b) $C_{6-10}$ aryl, and (c) $C_{1-6}$ alk-$C_{6-10}$ aryl; (20) $-(CH_2)_qSO_2NR^{E'}R^{F'}$, where q is an integer from zero to four and where each of $R^{E'}$ and $R^{F'}$ is, independently, selected from the group consisting of (a) hydrogen, (b) $C_{1-6}$ alkyl, (c) $C_{6-10}$ aryl, and (d) $C_{1-6}$ alk-$C_{6-10}$ aryl; (21) thiol; (22) $C_{6-10}$ aryloxy; (23) $C_{3-8}$ cycloalkoxy; (24) arylalkoxy; (25) $C_{1-6}$ alk-$C_{1-12}$ heterocyclyl (e.g., $C_{1-6}$ alk-$C_{1-12}$ heteroaryl); (26) oxo; (27) ($C_{1-12}$ heterocyclyl)imino; (28) $C_{2-20}$ alkenyl; and (29) $C_{2-20}$ alkynyl. In some embodiments, each of these groups can be further substituted as described herein. For example, the alkylene group of a $C_1$-alkaryl or a $C_1$-alkheterocyclyl can be further substituted with an oxo group to afford the respective aryloyl and (heterocyclyl)oyl substituent group.

**[00716]** The term "(heterocyclyl) imino," as used herein, represents a heterocyclyl group, as defined herein, attached to the parent molecular group through an imino group. In some embodiments, the heterocyclyl group can be substituted with 1, 2, 3, or 4 substituent groups as defined herein.

**[00717]** The term "(heterocyclyl)oxy," as used herein, represents a heterocyclyl group, as defined herein, attached to the parent molecular group through an oxygen atom. In some embodiments, the heterocyclyl group can be substituted with 1, 2, 3, or 4 substituent groups as defined herein.

**[00718]** The term "(heterocyclyl)oyl," as used herein, represents a heterocyclyl group, as defined herein, attached to the parent molecular group through a carbonyl group. In some embodiments, the heterocyclyl group can be substituted with 1, 2, 3, or 4 substituent groups as defined herein.

**[00719]** The term "hydrocarbon," as used herein, represents a group consisting only of carbon and hydrogen atoms.

**[00720]** The term "hydroxy," as used herein, represents an –OH group. In some embodiments, the hydroxy group can be substituted with 1, 2, 3, or 4 substituent groups (e.g., *O*-protecting groups) as defined herein for an alkyl.

**[00721]** The term "hydroxyalkenyl," as used herein, represents an alkenyl group, as defined herein, substituted by one to three hydroxy groups, with the proviso that no more than one hydroxy group may be attached to a single carbon atom of the alkyl group, and is exemplified by dihydroxypropenyl, hydroxyisopentenyl, and the like. In some embodiments, the hydroxyalkenyl group can be substituted with 1, 2, 3, or 4 substituent groups (e.g., *O*-protecting groups) as defined herein for an alkyl.

**[00722]** The term "hydroxyalkyl," as used herein, represents an alkyl group, as defined herein, substituted by one to three hydroxy groups, with the proviso that no more than one hydroxy group may be attached to a single carbon atom of the alkyl group, and is exemplified by hydroxymethyl, dihydroxypropyl, and the like. In some embodiments, the hydroxyalkyl group can be substituted with 1, 2, 3, or 4 substituent groups (e.g., *O*-protecting groups) as defined herein for an alkyl.

**[00723]** The term "hydroxyalkynyl," as used herein, represents an alkynyl group, as defined herein, substituted by one to three hydroxy groups, with the proviso that no more than one hydroxy group may be attached to a single carbon atom of the alkyl group. In some embodiments, the hydroxyalkynyl group can be substituted with 1, 2, 3, or 4 substituent groups (e.g., *O*-protecting groups) as defined herein for an alkyl.

**[00724]** The term "isomer," as used herein, means any tautomer, stereoisomer, enantiomer, or diastereomer of any compound of the invention. It is recognized that the compounds of the invention can have one or more chiral centers and/or double bonds and, therefore, exist as stereoisomers, such as double-bond isomers (i.e., geometric E/Z isomers) or diastereomers (e.g., enantiomers (i.e., (+) or (-)) or cis/trans isomers). According to the invention, the chemical structures depicted herein, and therefore the compounds of the invention, encompass all of the corresponding stereoisomers, that is, both the stereomerically pure form (e.g., geometrically pure, enantiomerically pure, or diastereomerically pure) and enantiomeric and stereoisomeric mixtures, e.g., racemates. Enantiomeric and stereoisomeric mixtures of compounds of the invention can typically be resolved into their component enantiomers or stereoisomers by well-known methods,

such as chiral-phase gas chromatography, chiral-phase high performance liquid chromatography, crystallizing the compound as a chiral salt complex, or crystallizing the compound in a chiral solvent. Enantiomers and stereoisomers can also be obtained from stereomerically or enantiomerically pure intermediates, reagents, and catalysts by well-known asymmetric synthetic methods.

**[00725]** The term "*N*-protected amino," as used herein, refers to an amino group, as defined herein, to which is attached one or two *N*-protecting groups, as defined herein.

**[00726]** The term "*N*-protecting group," as used herein, represents those groups intended to protect an amino group against undesirable reactions during synthetic procedures. Commonly used *N*-protecting groups are disclosed in Greene, "Protective Groups in Organic Synthesis," 3rd Edition (John Wiley & Sons, New York, 1999), which is incorporated herein by reference. *N*-protecting groups include acyl, aryloyl, or carbamyl groups such as formyl, acetyl, propionyl, pivaloyl, t-butylacetyl, 2-chloroacetyl, 2-bromoacetyl, trifluoroacetyl, trichloroacetyl, phthalyl, o-nitrophenoxyacetyl, α-chlorobutyryl, benzoyl, 4-chlorobenzoyl, 4-bromobenzoyl, 4-nitrobenzoyl, and chiral auxiliaries such as protected or unprotected D, L or D, L-amino acids such as alanine, leucine, phenylalanine, and the like; sulfonyl-containing groups such as benzenesulfonyl, p-toluenesulfonyl, and the like; carbamate forming groups such as benzyloxycarbonyl, p-chlorobenzyloxycarbonyl, p-methoxybenzyloxycarbonyl, p-nitrobenzyloxycarbonyl, 2-nitrobenzyloxycarbonyl, p-bromobenzyloxycarbonyl, 3,4-dimethoxybenzyloxycarbonyl, 3,5-dimethoxybenzyloxycarbonyl, 2,4-dimethoxybenzyloxycarbonyl, 4-methoxybenzyloxycarbonyl, 2-nitro-4,5-dimethoxybenzyloxycarbonyl, 3,4,5-trimethoxybenzyloxycarbonyl, 1-(p-biphenylyl)-1-methylethoxycarbonyl, α,α-dimethyl-3,5-dimethoxybenzyloxycarbonyl, benzhydryloxy carbonyl, t-butyloxycarbonyl, diisopropylmethoxycarbonyl, isopropyloxycarbonyl, ethoxycarbonyl, methoxycarbonyl, allyloxycarbonyl, 2,2,2,-trichloroethoxycarbonyl, phenoxycarbonyl, 4-nitrophenoxy carbonyl, fluorenyl-9-methoxycarbonyl, cyclopentyloxycarbonyl, adamantyloxycarbonyl, cyclohexyloxycarbonyl, phenylthiocarbonyl, and the like, alkaryl groups such as benzyl, triphenylmethyl, benzyloxymethyl, and the like and silyl groups, such as trimethylsilyl, and the like. Preferred *N*-protecting groups are formyl, acetyl, benzoyl, pivaloyl, t-butylacetyl, alanyl, phenylsulfonyl, benzyl, t-butyloxycarbonyl (Boc), and benzyloxycarbonyl (Cbz).

**[00727]** The term "nitro," as used herein, represents an $-NO_2$ group.

**[00728]** The term "*O*-protecting group," as used herein, represents those groups intended to protect an oxygen containing (e.g., phenol, hydroxyl, or carbonyl) group against undesirable reactions during synthetic procedures. Commonly used *O*-protecting groups are disclosed in Greene, "Protective Groups in Organic Synthesis," 3rd Edition (John Wiley & Sons, New York, 1999), which is incorporated herein by reference. Exemplary *O*-protecting groups include acyl, aryloyl, or carbamyl groups, such as formyl, acetyl, propionyl, pivaloyl, t-butylacetyl, 2-chloroacetyl, 2-bromoacetyl, trifluoroacetyl, trichloroacetyl, phthalyl, o-nitrophenoxyacetyl, α-chlorobutyryl, benzoyl, 4-chlorobenzoyl, 4-bromobenzoyl, *t*-butyldimethylsilyl, tri-*iso*-propylsilyloxymethyl, 4,4'-dimethoxytrityl, isobutyryl, phenoxyacetyl, 4-isopropylpehenoxyacetyl, dimethylformamidino, and 4-nitrobenzoyl; alkylcarbonyl groups, such as acyl, acetyl, propionyl, pivaloyl, and the like; optionally substituted arylcarbonyl groups, such as benzoyl; silyl groups, such as trimethylsilyl (TMS), tert-butyldimethylsilyl (TBDMS), tri-iso-propylsilyloxymethyl (TOM), triisopropylsilyl (TIPS), and the like; ether-forming groups with the hydroxyl, such methyl, methoxymethyl, tetrahydropyranyl, benzyl, p-methoxybenzyl, trityl, and the like; alkoxycarbonyls, such as methoxycarbonyl, ethoxycarbonyl, isopropoxycarbonyl, n-isopropoxycarbonyl, n-butyloxycarbonyl, isobutyloxycarbonyl, sec-butyloxycarbonyl, t-butyloxycarbonyl, 2-ethylhexyloxycarbonyl, cyclohexyloxycarbonyl, methyloxycarbonyl, and the like; alkoxyalkoxycarbonyl groups, such as methoxymethoxycarbonyl, ethoxymethoxycarbonyl, 2-methoxyethoxycarbonyl, 2-ethoxyethoxycarbonyl, 2-butoxyethoxycarbonyl, 2-methoxyethoxymethoxycarbonyl, allyloxycarbonyl, propargyloxycarbonyl, 2-butenoxycarbonyl, 3-methyl-2-butenoxycarbonyl, and the like; haloalkoxycarbonyls, such as 2-chloroethoxycarbonyl, 2-chloroethoxycarbonyl, 2,2,2-trichloroethoxycarbonyl, and the like; optionally substituted arylalkoxycarbonyl groups, such as benzyloxycarbonyl, p-methylbenzyloxycarbonyl, p-methoxybenzyloxycarbonyl, p-nitrobenzyloxycarbonyl, 2,4-dinitrobenzyloxycarbonyl, 3,5-dimethylbenzyloxycarbonyl, p-chlorobenzyloxycarbonyl, p-bromobenzyloxy-carbonyl, fluorenylmethyloxycarbonyl, and the like; and optionally substituted aryloxycarbonyl groups, such as phenoxycarbonyl, p-nitrophenoxycarbonyl, o-nitrophenoxycarbonyl, 2,4-dinitrophenoxycarbonyl, p-methyl-phenoxycarbonyl, m-methylphenoxycarbonyl, o-bromophenoxycarbonyl, 3,5-dimethylphenoxycarbonyl, p-chlorophenoxycarbonyl, 2-chloro-4-nitrophenoxy-carbonyl, and the like); substituted alkyl, aryl, and alkaryl ethers (e.g., trityl; methylthiomethyl; methoxymethyl; benzyloxymethyl; siloxymethyl; 2,2,2,-trichloroethoxymethyl; tetrahydropyranyl; tetrahydrofuranyl;

ethoxyethyl; 1-[2-(trimethylsilyl)ethoxy]ethyl; 2-trimethylsilylethyl; t-butyl ether; p-chlorophenyl, p-methoxyphenyl, p-nitrophenyl, benzyl, p-methoxybenzyl, and nitrobenzyl); silyl ethers (e.g., trimethylsilyl; triethylsilyl; triisopropylsilyl; dimethylisopropylsilyl; t-butyldimethylsilyl; t-butyldiphenylsilyl; tribenzylsilyl; triphenylsilyl; and diphenymethylsilyl); carbonates (e.g., methyl, methoxymethyl, 9-fluorenylmethyl; ethyl; 2,2,2-trichloroethyl; 2-(trimethylsilyl)ethyl; vinyl, allyl, nitrophenyl; benzyl; methoxybenzyl; 3,4-dimethoxybenzyl; and nitrobenzyl); carbonyl-protecting groups (e.g., acetal and ketal groups, such as dimethyl acetal, 1,3-dioxolane, and the like; acylal groups; and dithiane groups, such as 1,3-dithianes, 1,3-dithiolane, and the like); carboxylic acid-protecting groups (e.g., ester groups, such as methyl ester, benzyl ester, t-butyl ester, orthoesters, and the like; and oxazoline groups.

**[00729]** The term "oxo" as used herein, represents =O.

**[00730]** The term "perfluoroalkyl," as used herein, represents an alkyl group, as defined herein, where each hydrogen radical bound to the alkyl group has been replaced by a fluoride radical. Perfluoroalkyl groups are exemplified by trifluoromethyl, pentafluoroethyl, and the like.

**[00731]** The term "perfluoroalkoxy," as used herein, represents an alkoxy group, as defined herein, where each hydrogen radical bound to the alkoxy group has been replaced by a fluoride radical. Perfluoroalkoxy groups are exemplified by trifluoromethoxy, pentafluoroethoxy, and the like.

**[00732]** The term "spirocyclyl," as used herein, represents a $C_{2-7}$ alkylene diradical, both ends of which are bonded to the same carbon atom of the parent group to form a spirocyclic group, and also a $C_{1-6}$ heteroalkylene diradical, both ends of which are bonded to the same atom. The heteroalkylene radical forming the spirocyclyl group can containing one, two, three, or four heteroatoms independently selected from the group consisting of nitrogen, oxygen, and sulfur. In some embodiments, the spirocyclyl group includes one to seven carbons, excluding the carbon atom to which the diradical is attached. The spirocyclyl groups of the invention may be optionally substituted with 1, 2, 3, or 4 substituents provided herein as optional substituents for cycloalkyl and/or heterocyclyl groups.

**[00733]** The term "stereoisomer," as used herein, refers to all possible different isomeric as well as conformational forms which a compound may possess (e.g., a compound of any formula described herein), in particular all possible stereochemically and conformationally isomeric forms, all diastereomers, enantiomers and/or conformers of the basic molecular structure. Some compounds of

the present invention may exist in different tautomeric forms, all of the latter being included within the scope of the present invention.

**[00734]** The term "sulfoalkyl," as used herein, represents an alkyl group, as defined herein, substituted by a sulfo group of $-SO_3H$. In some embodiments, the alkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein, and the sulfo group can be further substituted with one or more *O*-protecting groups (e.g., as described herein).

**[00735]** The term "sulfonyl," as used herein, represents an $-S(O)_2-$ group.

**[00736]** The term "thioalkaryl," as used herein, represents a chemical substituent of formula –SR, where R is an alkaryl group. In some embodiments, the alkaryl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein.

**[00737]** The term "thioalkheterocyclyl," as used herein, represents a chemical substituent of formula –SR, where R is an alkheterocyclyl group. In some embodiments, the alkheterocyclyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein.

**[00738]** The term "thioalkoxy," as used herein, represents a chemical substituent of formula –SR, where R is an alkyl group, as defined herein. In some embodiments, the alkyl group can be further substituted with 1, 2, 3, or 4 substituent groups as described herein.

**[00739]** *Compound:* As used herein, the term "compound," is meant to include all stereoisomers, geometric isomers, tautomers, and isotopes of the structures depicted.

**[00740]** 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.

**[00741]** 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.

**[00742]** 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.

**[00743]** 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.

**[00744]** *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.

**[00745]** 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.

**[00746]** *Cyclic or Cyclized:* As used herein, the term "cyclic" refers to the presence of a continuous loop. Cyclic molecules need not be circular, only joined to form an unbroken chain of subunits. Cyclic molecules such as the mRNA of the present invention may be single units or multimers or comprise one or more components of a complex or higher order structure.

**[00747]** *Cytostatic*: As used herein, "cytostatic" refers to inhibiting, reducing, suppressing the growth, division, or multiplication of a cell (*e.g.*, a mammalian cell (*e.g.*, a human cell)), bacterium, virus, fungus, protozoan, parasite, prion, or a combination thereof.

**[00748]** *Cytotoxic*: As used herein, "cytotoxic" refers to killing or causing injurious, toxic, or deadly effect on a cell (*e.g.*, a mammalian cell (*e.g.*, a human cell)), bacterium, virus, fungus, protozoan, parasite, prion, or a combination thereof.

**[00749]** *Delivery:* As used herein, "delivery" refers to the act or manner of delivering a compound, substance, entity, moiety, cargo or payload.

**[00750]** *Delivery Agent*: As used herein, "delivery agent" refers to any substance which facilitates, at least in part, the *in vivo* delivery of a polynucleotide to targeted cells.

**[00751]** *Destabilized:* As used herein, the term "destable," "destabilize," or "destabilizing region" means a region or molecule that is less stable than a starting, wild-type or native form of the same region or molecule.

**[00752]** *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, streptavidin 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.

**[00753]** *Digest*: As used herein, the term "digest" means to break apart into smaller pieces or components. When referring to polypeptides or proteins, digestion results in the production of peptides.

**[00754]** *Distal:* As used herein, the term "distal" means situated away from the center or away from a point or region of interest.

**[00755]** *Encoded protein cleavage signal*: As used herein, "encoded protein cleavage signal" refers to the nucleotide sequence which encodes a protein cleavage signal.

**[00756]** *Engineered:* As used herein, embodiments of the invention are "engineered" when they are designed to have a feature or property, whether structural or chemical, that varies from a starting point, wild type or native molecule.

**[00757]** *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.

**[00758]** *Feature:* As used herein, a "feature" refers to a characteristic, a property, or a distinctive element.

**[00759]** *Formulation*: As used herein, a "formulation" includes at least a polynucleotide and a delivery agent.

**[00760]** *Fragment:* A "fragment," as used herein, refers to a portion. For example, fragments of proteins may comprise polypeptides obtained by digesting full-length protein isolated from cultured cells.

**[00761]** *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.

**[00762]** *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%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% identical or similar. The term "homologous" necessarily refers to a comparison between at least two sequences (polynucleotide or polypeptide sequences). In accordance with the invention, two polynucleotide sequences are considered to be homologous if the polypeptides they encode are at least about 50%, 60%, 70%, 80%, 90%, 95%, or even 99% for at least one stretch of at least about 20 amino acids. 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%, 60%, 70%, 80%, or 90% identical for at least one stretch of at least about 20 amino acids.

**[00763]** *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, 4:11-17), 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 NWSgapdna.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 Altschul, S. F. *et al.*, *J. Molec. Biol.*, 215, 403 (1990)).

**[00764]** *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.

**[00765]** *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).

**[00766]** *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).

**[00767]** *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.

**[00768]** *Linker:* As used herein, a linker 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, alkylamino, oxygen, sulfur, sulfoxide, sulfonyl, 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., a 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. The linker can be used for any useful purpose, such as to form multimers (e.g., through linkage of two or more polynucleotides) or conjugates, as well as to administer a payload, as described herein. Examples of chemical groups that can be incorporated into the linker include, but are not limited to, alkyl, alkenyl, alkynyl, amido, amino, ether, thioether, ester, alkylene, heteroalkylene, aryl, or heterocyclyl, each of which can be optionally substituted, as described herein. Examples of linkers include, but are not limited to, unsaturated alkanes, polyethylene glycols (e.g., ethylene or propylene glycol monomeric units, e.g., diethylene glycol, dipropylene glycol, triethylene glycol, tripropylene glycol, tetraethylene glycol, or tetraethylene glycol), and dextran polymers, 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. Non-limiting examples of a selectively cleavable bond include an amido bond can be cleaved for example by the use of tris(2-carboxyethyl)phosphine (TCEP), or other reducing agents, and/or photolysis, as well as an ester bond can be cleaved for example by acidic or basic hydrolysis.

**[00769]** *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, e.g., as it relates to the natural ribonucleotides A, U, G, and C. Noncanonical nucleotides such as the cap structures are not considered "modified" although they differ from the chemical structure of the A, C, G, U ribonucleotides.

**[00770]** *Naturally occurring:* As used herein, "naturally occurring" means existing in nature without artificial aid.

**[00771]** *Non-human vertebrate:* As used herein, a "non human vertebrate" includes all vertebrates except *Homo sapiens*, including wild and domesticated species. Examples of non-human vertebrates include, but are not limited to, mammals, such as alpaca, banteng, bison, camel, cat, cattle, deer, dog, donkey, gayal, goat, guinea pig, horse, llama, mule, pig, rabbit, reindeer, sheep water buffalo, and yak.

**[00772]** *Off-target:* As used herein, "off target" refers to any unintended effect on any one or more target, gene, or cellular transcript.

**[00773]** *Open reading frame:* As used herein, "open reading frame" or "ORF" refers to a sequence which does not contain a stop codon in a given reading frame.

**[00774]** *Operably linked:* As used herein, the phrase "operably linked" refers to a functional connection between two or more molecules, constructs, transcripts, entities, moieties or the like.

**[00775]** *Paratope:* As used herein, a "paratope" refers to the antigen-binding site of an antibody.

**[00776]** *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.

**[00777]** *Optionally substituted:* Herein a phrase of the form "optionally substituted X" (*e.g.*, optionally substituted alkyl) is intended to be equivalent to "X, wherein X is optionally substituted" (*e.g.*, "alkyl, wherein said alkyl is optionally substituted"). It is not intended to mean that the feature "X" (*e.g.* alkyl) *per se* is optional.

**[00778]** *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.

**[00779]** *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.

**[00780]** *Pharmaceutically acceptable excipients:* The phrase "pharmaceutically acceptable excipient," as used herein, refers any ingredient other than the compounds described herein (for example, a vehicle capable of suspending or dissolving the active compound) and having the properties of being substantially nontoxic and non-inflammatory in a patient. Excipients may include, for example: antiadherents, antioxidants, binders, coatings, compression aids, disintegrants,

dyes (colors), emollients, emulsifiers, fillers (diluents), film formers or coatings, flavors, fragrances, glidants (flow enhancers), lubricants, preservatives, printing inks, sorbents, suspensing or dispersing agents, sweeteners, and waters of hydration. Exemplary excipients include, but are not limited to: butylated hydroxytoluene (BHT), calcium carbonate, calcium phosphate (dibasic), calcium stearate, croscarmellose, crosslinked polyvinyl pyrrolidone, citric acid, crospovidone, cysteine, ethylcellulose, gelatin, hydroxypropyl cellulose, hydroxypropyl methylcellulose, lactose, magnesium stearate, maltitol, mannitol, methionine, methylcellulose, methyl paraben, microcrystalline cellulose, polyethylene glycol, polyvinyl pyrrolidone, povidone, pregelatinized starch, propyl paraben, retinyl palmitate, shellac, silicon dioxide, sodium carboxymethyl cellulose, sodium citrate, sodium starch glycolate, sorbitol, starch (corn), stearic acid, sucrose, talc, titanium dioxide, vitamin A, vitamin E, vitamin C, and xylitol.

**[00781]** *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 (e.g., by reacting the free base group with a suitable organic acid). 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. Representative acid addition salts include acetate, adipate, alginate, ascorbate, aspartate, benzenesulfonate, benzoate, bisulfate, borate, butyrate, camphorate, camphorsulfonate, citrate, cyclopentanepropionate, digluconate, dodecylsulfate, ethanesulfonate, fumarate, glucoheptonate, glycerophosphate, hemisulfate, heptonate, hexanoate, hydrobromide, hydrochloride, hydroiodide, 2-hydroxy-ethanesulfonate, lactobionate, lactate, laurate, lauryl sulfate, malate, maleate, malonate, methanesulfonate, 2-naphthalenesulfonate, nicotinate, nitrate, oleate, oxalate, palmitate, pamoate, pectinate, persulfate, 3-phenylpropionate, phosphate, picrate, pivalate, propionate, stearate, succinate, sulfate, tartrate, thiocyanate, toluenesulfonate, undecanoate, valerate salts, and the like. Representative alkali or alkaline earth metal salts include sodium, lithium, potassium, calcium, magnesium, and the like, as well as nontoxic ammonium, quaternary ammonium, and amine cations, including, but not limited to ammonium, tetramethylammonium, tetraethylammonium, methylamine, dimethylamine, trimethylamine, triethylamine, ethylamine, 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, *Pharmaceutical Salts: Properties, Selection, and Use*, P.H. Stahl and C.G. Wermuth (eds.), Wiley-VCH, 2008, and Berge et al., *Journal of Pharmaceutical Science*, 66, 1-19 (1977), each of which is incorporated herein by reference in its entirety.

**[00782]** *Pharmacokinetic:* As used herein, "pharmacokinetic" refers to any one or more properties of a molecule or compound as it relates to the determination of the fate of substances administered to a living organism. Pharmacokinetics is divided into several areas including the extent and rate of absorption, distribution, metabolism and excretion. This is commonly referred to as ADME where: (A) Absorption is the process of a substance entering the blood circulation; (D) Distribution is the dispersion or dissemination of substances throughout the fluids and tissues of the body; (M) Metabolism (or Biotransformation) is the irreversible transformation of parent compounds into daughter metabolites; and (E) Excretion (or Elimination) refers to the elimination of the substances from the body. In rare cases, some drugs irreversibly accumulate in body tissue.

**[00783]** *Pharmaceutically acceptable solvate*: The term "pharmaceutically acceptable solvate," as used herein, means a compound of the invention wherein molecules of a suitable solvent are incorporated in the crystal lattice. A suitable solvent is physiologically tolerable at the dosage administered. For example, solvates may be prepared by crystallization, recrystallization, or precipitation from a solution that includes organic solvents, water, or a mixture thereof. Examples of suitable solvents are ethanol, water (for example, mono-, di-, and tri-hydrates), *N*-methylpyrrolidinone (NMP), dimethyl sulfoxide (DMSO), *N,N'*-dimethylformamide (DMF), *N,N'*-dimethylacetamide (DMAC), 1,3-dimethyl-2-imidazolidinone (DMEU), 1,3-dimethyl-3,4,5,6-tetrahydro-2-(1H)-pyrimidinone (DMPU), acetonitrile (ACN), propylene glycol, ethyl acetate, benzyl alcohol, 2-pyrrolidone, benzyl benzoate, and the like. When water is the solvent, the solvate is referred to as a "hydrate."

**[00784]** *Physicochemical:* As used herein, "physicochemical" means of or relating to a physical and/or chemical property.

**[00785]** *Preventing*: As used herein, the term "preventing" refers to partially or completely delaying onset of an infection, disease, disorder and/or condition; partially or completely delaying onset of one or more symptoms, features, or clinical manifestations of a particular infection, disease, disorder, and/or condition; partially or completely delaying onset of one or more symptoms, features, or manifestations of a particular infection, disease, disorder, and/or condition; partially or completely delaying progression from an infection, a particular disease, disorder and/or condition; and/or decreasing the risk of developing pathology associated with the infection, the disease, disorder, and/or condition.

**[00786]** *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.

**[00787]** *Proliferate:* As used herein, the term "proliferate" means to grow, expand or increase or cause to grow, expand or increase rapidly. "Proliferative" means having the ability to proliferate. "Anti-proliferative" means having properties counter to or inapposite to proliferative properties.

**[00788]** *Protein cleavage site*: As used herein, "protein cleavage site" refers to a site where controlled cleavage of the amino acid chain can be accomplished by chemical, enzymatic or photochemical means.

**[00789]** *Protein cleavage signal:* As used herein "protein cleavage signal" refers to at least one amino acid that flags or marks a polypeptide for cleavage.

**[00790]** *Protein of interest:* As used herein, the terms "proteins of interest" or "desired proteins" include those provided herein and fragments, mutants, variants, and alterations thereof.

**[00791]** *Proximal:* As used herein, the term "proximal" means situated nearer to the center or to a point or region of interest.

**[00792]** *Purified:* As used herein, "purify," "purified," "purification" means to make substantially pure or clear from unwanted components, material defilement, admixture or imperfection.

**[00793]** *Sample:* As used herein, the term "sample" or "biological 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.

**[00794]** *Signal Sequences:* As used herein, the phrase "signal sequences" refers to a sequence which can direct the transport or localization of a protein.

**[00795]** *Significant or Significantly:* As used herein, the terms "significant" or "significantly" are used synonymously with the term "substantially."

**[00796]** *Single unit dose*: As used herein, a "single unit dose" is a dose of any therapeutic administed in one dose/at one time/single route/single point of contact, i.e., single administration event.

**[00797]** *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.

**[00798]** *Split dose*: As used herein, a "split dose" is the division of single unit dose or total daily dose into two or more doses.

**[00799]** *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.

**[00800]** *Stabilized:* As used herein, the term "stabilize", "stabilized," "stabilized region" means to make or become stable.

**[00801]** *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.

**[00802]** *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.

**[00803]** *Substantially equal*: As used herein as it relates to time differences between doses, the term means plus/minus 2%.

**[00804]** *Substantially simultaneously*: As used herein and as it relates to plurality of doses, the term means within 2 seconds.

**[00805]** *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.

**[00806]** *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.

**[00807]** *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.

**[00808]** *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.

**[00809]** *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.

**[00810]** *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 an infection, disease, disorder, and/or condition, to treat, improve symptoms of, diagnose, prevent, and/or delay the onset of the infection, disease, disorder, and/or condition.

**[00811]** *Therapeutically effective outcome*: As used herein, the term "therapeutically effective outcome" means an outcome that is sufficient in a subject suffering from or susceptible to an infection, disease, disorder, and/or condition, to treat, improve symptoms of, diagnose, prevent, and/or delay the onset of the infection, disease, disorder, and/or condition.

**[00812]** *Total daily dose:* As used herein, a "total daily dose" is an amount given or prescribed in 24 hr period. It may be administered as a single unit dose.

**[00813]** *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.

**[00814]** *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 infection, 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.

**[00815]** *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**

**[00816]** 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.

**[00817]** 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.

**[00818]** It is also noted that the term "comprising" is intended to be open and permits but does not require the inclusion of additional elements or steps. When the term "comprising" is used herein, the term "consisting of" is thus also encompassed and disclosed.

**[00819]** 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.

**[00820]** 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.

**[00821]** 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**

**[00822]** The present disclosure is further described in the following examples, which do not limit the scope of the disclosure described in the claims.

**Example 1. Modified mRNA *In Vitro* Transcription**

**A. Materials and Methods**

**[00823]** Modified mRNAs according to the invention are made using standard laboratory methods and materials for in vitro transcription with the exception that the nucleotide mix contains 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 mRNAs not

incorporating adenosine analogs. Adenosine-containing mRNAs are synthesized without an oligo (dT) sequence to allow for post-transcription poly (A) polymerase poly-(A) tailing.

**[00824]** The modified mRNAs may be modified to reduce the cellular innate immune response. The modifications to reduce the cellular response may include pseudouridine ($\psi$) and 5-methyl-cytidine (5meC, 5mc or $m^5C$). (See, Kariko K et al. Immunity 23:165-75 (2005), Kariko K et al. Mol Ther 16:1833-40 (2008), Anderson BR et al. NAR (2010); herein incorporated by reference).

**[00825]** The ORF may also include various upstream or downstream additions (such as, but not limited to, β-globin, tags, etc.) may be ordered from an optimization service such as, but limited to, DNA2.0 (Menlo Park, CA) and may contain multiple cloning sites which may have XbaI recognition. Upon receipt of the construct, it may be reconstituted and transformed into chemically competent *E. coli*.

**[00826]** For the present invention, NEB DH5-alpha Competent *E. coli* are used. Transformations are performed according to NEB instructions using 100 ng of plasmid. The protocol is as follows:

**[00827]** Thaw a tube of NEB 5-alpha Competent *E. coli* cells on ice for 10 minutes.

**[00828]** 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.

**[00829]** Place the mixture on ice for 30 minutes. Do not mix.

**[00830]** Heat shock at 42°C for exactly 30 seconds. Do not mix.

**[00831]** Place on ice for 5 minutes. Do not mix.

**[00832]** Pipette 950 µl of room temperature SOC into the mixture.

**[00833]** Place at 37°C for 60 minutes. Shake vigorously (250 rpm) or rotate.

**[00834]** Warm selection plates to 37°C.

**[00835]** Mix the cells thoroughly by flicking the tube and inverting.

**[00836]** 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.

**[00837]** 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.

**[00838]** 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.

**[00839]** In order to generate cDNA for *In Vitro* Transcription (IVT), the plasmid (an Example of which is shown in Figure 3) is first linearized using a restriction enzyme such as XbaI. A typical restriction digest with XbaI will comprise the following: Plasmid 1.0 µg; 10x Buffer 1.0 µl; XbaI 1.5 µl; $dH_20$ 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.

**B. Agarose Gel Electrophoresis of modified mRNA**

**[00840]** Individual modified mRNAs (200-400 ng in a 20 µl volume) 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.

**C. Agarose Gel Electrophoresis of RT-PCR products**

**[00841]** Individual reverse transcribed-PCR products (200-400ng) are loaded into a well of a non-denaturing 1.2% Agarose E-Gel (Invitrogen, Carlsbad, CA) and run for 12-15 minutes according to the manufacturer protocol.

**D. Nanodrop modified mRNA quantification and UV spectral data**

**[00842]** Modified mRNAs in TE buffer (1 µl) are used for Nanodrop UV absorbance readings to quantitate the yield of each modified mRNA from an *in vitro* transcription reaction (UV absorbance traces are not shown).

**Example 2. Modified mRNA Transfection**

**A. Reverse Transfection**

**[00843]** For experiments performed in a 24-well collagen-coated tissue culture plate, Keratinocytes are seeded at a cell density of $1 \times 10^5$. For experiments performed in a 96-well collagen-coated tissue culture plate, Keratinocytes are seeded at a cell density of $0.5 \times 10^5$. For each modified mRNA to be transfected, modified mRNA: RNAIMAX™ are prepared as described and mixed with the cells in the multi-well plate within 6 hours of cell seeding before cells had adhered to the tissue culture plate.

**B. Forward Transfection**

**[00844]** In a 24-well collagen-coated tissue culture plate, Keratinocytes are seeded at a cell density of 0.7 x $10^5$. For experiments performed in a 96-well collagen-coated tissue culture plate, Keratinocytes are seeded at a cell density of 0.3 x $10^5$. Keratinocytes are then grown to a confluency of >70% for over 24 hours. For each modified mRNA to be transfected, modified mRNA: RNAIMAX™ are prepared as described and transfected onto the cells in the multi-well plate over 24 hours after cell seeding and adherence to the tissue culture plate.

**C. Modified mRNA Translation Screen: G-CSF ELISA**

**[00845]** Keratinocytes are grown in EpiLife medium with Supplement S7 from Invitrogen at a confluence of >70%. Keratinocytes are reverse transfected with 300 ng of the indicated chemically modified mRNA complexed with RNAIMAX™ from Invitrogen. Alternatively, keratinocytes are forward transfected with 300 ng modified mRNA complexed with RNAIMAX™ from Invitrogen. The RNA: RNAIMAX™ complex is formed by first incubating the RNA with Supplement-free EPILIFE® media in a 5X volumetric dilution for 10 minutes at room temperature.

**[00846]** In a second vial, RNAIMAX™ reagent is incubated with Supplement-free EPILIFE® Media in a 10X volumetric dilution for 10 minutes at room temperature. The RNA vial is then mixed with the RNAIMAX™ vial and incubated for 20-30 at room temperature before being added to the cells in a drop-wise fashion. Secreted huG-CSF concentration in the culture medium is measured at 18 hours post-transfection for each of the chemically modified mRNAs in triplicate. Secretion of Human Granulocyte-Colony Stimulating Factor (G-CSF) from transfected human keratinocytes is quantified using an ELISA kit from Invitrogen or R&D Systems (Minneapolis, MN) following the manufacturers recommended instructions.

**D. Modified mRNA Dose and Duration: G-CSF ELISA**

**[00847]** Keratinocytes are grown in EPILIFE® medium with Supplement S7 from Invitrogen at a confluence of >70%. Keratinocytes are reverse transfected with 0ng, 46.875ng, 93.75ng, 187.5ng, 375ng, 750ng, or 1500ng modified mRNA complexed with RNAIMAX™ from Invitrogen. The modified mRNA: RNAIMAX™ complex is formed as described. Secreted huG-CSF concentration in the culture medium is measured at 0, 6, 12, 24, and 48 hours post-transfection for each concentration of each modified mRNA in triplicate. Secretion of Human Granulocyte-Colony Stimulating Factor (G-CSF) from transfected human keratinocytes is quantified using an ELISA kit from Invitrogen or R&D Systems following the manufacturers recommended instructions.

**Example 3. Cellular Innate Immune Response to Modified Nucleic Acids: IFN-beta ELISA and TNF-alpha ELISA**

**[00848]** An enzyme-linked immunosorbent assay (ELISA) for Human Tumor Necrosis Factor-α (TNF-α), Human Interferon-β (IFN-β) and Human Granulocyte-Colony Stimulating Factor (G-CSF) secreted from *in vitro*-transfected Human Keratinocyte cells is tested for the detection of a cellular innate immune response.

**[00849]** Keratinocytes are grown in EPILIFE® medium with Human Keratinocyte Growth Supplement in the absence of hydrocortisone from Invitrogen at a confluence of >70%. Keratinocytes are reverse transfected with 0ng, 93.75ng, 187.5ng, 375ng, 750ng, 1500ng or 3000ng of the indicated chemically modified mRNA complexed with RNAIMAX™ from Invitrogen as described in triplicate. Secreted TNF-α in the culture medium is measured 24 hours post-transfection for each of the chemically modified mRNAs using an ELISA kit from Invitrogen according to the manufacturer protocols.

**[00850]** Secreted IFN-β is measured 24 hours post-transfection for each of the chemically modified mRNAs using an ELISA kit from Invitrogen according to the manufacturer protocols. Secreted hu-G-CSF concentration is measured at 24 hours post-transfection for each of the chemically modified mRNAs. Secretion of Human Granulocyte-Colony Stimulating Factor (G-CSF) from transfected human keratinocytes is quantified using an ELISA kit from Invitrogen or R&D Systems (Minneapolis, MN) following the manufacturers recommended instructions. These data indicate which modified mRNA are capable eliciting a reduced cellular innate immune response in comparison to natural and other chemically modified polynucleotides or reference compounds by measuring exemplary type 1 cytokines TNF-alpha and IFN-beta.

**Example 4. Human Granulocyte-Colony Stimulating Factor-modified mRNA-induced Cell Proliferation Assay**

**[00851]** Human keratinocytes are grown in EPILIFE® medium with Supplement S7 from Invitrogen at a confluence of >70% in a 24-well collagen-coated TRANSWELL® (Corning, Lowell, MA) co-culture tissue culture plate. Keratinocytes are reverse transfected with 750ng of the indicated chemically modified mRNA complexed with RNAIMAX™ from Invitrogen as described in triplicate. The modified mRNA: RNAIMAX™ complex is formed as described. Keratinocyte media is exchanged 6-8 hours post-transfection. 42-hours post-transfection, the 24-well

TRANSWELL® plate insert with a 0.4μm-pore semi-permeable polyester membrane is placed into the hu-G-CSF modified mRNA-transfected keratinocyte containing culture plate.

**[00852]** Human myeloblast cells, Kasumi-1 cells or KG-1 ($0.2 \times 10^5$ cells), are seeded into the insert well and cell proliferation is quantified 42 hours post-co-culture initiation using the CyQuant Direct Cell Proliferation Assay (Invitrogen) in a 100-120 μl volume in a 96-well plate. modified mRNA-encoding hu-G-CSF-induced myeloblast cell proliferation is expressed as a percent cell proliferation normalized to untransfected keratinocyte/myeloblast co-culture control wells. Secreted hu-G-CSF concentration in both the keratinocyte and myeloblast insert co-culture wells is measured at 42 hours post-co-culture initiation for each modified mRNA in duplicate. Secretion of Human Granulocyte-Colony Stimulating Factor (G-CSF) is quantified using an ELISA kit from Invitrogen following the manufacturers recommended instructions.

**[00853]** Transfected hu-G-CSF modified mRNA in human keratinocyte feeder cells and untransfected human myeloblast cells are detected by RT-PCR. Total RNA from sample cells is extracted and lysed using RNAEASY® kit (Qiagen, Valencia, CA) according to the manufacturer instructions. Extracted total RNA is submitted to RT-PCR for specific amplification of modified mRNA-G-CSF using PROTOSCRIPT® M-MuLV *Taq* RT-PCR kit (New England BioLabs, Ipswich, MA) according to the manufacturer instructions with hu-G-CSF-specific primers. RT-PCR products are visualized by 1.2% agarose gel electrophoresis.

**Example 5. Cytotoxicity and Apoptosis**

**[00854]** This experiment demonstrates cellular viability, cytotoxity and apoptosis for distinct modified mRNA-*in vitro* transfected Human Keratinocyte cells. Keratinocytes are grown in EPILIFE® medium with Human Keratinocyte Growth Supplement in the absence of hydrocortisone from Invitrogen at a confluence of >70%. Keratinocytes are reverse transfected with 0ng, 46.875ng, 93.75ng, 187.5ng, 375ng, 750ng, 1500ng, 3000ng, or 6000ng of modified mRNA complexed with RNAIMAX™ from Invitrogen. The modified mRNA: RNAIMAX™ complex is formed. Secreted huG-CSF concentration in the culture medium is measured at 0, 6, 12, 24, and 48 hours post-transfection for each concentration of each modified mRNA in triplicate. Secretion of Human Granulocyte-Colony Stimulating Factor (G-CSF) from transfected human keratinocytes is quantified using an ELISA kit from Invitrogen or R&D Systems following the manufacturers recommended instructions. Cellular viability, cytotoxicity and apoptosis is measured at 0, 12, 48, 96, and 192

hours post-transfection using the APOTOX-GLO™ kit from Promega (Madison, WI) according to manufacturer instructions.

**Example 6. Co-Culture Environment**

**[00855]** The modified mRNA comprised of chemically-distinct modified nucleotides encoding human Granulocyte-Colony Stimulating Factor (G-CSF) may stimulate the cellular proliferation of a transfection incompetent cell in co-culture environment. The co-culture includes a highly transfectable cell type such as a human keratinocyte and a transfection incompetent cell type such as a white blood cell (WBC). The modified mRNA encoding G-CSF may be transfected into the highly transfectable cell allowing for the production and secretion of G-CSF protein into the extracellular environment where G-CSF acts in a paracrine-like manner to stimulate the white blood cell expressing the G-CSF receptor to proliferate. The expanded WBC population may be used to treat immune-compromised patients or partially reconstitute the WBC population of an immunosuppressed patient and thus reduce the risk of opportunistic infections.

**[00856]** In another example, a highly transfectable cell such as a fibroblast are transfected with certain growth factors to support and simulate the growth, maintenance, or differentiation of poorly transfectable embryonic stem cells or induced pluripotent stem cells.

**Example 7. 5'-Guanosine Capping on Modified Nucleic Acids (modified mRNAs)**

**A. Materials and Methods**

**[00857]** The cloning, gene synthesis and vector sequencing was performed by DNA2.0 Inc. (Menlo Park, CA). The ORF was restriction digested using XbaI and used for cDNA synthesis using tailed-or tail-less-PCR. The tailed-PCR cDNA product was used as the template for the modified mRNA synthesis reaction using 25mM each modified nucleotide mix (all modified nucleotides were custom synthesized or purchased from TriLink Biotech, San Diego, CA except pyrrolo-C triphosphate purchased from Glen Research, Sterling VA; unmodifed nucleotides were purchased from Epicenter Biotechnologies, Madison, WI) and CellScript MEGASCRIPT™ (Epicenter Biotechnologies, Madison, WI) complete mRNA synthesis kit. The *in vitro* transcription reaction was run for 4 hours at 37°C. Modified mRNAs incorporating adenosine analogs were poly (A) tailed using yeast Poly (A) Polymerase (Affymetrix, Santa Clara, CA). PCR reaction used HiFi PCR 2X MASTER MIX™ (Kapa Biosystems, Woburn, MA). Modified mRNAs were post-transcriptionally capped using recombinant Vaccinia Virus Capping Enzyme (New England BioLabs, Ipswich, MA) and a recombinant 2'-o-methyltransferase (Epicenter Biotechnologies,

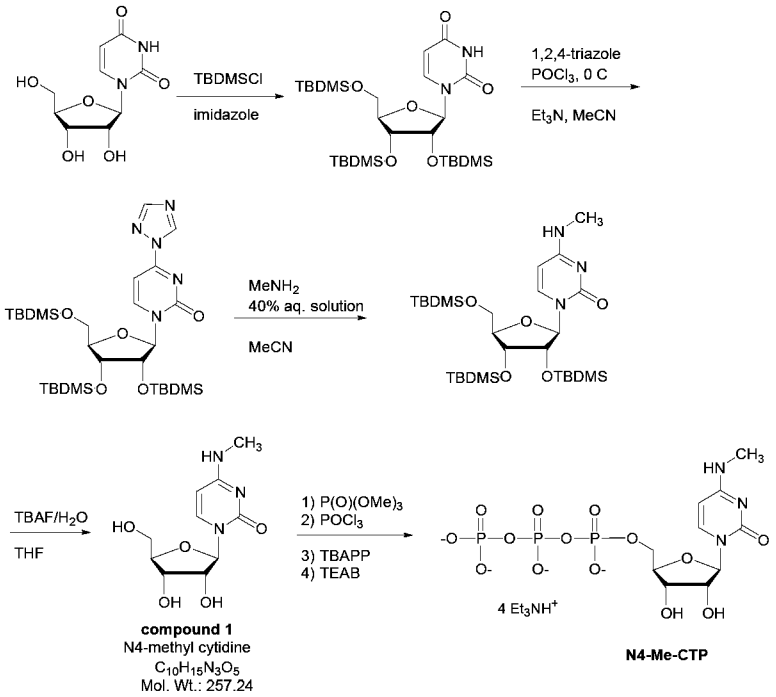
Madison, WI) to generate the 5'-guanosine Cap1 structure. Cap 2 structure and Cap 2 structures may be generated using additional 2'-o-methyltransferases. The *In vitro* transcribed mRNA product was run on an agarose gel and visualized. Modified mRNA was purified with Ambion/Applied Biosystems (Austin, TX) MEGAClear RNA™ purification kit. PCR used PURELINK™ PCR purification kit (Invitrogen, Carlsbad, CA). The product was quantified on NANODROP™ UV Absorbance (ThermoFisher, Waltham, MA). Quality, UV absorbance quality and visualization of the product was performed on an 1.2% agarose gel. The product was resuspended in TE buffer.

**B. 5' Capping Modified Nucleic Acid (mRNA) Structure**

**[00858]** 5'-capping of modified mRNA 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-$m^7$G(5')ppp(5')G (the ARCA cap); G(5')ppp(5')A; G(5')ppp(5')G; $m^7$G(5')ppp(5')A; $m^7$G(5')ppp(5')G (New England BioLabs, Ipswich, MA). 5'-capping of modified mRNA may be completed post-transcriptionally using a Vaccinia Virus Capping Enzyme to generate the "Cap 0" structure: $m^7$G(5')ppp(5')G (New England BioLabs, Ipswich, MA). Cap 1 structure may be generated using both Vaccinia Virus Capping Enzyme and a 2'-O methyl-transferase to generate: m7G(5')ppp(5')G-2'-O-methyl. Cap 2 structure may be generated from the Cap 1 structure followed by the 2'-o-methylation of the 5'-antepenultimate nucleotide using a 2'-O 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'-O methyl-transferase. Enzymes are preferably derived from a recombinant source.

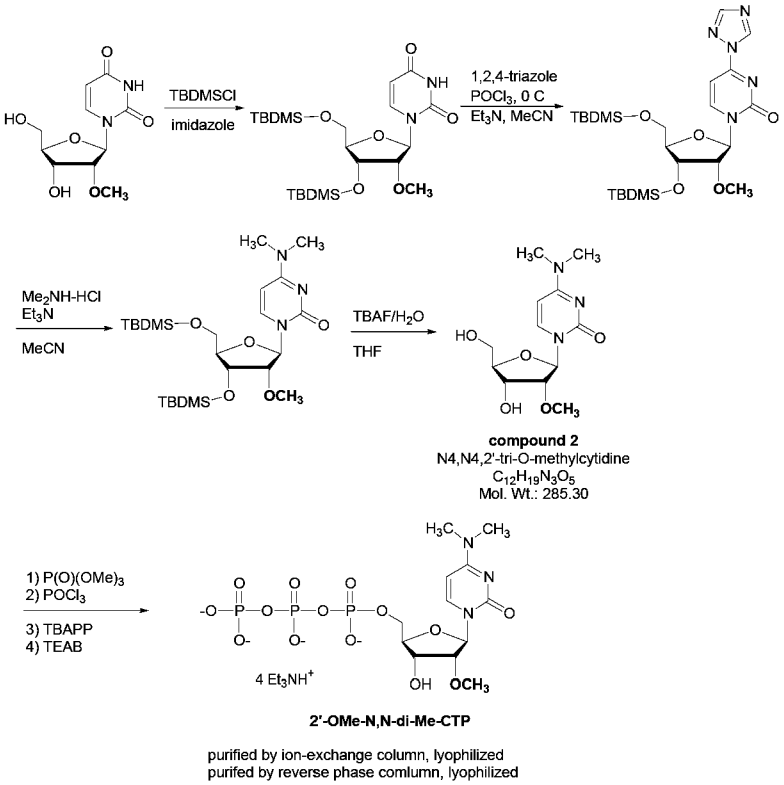
**[00859]** When transfected into mammalian cells, the modified mRNAs have a stability of 12-18 hours or more than 18 hours, e.g., 24, 36, 48, 60, 72 or greater than 72 hours.

**Example 8. Synthesis of N4-methyl cytidine (compound 1) and N4-methyl CTP (NTP of said compound)**




**[00860]** Uridine was silylated to provide a trisilylated compound, which was purified by column, activated with re-distilled $POCl_3$/triazole under anhydrous condition, and then followed by nucleophilic substitution with 40% methylamine aqueous solution. N4-Methyl-2',3',5'-tri-O-TBDMS-cytidine was thus obtained after chromatographic purification. The resultant product was deprotected with TBAF and then purified with an ethanol-ethyl acetate (3:1) solvent system to obtain compound 1. The final product was characterized by NMR (in DMSO); MS: 258 $(M + H)^+$, 280 $(M + Na)^+$, and 296 $(M + K)^+$; and HPLC: purity, 99.35% (FIGS. 1A-1D). HPLC, purity 98% (FIG. 2).

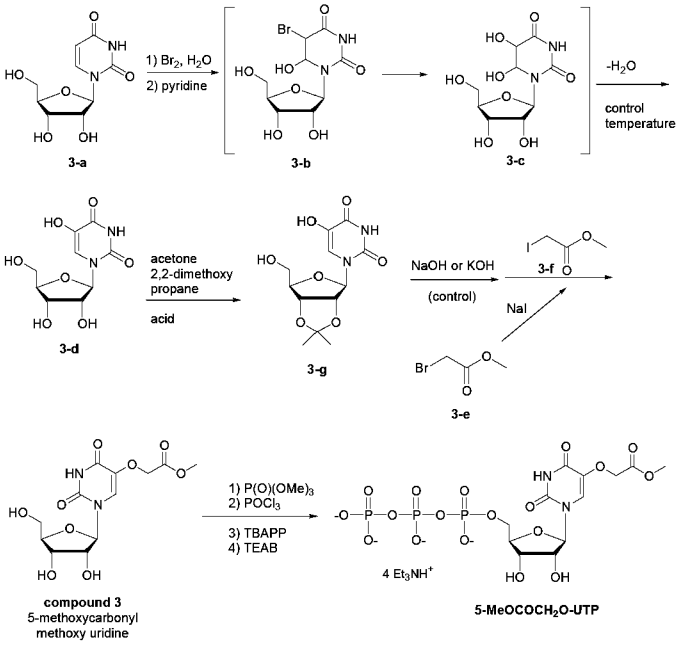
**Example 9. Synthesis of 2'-OMe-N,N-di-Me-cytidine (compound 2) and 2'-OMe-N,N-di-Me-CTP (NTP of said compound)**




**[00861]** 2'-O-Methyluridine was silylated to give the di-silylated compound. Purified 2'-O-methyl-3',5'-di-O-TBDMS uridine was activated with re-distilled $POCl_3$ and imidazole under anhydrous condition, followed by the nucleophilic substitution with dimethylamine hydrochloride under triethylamine environment to trap HCl. Intermediate compound N4,N4,2'-tri-O-methyl-3',5'-bis-O-TBDMS uridine was purified by flash chromatography and obtained as a white foam. The resultant compound was de-protected with TBAF and then purified to provide ~400 mg final product compound 2 as white foam. ES MS: m/z 308 $(M + Na)^+$, 386 $(M + H)^+$; HPLC: purity, 99.49% (FIGS. 3A-3C).

**[00862]** To synthesize the corresponding NTP, 70 mg of nucleoside compound 2 provided 23 mg of 2'-OMe-N,N-di-Me-CTP after purification via ion-exchange and reverse phase columns. HPLC: purity, 95% (FIG. 4).

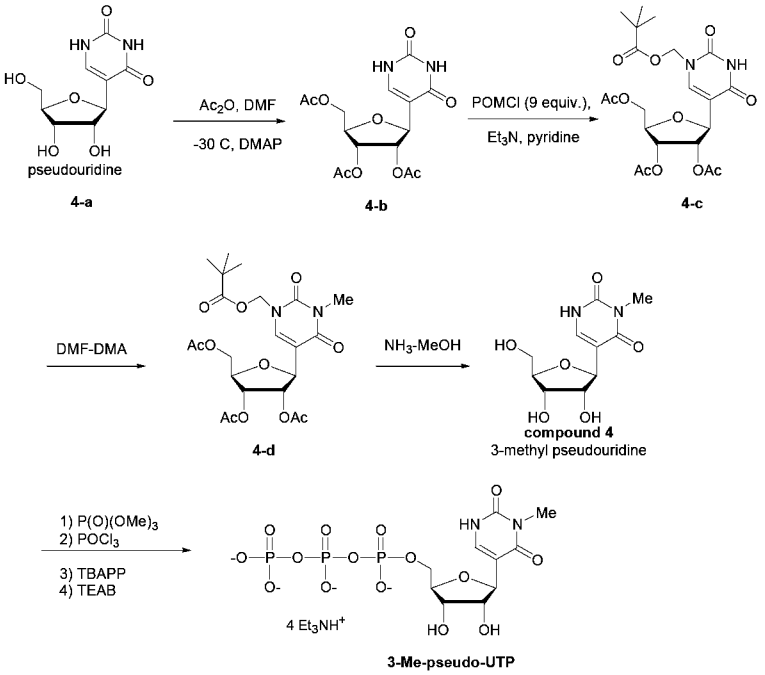
**Example 10. Synthesis of 5-methoxycarbonylmethoxy uridine (compound 3) and 5-methoxycarbonylmethoxy-UTP (NTP of said compound)**




**[00863]** Uridine 3-a in water was treated with excess amount of bromine and then flushed with air to remove bromine. The reaction mixture was treated with pyridine at a controlled speed and temperature. During the reaction, unstable bromo-intermediate 3-b gradually converted to di-hydroxyl intermediate 3-c, which presumably dehydrated to the stable 5-hydroxyuridine 3-d. Then, the 5-hydroxyuridine was protected with a 2',3'-isopropylidene group to provide compound 3-g. Reaction with compound 3-f provided compound 3.

**[00864]** 60-70 mg of the nucleoside provided >21 mg of the desired triphosphate after two HPLC column purification and two lyophilization steps. HPLC: purity, 98% (FIG. 5).

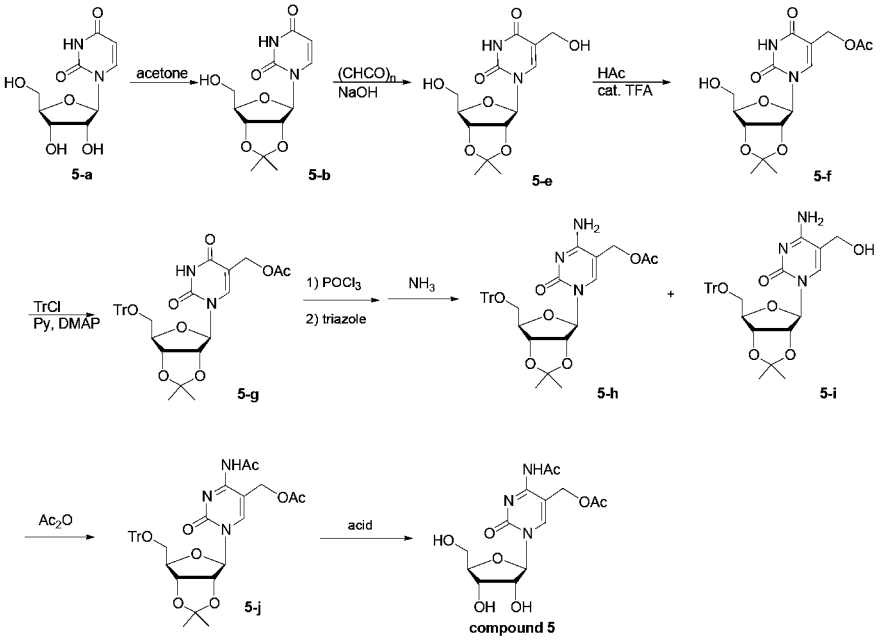
**Example 11. Synthesis of 3-methyl pseudouridine (compound 4) and 3-methyl pseudo-UTP (NTP of said compound)**




**[00865]** Pseudouridine 4-a was reacted with $Ac_2O$ to provide acetyl-protected pseudouridine 4-b. Then, N1 was selectively protected with POM to provide compound 4-c. Methylation of N3, followed by deprotected, provided compound 4 (~400 mg). Molecular formula: C10H14N2O6, molecular weight: 258.23g/mol; appearance: white solid; storage conditions: store at 25 °C; HPLC: purity, 98.51%; $^1$H NMR (DMSO-$d_6$): δ 11.17 (d, 1H, $J$ = 3.0 Hz), 7.56 (d, 1H, $J$ = 3.6 Hz), 4.91 (d, 1H, $J$ = 3.6 Hz), 4.79 (t, 1H, $J$ = 4.2Hz), 4.70 (d, 1H, $J$ = 4.2Hz), 4.49 (d, 1H, $J$ = 3.0Hz), 3.82-3.88 (m, 2H), 3.66-3.67 (m, 1H), 3.57-3.61 (m, 1H), 3.40-3.47 (m, 1H), 3.09 (s, 3H); MS: 281 $(M + Na)^+$) (FIGS. 6A and 6B).

**[00866]** Alternative routes could be applied to obtain compound 4. For example, pseudouridine could be reacted with an *O*-protecting group (e.g., as described herein, such as TMS) and reacted with an *N*-protecting group (e.g., as described herein, such as acetyl at N1). Then, N3 of the nucleobase could be reacted with an alkylating agent (e.g., dimethylamine/dimethoxymethyl) to provide compound 4 having *N*- and *O*-protecting groups. Finally, the resultant compound would be deprotected (e.g., under basic conditions, such as $NH_3$/MeOH) to provide compound 4.

**Example 12. Synthesis of N-Ac, 5-Ac-$OCH_2$-cytidine (compound 5)**




**[00867]** Uridine 5-a was protected to obtain isopropylidene compound 5-b, which was reacted with $(CHCO)_n$. Acetic acid with catalyst amount of TFA was employed to obtain the desired selectively acylated compound 5-f (30% yield). Further tritylation of the 5'-OH group resulted in the desired orthogonally protected compound 5-g.

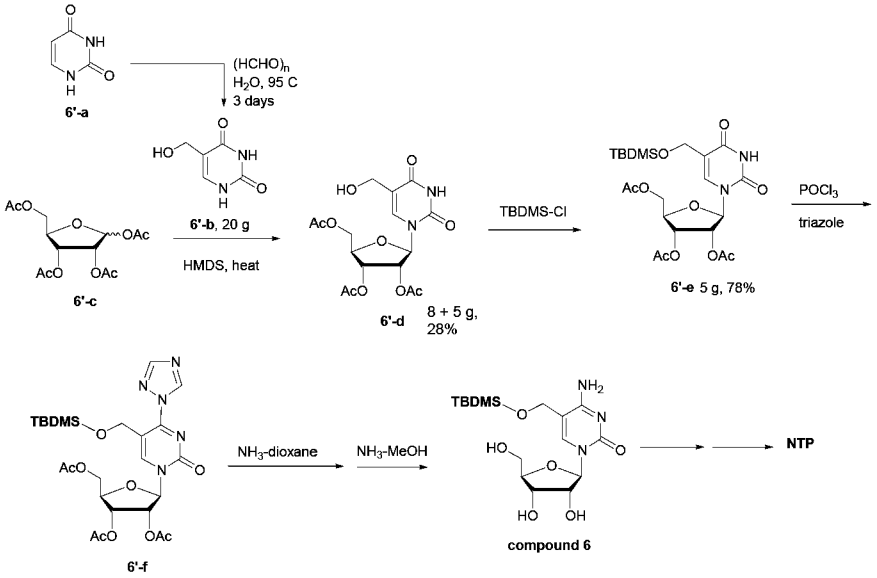
**[00868]** Compound 5-g was treated with $POCl_3$ and triazole to provide compound 5-h together with de-acylated compound 5-i. Acetylation of these two compounds provided di-acylated, fully protected compound 5-j. Deprotection of compound 5-j with acetic acid under heating condition resulted in three products, one of which was compound 5.

**[00869]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**[00870]** Alternative routes could be applied to obtain compound 5, such as by beginning with cytidine as the starting material. In such methods, the 5-position could be reacted with a halogen or a halogenation agent (e.g., any described herein, such as $I_2$/ meta-chloroperoxybenzoic acid), which

can be displaced with an alkylating agent. Further, such methods could include the use of one or more *N*- or *O*-protecting groups (e.g., any described herein, such as silylation or acetylation) to protect the amino group of cytidine and/or hydroxyl groups of the sugar moiety.

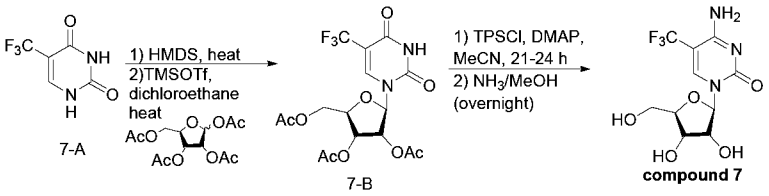
**Example 13. Synthesis of 5-TBDMS-$OCH_2$-cytidine (compound 6)**




**[00871]** A 5-hydroxyuracil compound '-b was glycosylated to obtain compound 6'-d (28% yield), which was silylated to provide compound 6'-e. Activation of the protected uridine provided the desired compound 6 after further amination and deprotection (800 mg of the final compound). Molecular formula: C16H29N3O6Si; molecular weight: 387.50 g/mol; appearance: white solid; storage conditions: store at 25 °C; HPLC: purity, 97.57%; $^1$H NMR ($CDCl_3$): d 7.81 (s, 1H), 7.40 (bs, 1H), 6.49 (bs, 1H), 5.79 (d, 1H, $J$ = 2.4 Hz), 5.3–5.32 (m, 1H), 5.00–5.07 (m, 2H), 4.30–4.45 (m, 2H), 3.90–3.94 (m, 2H), 3.80–3.83 (m, 1H), 3.50–3.70 (m, 2H), 0.87 (s, 9H), 0.05 (S, 6H); MS: 388 $(M + H)^+$, 410 $(M + Na)^+$) (FIGS. 7A-7C).

**[00872]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

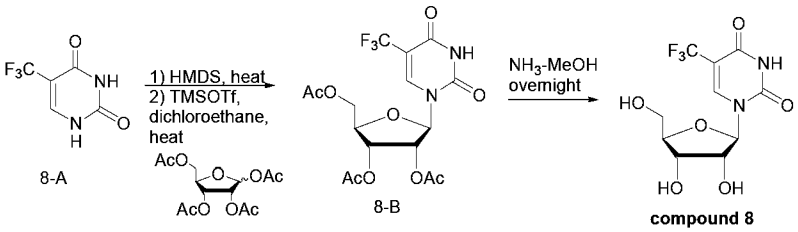
**Example 14. Synthesis of 5-trifluoromethyl cytidine (compound 7)**




**[00873]** Compound 7-A was glycosylated to provide compound 7-B, which was treated with 2,4,6-triisopropylbenzene sulfonyl chloride (TPSCl) to activate the carbonyl group and to promote reductive amination. Deprotection provided compound 7. Alternative activating agents could be used instead of TPSCl, such as 2,4,6-trimethylbenzene sulfonyl chloride.

**[00874]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

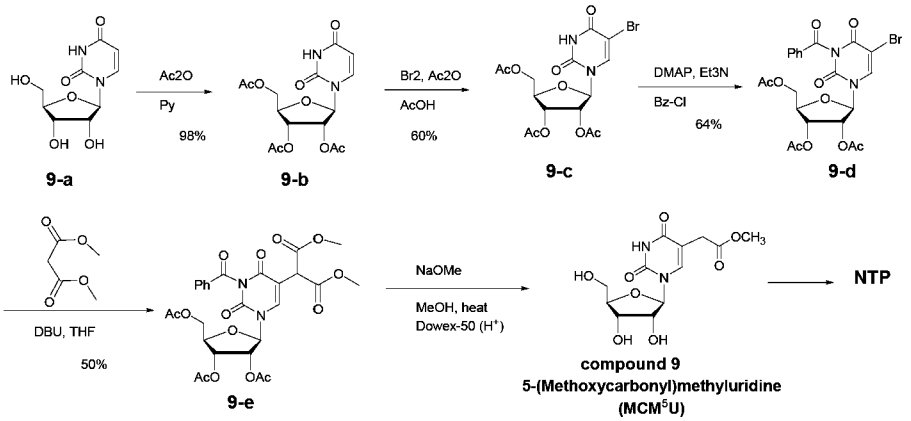
**Example 15. Synthesis of 5-trifluoromethyl uridine (compound 8)**




**[00875]** 5-Trifluoromethyluracil 8-A was glycosylated with tetra-O-acetyl ribose, and the desired triprotected 5-trifluoromethyluridine 8-B was obtained in good yield. Further deprotection gave desired compound 8, which was characterized with NMR, MS and HPLC results. MS: 313 $(M + H)^+$, 335 $(M + Na)^+$; HPLC: purity, 98.87%, ((FIGS 8A-8C).

**[00876]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

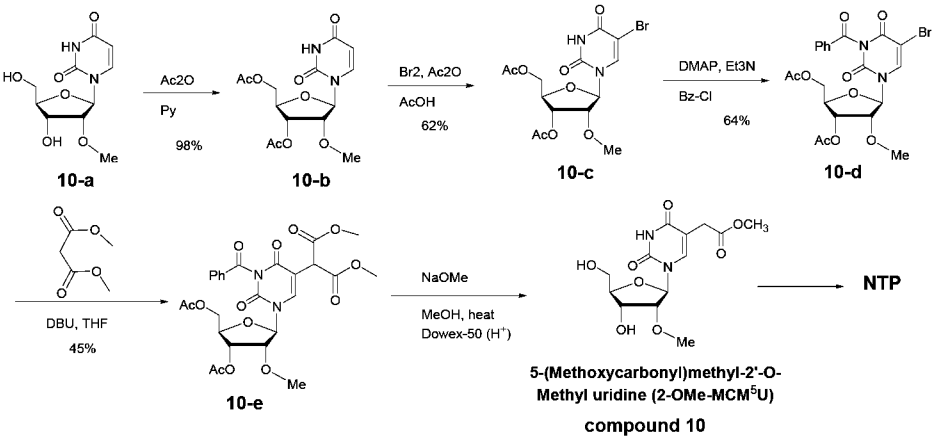
**Example 16. Synthesis of 5-(methoxycarbonyl)methyl uridine (compound 9)**




**[00877]** Uridine 9-a was protected to provide compound 9-b (98% yield). This compound was brominated with excess bromine in the presence of acetic anhydride and acetic acid. The 5-bromo analog 9-c was obtained (60% yield) and further benzoylated to provide desired compound 9-d (64% yield). 5-Bromo compound 9-d was condensed with dimethyl malonate under basic condition to give the arylated malonate and the fully protected diester 9-e (50% yield). After de-carboxylation and deprotection, compound 9 was obtained verified by NMR (FIG. 9).

**[00878]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

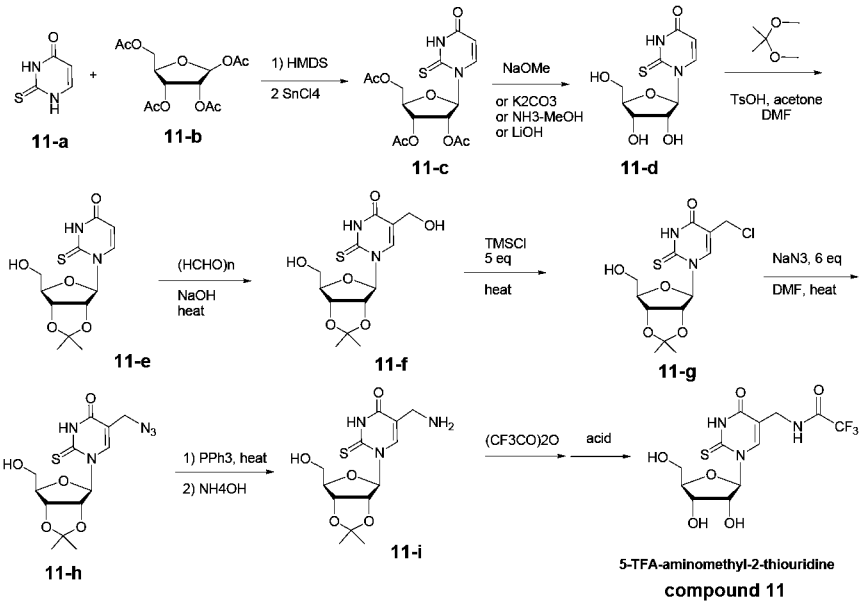
**Example 17. Synthesis of 5-(methoxycarbonyl)methyl-2'-O-methyl uridine (2-OMe-MCM5U) (compound 10)**




**[00879]** Similar strategy to the synthesis of compound 9 above, 2'-O-methyluridine 10-a was acylated and brominated to obtain compound 10-c. Further benzoylation provided 5-bromo analog 10-d, which was condensed with dimethyl malonate provide the desired product 10-e (45% yield). Decarboxylation and deprotection provided compound 10.

**[00880]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

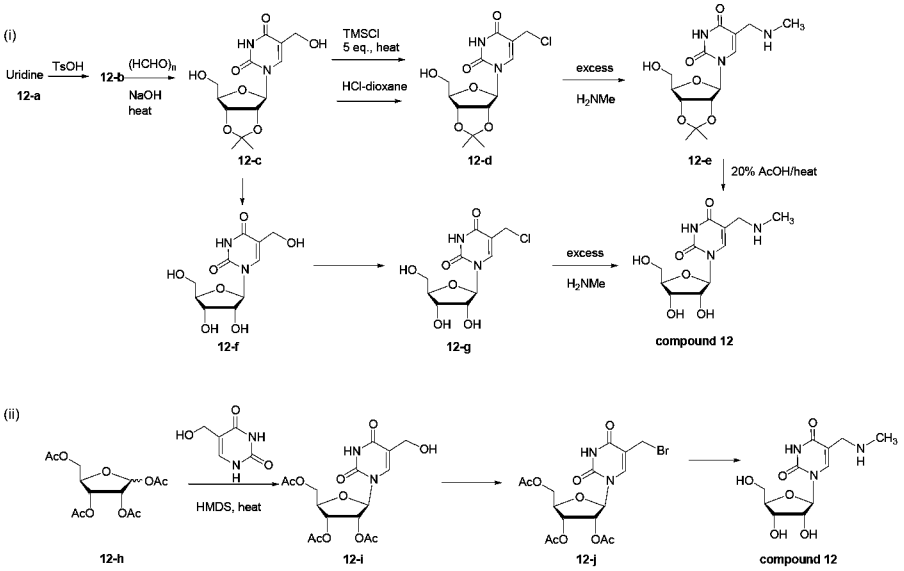
**Example 18. Synthesis of 5-trifluoroacetyl-aminomethyl-2-thiouridine (compound 11)**




**[00881]** Glycosylation of 2-thiouracil 11-a provided compound 11-c, which can be deprotected with any useful deprotection reagent. In particular, LiOH provided desired product 11-d (80-90% yield). Isopropylidene protection provided compound 11-e (90% yield). Further 5-hydroxylmethylation provided compound 11-f. Chlorination, azidation, and further reduction provided methylamine compound 11-i, which was acetylated to provided compound 11.

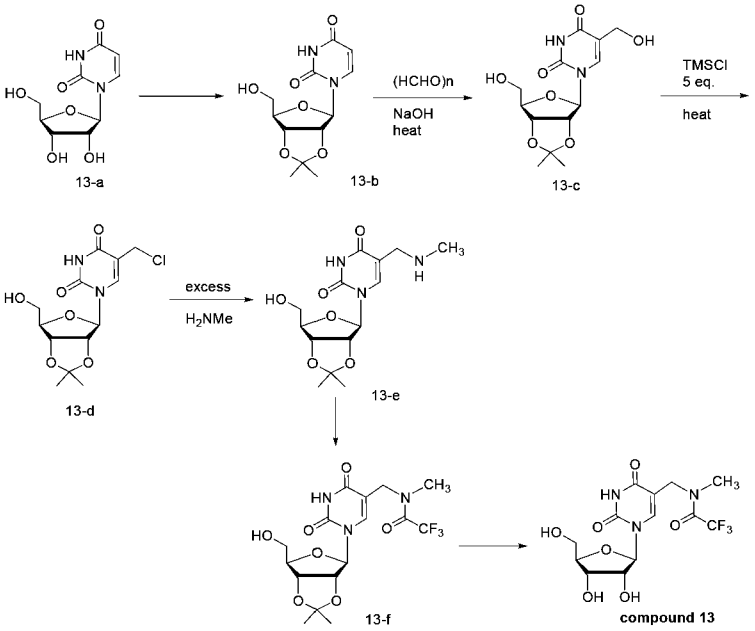
**[00882]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 19. Synthesis of 5-methylaminomethyl-2-uridine (compound 12)**




**[00883]** Compound 12 can be obtained by any useful method (e.g., see schemes (i) and (ii) above). For example, protected uracil can be glycosylated and subsequently aminated to provide compound 12. Additional protecting, deprotecting, and activating steps can be conducted as needed. To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

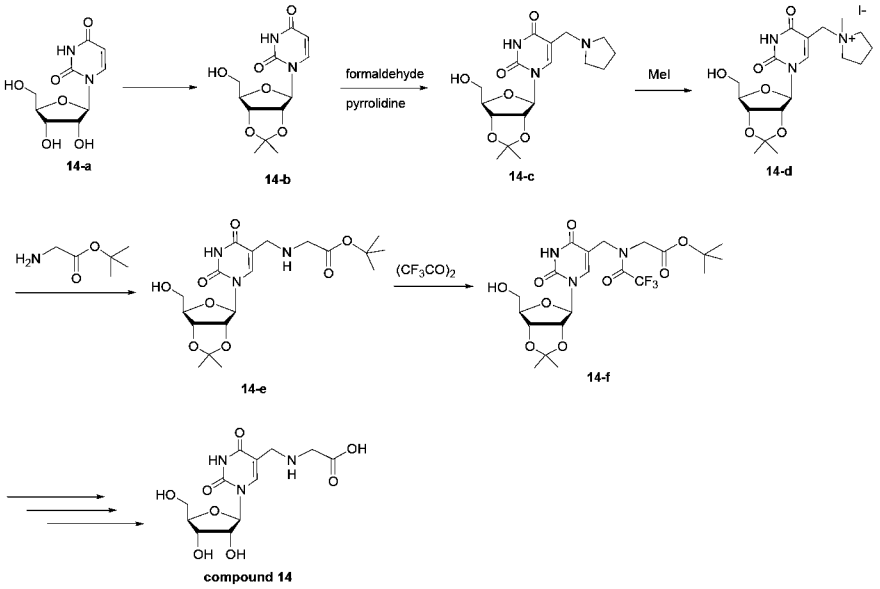
**Example 20. Synthesis of 5-TFA-methylaminomethyl-2-uridine (compound 13)**




**[00884]** Uridine 13-a was protected with isopropylidene to provide compound 13-b and then 5-hydroxymethylated to provide compound 13-c. Chlorination and subsequent amination provided compound 13-e, which can be protected to provided 13-f. Subsequent deprotection provided compound 13.

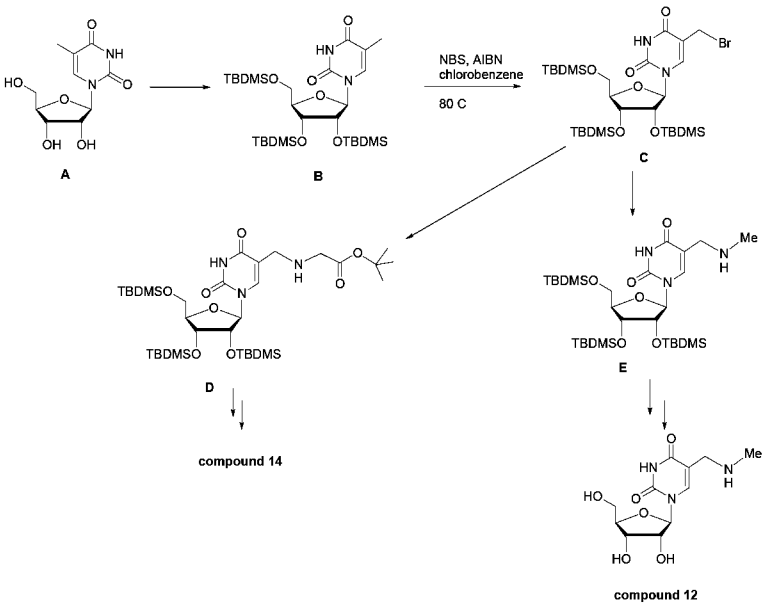
**[00885]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 21. Synthesis of 5-carboxymethylaminomethyl uridine (compound 14)**



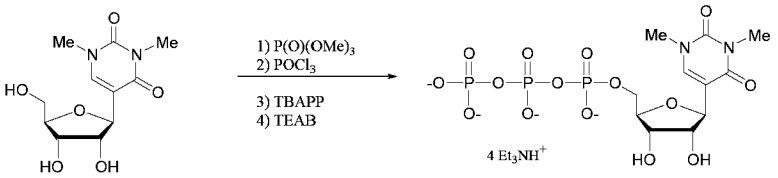

**[00886]** Uridine 14-a was protected with isopropylidene to provide compound 14-b and then 5-aminoalkylated with the Mannich reaction to provide compound 14-c. Methylation provided quaternary amine 14-d. Subsequent amination and deprotection steps can be used to provide compound 14. To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 22. Alternative synthesis of 5-methylaminomethyl-2-uridine (compound 12) and 5-carboxymethylaminomethyl-2-uridine (compound 14)**



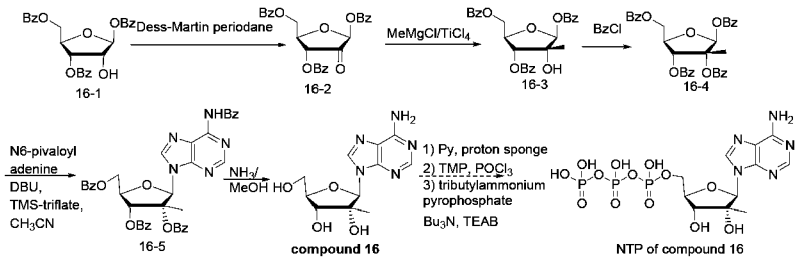

**[00887]** In addition to those strategies provided above for compounds 12 and 14, the following strategy can also be implemented. 5-Methyluridine A can be silylated to provide compound B. After radical monobromination, the resultant intermediate bromide C can be used for the preparation of compound 12 and compound 14 analogs. Subsequent alkylamination of bromide compound C could provide compounds D and E, which can be deprotected to provide compounds 14 and 12, respectively. To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 23. Synthesis of dimethyl-pseudouridine (compound 15) and dimethyl-pseudo-UTP (NTP of said compound)**



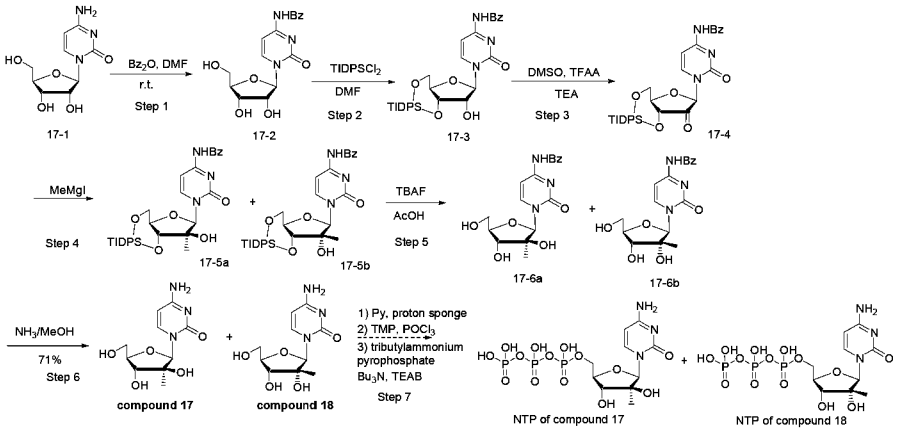

**[00888]** Nucleosides can be phosphorylated by any useful method. For example, as shown above, nucleosides can be reacted with phosphorus oxychloride and subsequently treated with a monophosphate intermediate with bis(tributylammonium)pyrophosphate (TBAPP) to give the triphosphate.

**Example 24. Synthesis of 2'-C-methyl adenosine (compound 16) and 2'-C-methyl ATP (NTP of said compound)**



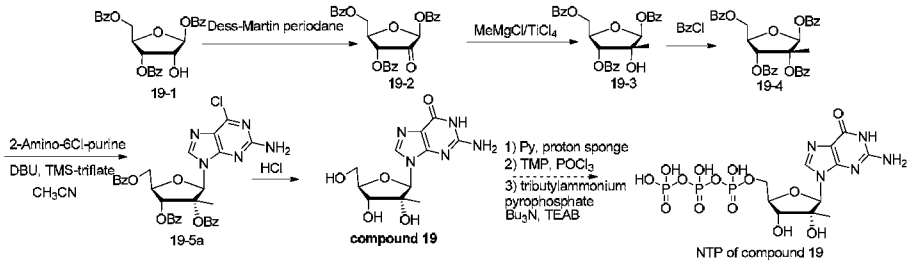

**[00889]** About 5g of compound 16-2 was prepared from 5g of compound 16-1 via a Dess-Martin periodane reaction. Compound 16-2 was reacted with MeMgI/TiCl4/-78°C to provide compound 16-3, and crude compound 16-3 (6g) was directly reacted with benzylchloride to prepare compound 16-4. Reaction with the nucleobase and deprotection provided compound 16 (0.56g).

**Example 25. Synthesis of 2'-C-methyl-cytidine isomers (compound 17 and compound 18) and 2'-C-methyl UTP (NTP of said compounds)**




**[00890]** About 17.4g of compound 17-3 was prepared from 20g of compound 17-1. Then, 2'-oxidation and alkylation with MeMgI provided 300mg of compound 17-5a and 80mg of compound 17-5b. About 9g of compound 17-5a (about 90% pure) and 2.1g of compound 17-5b (pure) were prepared from 17.4g of compound 17-3 in 2 batches. N- and O-deprotection provided compounds 17 and 18.

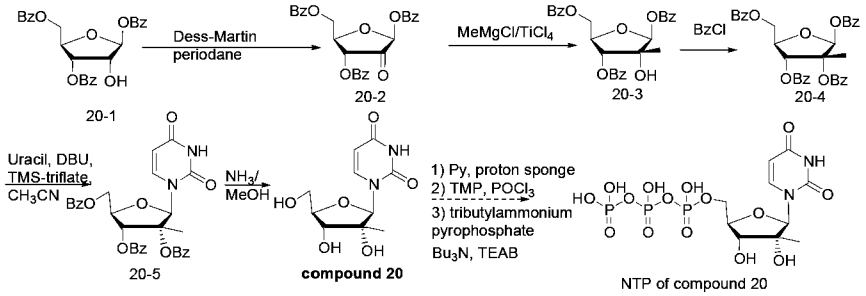
**Example 26. Synthesis of 2'-C-methyl guanosine (compound 19) and 2'-C-methyl GTP (NTP of said compound)**




**[00891]** 2'-Oxidation of protected ribose 19-1 and subsequent alkylation with MeMgCl provided compound 19-3. The resultant compound was further protected to provided compound 19-4, and 1.56g of compound 19-5a was prepared from 3.1 g of compound 19-4. Subsequent oidation and deprotection provided compound 19 (about 90% pure, 50 mg).

**[00892]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

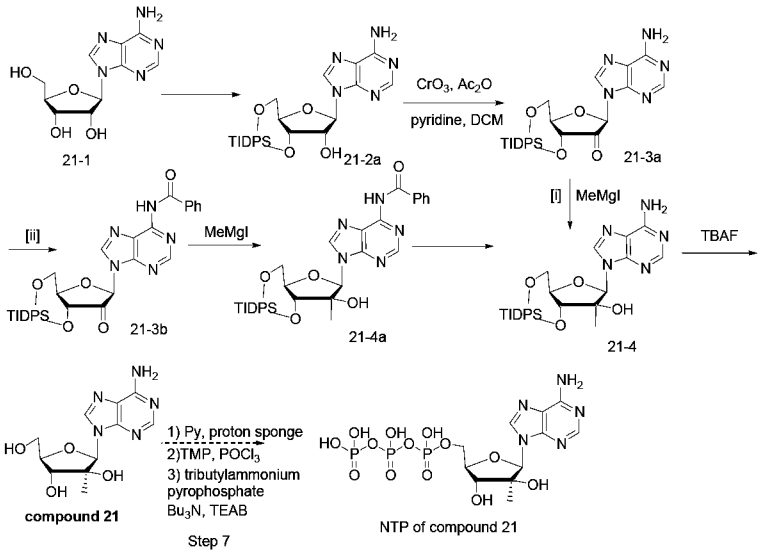
**Example 27. Synthesis of 2'-C-methyl uridine (compound 20) and 2'-C-methyl UTP (NTP of said compound)**




**[00893]** 2'-Oxidation of protected ribose 20-1 and subsequent alkylation with MeMgCl provided compound 20-3. The resultant compound was further protected to provide compound 20-4. Reaction with uracil and deprotection provided pure compound 20 (50mg).

**[00894]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

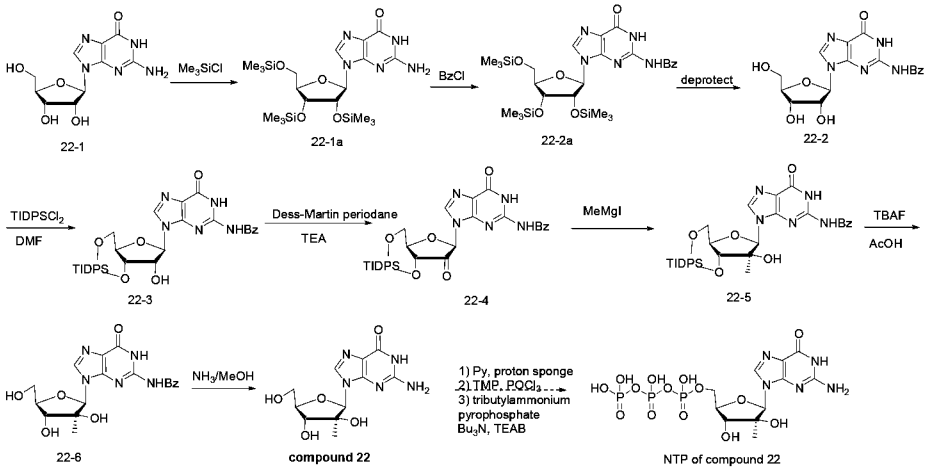
**Example 28. Synthesis of (S)-2'-C-methyl adenosine (compound 21) and (S)-2'-C-methyl ATP (NTP of said compound)**




**[00895]** Compound 21-1 (5g) was protected to form compound 21-2a, and chromium oxidation provided compound 21-3a. Alkylation via route [i] (5eq. MeMgI in ether at -50°C) provided compound 21-4. Optionally, yield could be improved via route [ii] by protecting the amino group to provide compound 21-3b and then alkylating at the 2'-C position to provide compound 21-4a. Compound 21-3a was alkylated to provide crude compound 21-4 (3g, 20% of compound 3a in this crude product), where the product can be optionally purified. Deprotection of compound 21-4 afforded compound 21 (50% yield).

**[00896]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

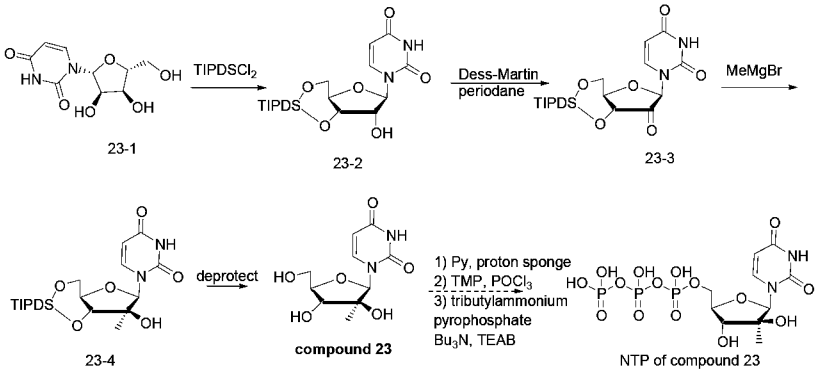
**Example 29. Synthesis of (S)-2'-C-methyl guanosine (compound 22) and (S)-2'-methyl GTP (NTP of said compound)**




**[00897]** About 30g of compound 22-1 was silylated to provide compound 22-2 in three steps. Further protection provided compound 22-3, and Dess-Martin periodane oxidation provided compound 22-4 (1.6g) in two batches. 2'-C alkylation (5eq. MeMgI in ether, -50°C to RT) provided compound 22-5, and further deprotection steps provided compound 22.

**[00898]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

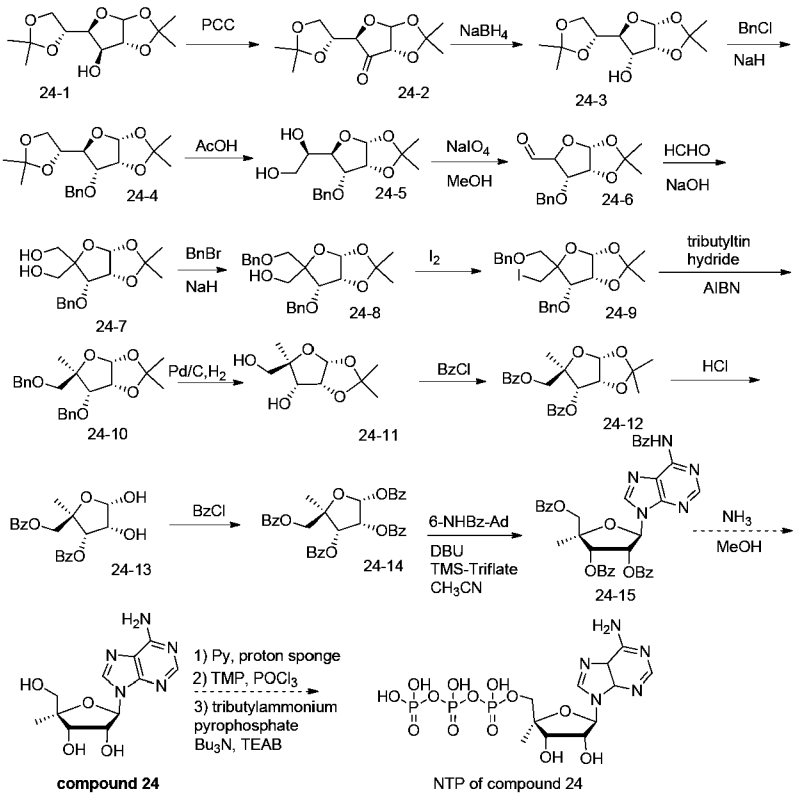
**Example 30. Synthesis of (S)-2'-C-methyl uridine (compound 23) and of (S)-2'-C-methyl UTP (NTP of said compound)**




**[00899]** Uridine 23-1 (2.0g) was protected with $TIPDSCl_2$ (1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane) to provide compound 23-2. Oxidation provided compound 23-3, and 2'-C alkylation provided compound 23-4, which can be optionally purified with Prep-HPLC prior to the next step. Then, deprotection provided desired compound 23.

**[00900]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

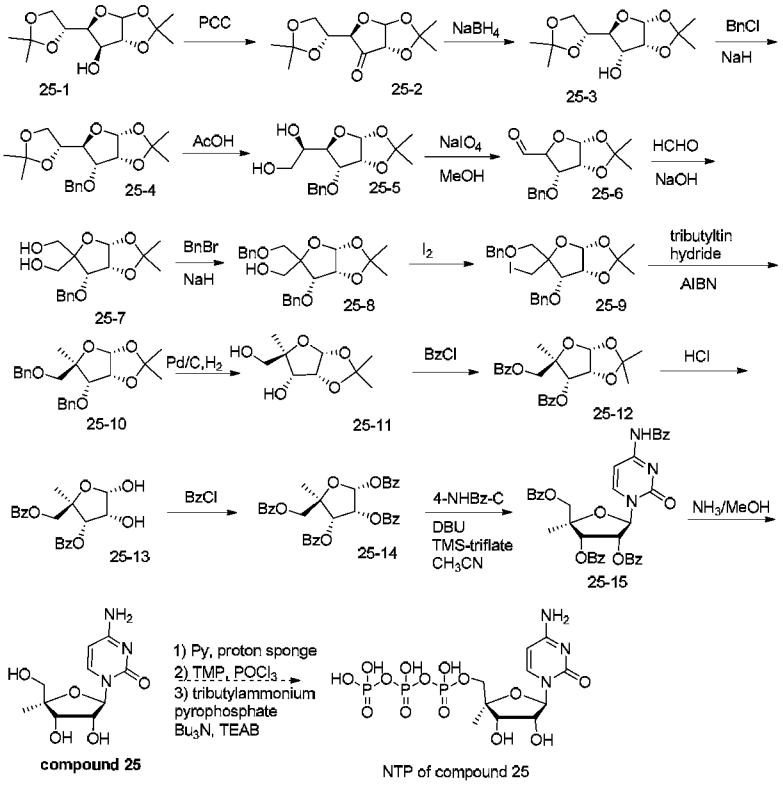
**Example 31. Synthesis of 4'-C-methyl adenosine (compound 24) and 4'-C-methyl ATP (NTP of said compound)**




**[00901]** 1,2:5,6-Di-O-isopropylidene-α-D-glucofuranose 24-1 was converted via sequential oxidation, reduction, and protection steps to provide compound 24-4. The first oxidation step to provide compound 24-2 can be implemented with any useful reagents, such as 0.75 eq. pyridinium dichromate (PDC) with 1 eq. $Ac_2O$ or 1.2 eq. of Dess-Martin periodane. Subsequent deprotection, formylation, and reduction provided compound 24-7, which was followed with protection and deoxygenation steps to provide compound 24-10. About 0.4g of compound 24-14 was prepared from 1g of compound 24-10 via sequential protection and deprotection steps. Addition of N6-benzoyladenine and subsequent deprotection provided compound 24.

**[00902]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

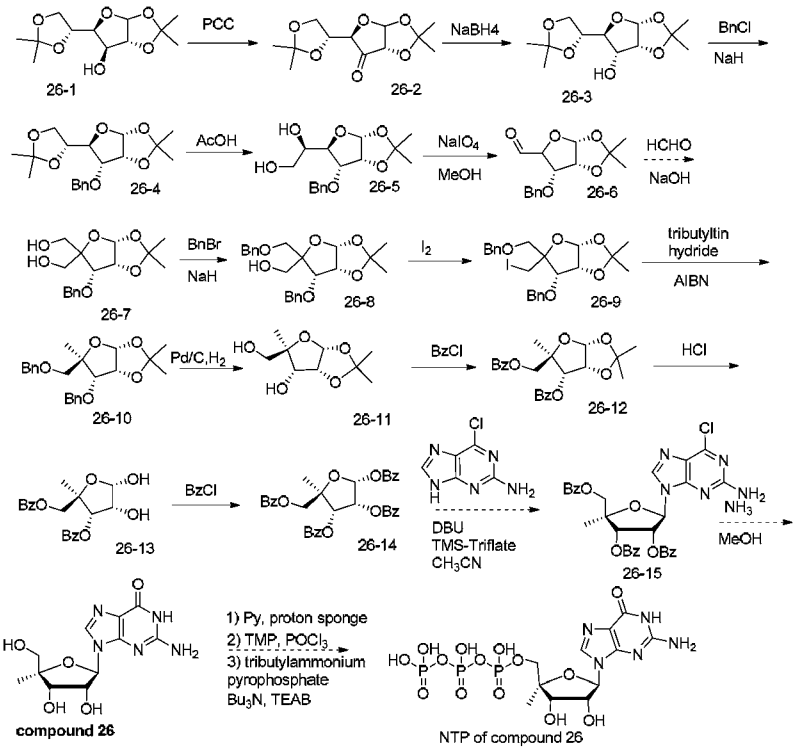
**Example 32. Synthesis of 4'-C-methyl cytidine (compound 25) and 4'-C-methyl CTP (NTP of said compound)**




**[00903]** Similar to the strategy provided above for compound 24, compound 25-14 was produced with compound 25-1. Addition of cytidine and subsequent deprotection provided compound 25.

**[00904]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

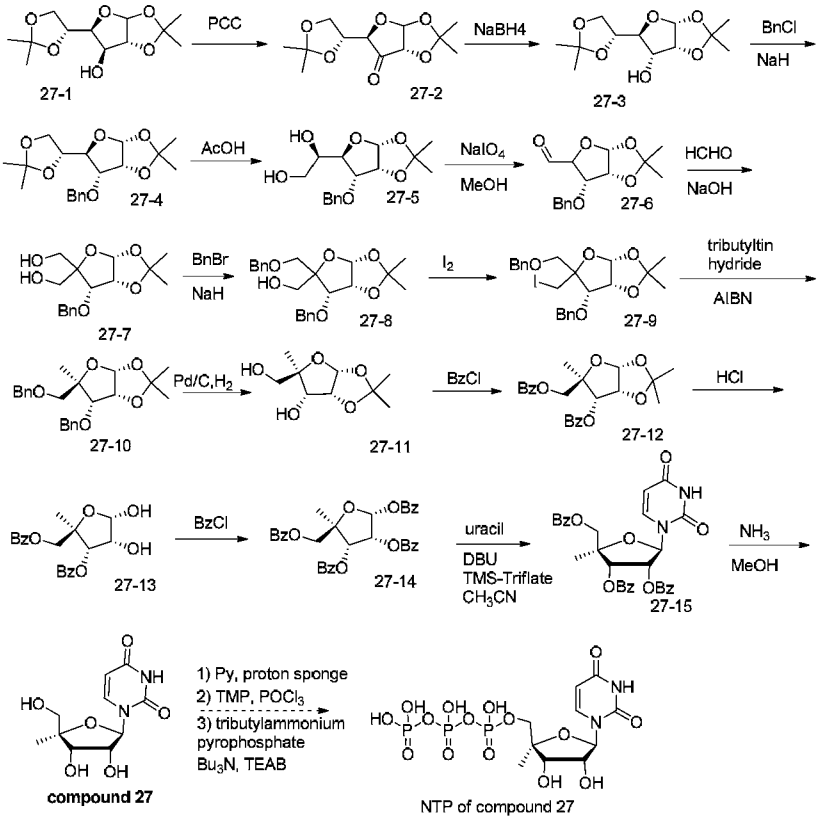
**Example 33. Synthesis of 4'-C-methyl guanosine (compound 26) and 4'-C-methyl GTP (NTP of said compound)**




**[00905]** Similar to the strategy provided above for compound 24, compound 26-14 was produced with compound 26-1. Addition of 2-amino-6-chloropurine, subsequent oxidation, and then deprotection provided compound 26.

**[00906]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

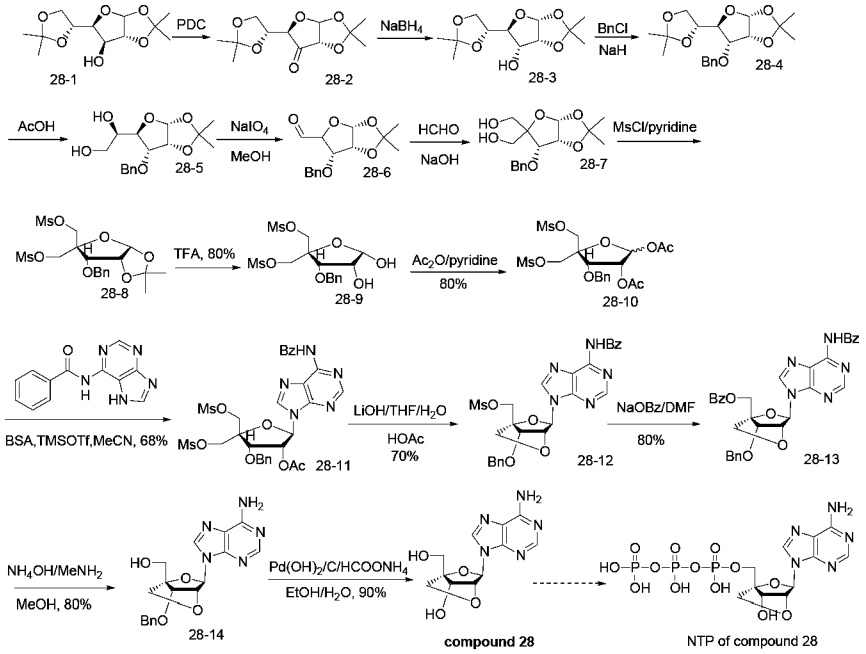
**Example 34. Synthesis of 4'-C-methyl uridine (compound 27) and 4'-C-methyl UTP (NTP of said compound)**




**[00907]** Similar to the strategy provided above for compound 24, compound 27-14 was produced with compound 27-1. Addition of uracil and subsequent deprotection provided compound 27.

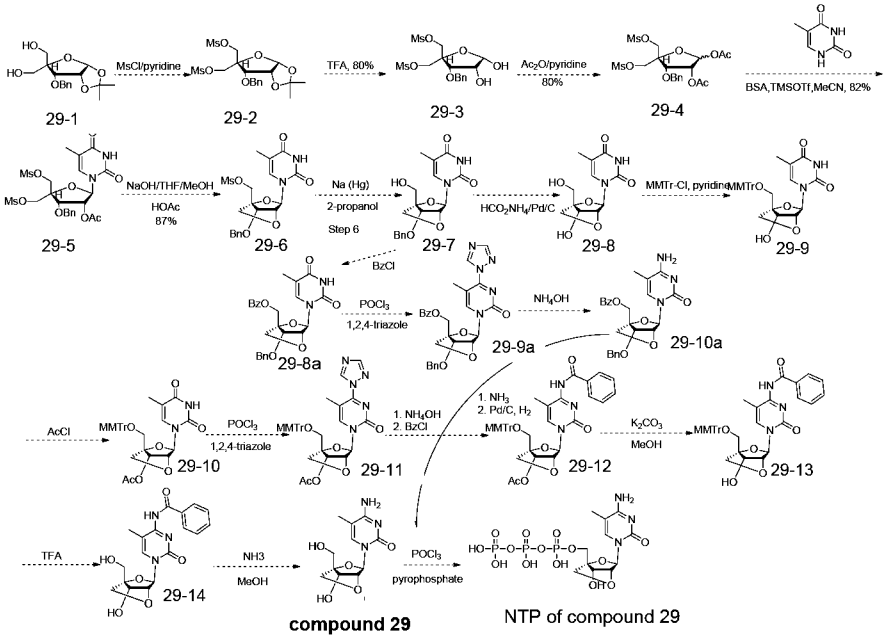
**[00908]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 35. Synthesis of 2'-O,4'-C-methylene adenosine (compound 28) and 2'-O,4'-C-methylene ATP (NTP of said compound)**



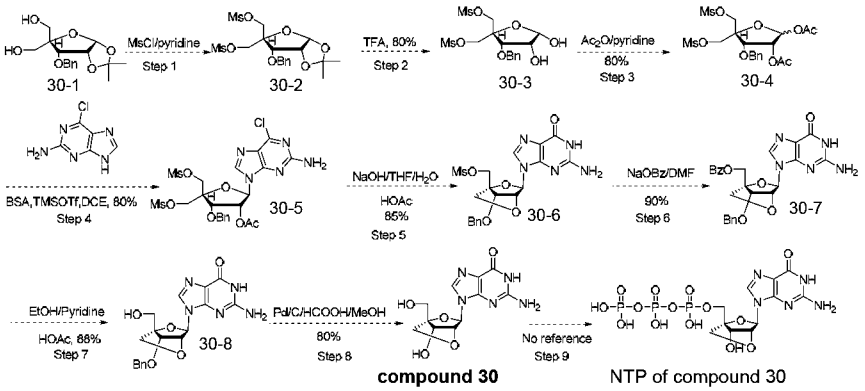

**[00909]** Similar to the strategy provided above for compound 24, compound 28-7 was produced with compound 28-1. Subsequent mesylation, deprotection, and acetylation provided compound 28-10, which was followed by addition of N6-benzoyladenine and subsequent internal cyclization. Various protection and deprotection steps provided compound 28.

**Example 36. Synthesis of 5-methyl-2'-O,4'-C-methylene cytidine (compound 29) and 5-methyl-2'-O,4'-C-methylene CTP (NTP of said compound)**




**[00910]** Aldofuranose compound 29-1 was reacted via various protection steps, and then 5-methyluracil was added to provide compound 29-5. Subsequent internal cyclization, deprotection, protection, and amination steps provided compound 29.

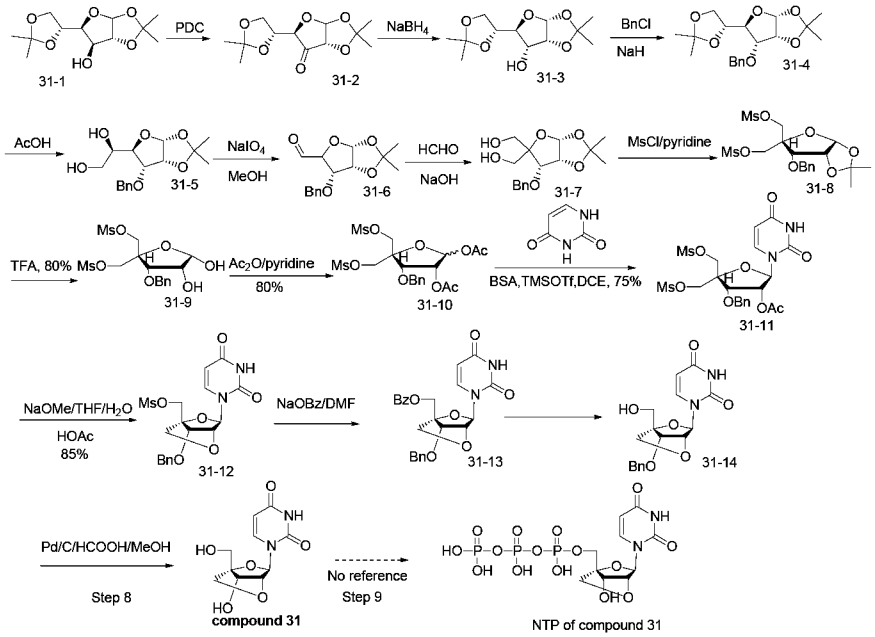
**Example 37. Synthesis of 2'-O,4'-C-methylene guanosine (compound 30) and 2'-O,4'-C-methylene GTP (NTP of said compound)**




**[00911]** Similar to the strategy provided above for compound 29, aldofuranose compound 30-1 was reacted via various protection steps, and then 2-amino-6-chloropurine was added to provide compound 30-5. Subsequent internal cyclization, amination, and deprotection steps provided compound 30.

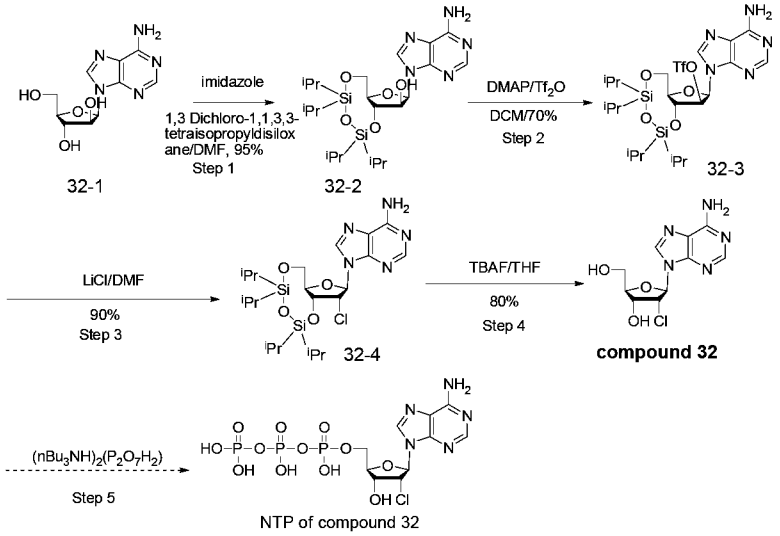
**[00912]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 38. Synthesis of 2'-O,4'-C-methylene uridine (compound 31) and 2'-O,4'-C-methylene UTP (NTP of said compound)**



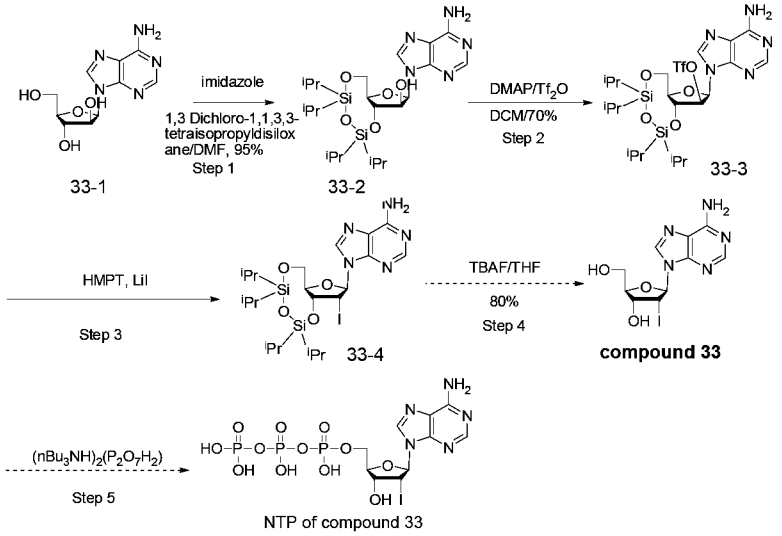

**[00913]** Similar to the strategy provided above for compound 24, compound 31-7 was produced with compound 31-1. Subsequent mesylation, deprotection, and acetylation provided compound 30-10. Addition of uracil and subsequent internal cyclization provided compound 31-12, and various protection and deprotection steps provided compound 31. A subsequent triphosphate reaction (e.g., as described herein) provided the NTP of compound 31, which can be optionally purified (e.g., with HPLC).

**Example 39. Synthesis of 2'-chloro adenosine (compound 32) and 2'-chloro ATP (NTP of said compound)**



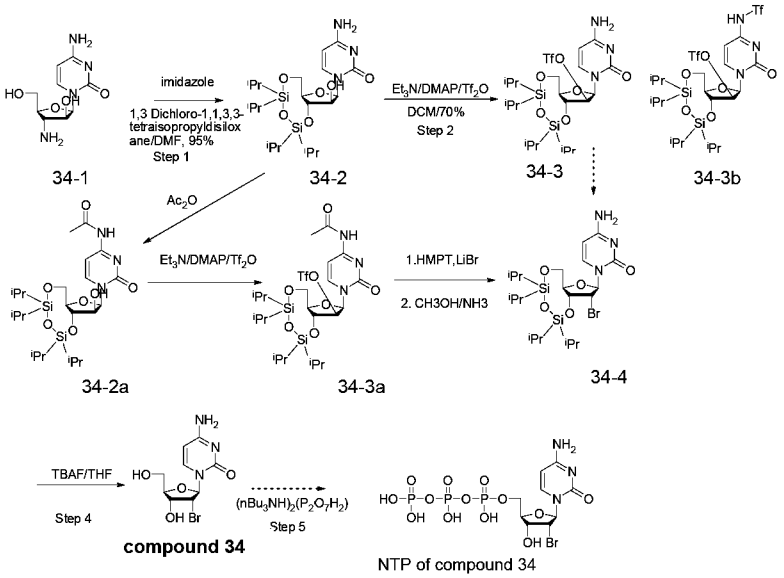

**[00914]** Arabinoadenosine 32-1 was protected via steps 1 and 2 and then chlorinated to provide compound 32-4. Subsequent deprotection provided compound 32, and the triphosphate reaction provided the NTP of compound 32.

**Example 40. Synthesis of 2'-iodo adenosine (compound 33) and 2'-iodo ATP (NTP of said compound)**



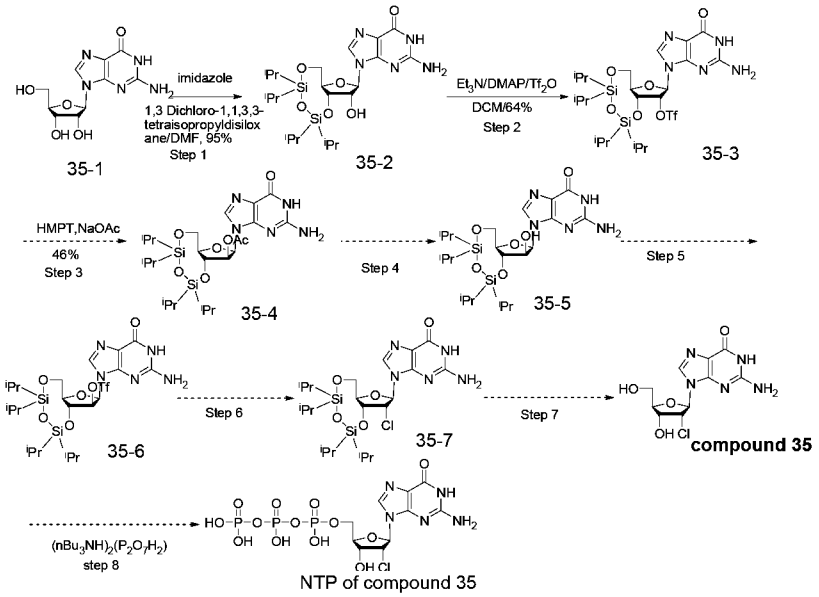

**[00915]** Arabinoadenosine 33-1 was protected via steps 1 and 2 and then iodinated to provide compound 33-4. Subsequent deprotection provided compound 33, and the triphosphate reaction in DMF provided the NTP of compound 33.

**Example 41. Synthesis of 2'-bromo cytidine (compound 34) and 2'-bromo CTP (NTP of said compound)**




**[00916]** Arabinocytidine 34-1 was protected under various conditions and then brominated to provide compound 34-4. Optionally, the reaction can provide compound 34-4 via compound 34-3a under any useful protection reactions, such as (i) 1.5 eq. $Et_3N$, 1 eq. DMAP, 1.2 eq. TfCl, in DCM (10mL); (ii) 3 eq. DMAP, 1.2 eq. TfCl in DCM (15mL); or (iii) 15 eq. DMAP, 1.5 eq. $Tf_2O$, in DCM (15mL) at -10°C to 0°C for 2 hour. In particular, 55mg of compound 34-3a was obtained from reaction condition (iii). Subsequent deprotection provided compound 34, and the triphosphate reaction in DMF provided the NTP of compound 34. Crude product 34 could be optionally purified prior to phosphorylation.

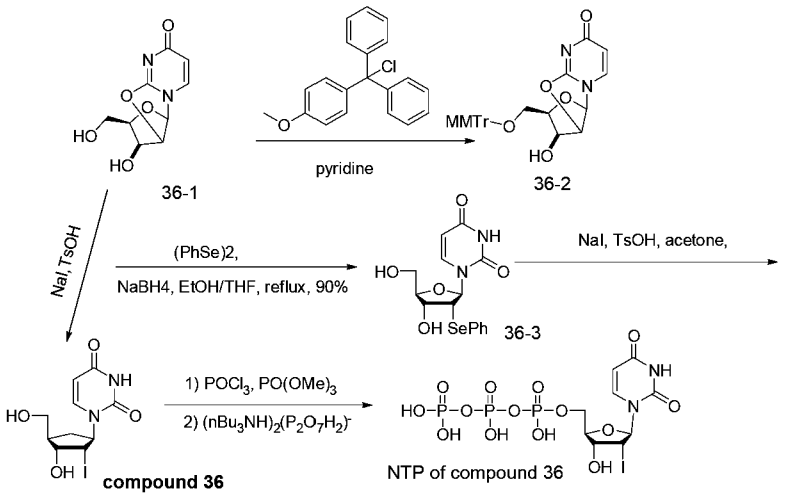
**Example 42. Synthesis of 2'-chloro guanosine (compound 35) and 2'-chloro GTP (NTP of said compound)**




**[00917]** Guanosine 35-1 was protected under various conditions and then acetylated to provide compound 35-4. The reaction from compound 35-2 to compound 35-3 was conducted with 2 eq. DMAP, 2 eq. $Et_3N$, 3 eq. $Tf_2O$ in 1,2-dichloroethane (10mL) at 40°C for 4 hours. About 55mg of compound 35-3 was obtained after the purification.

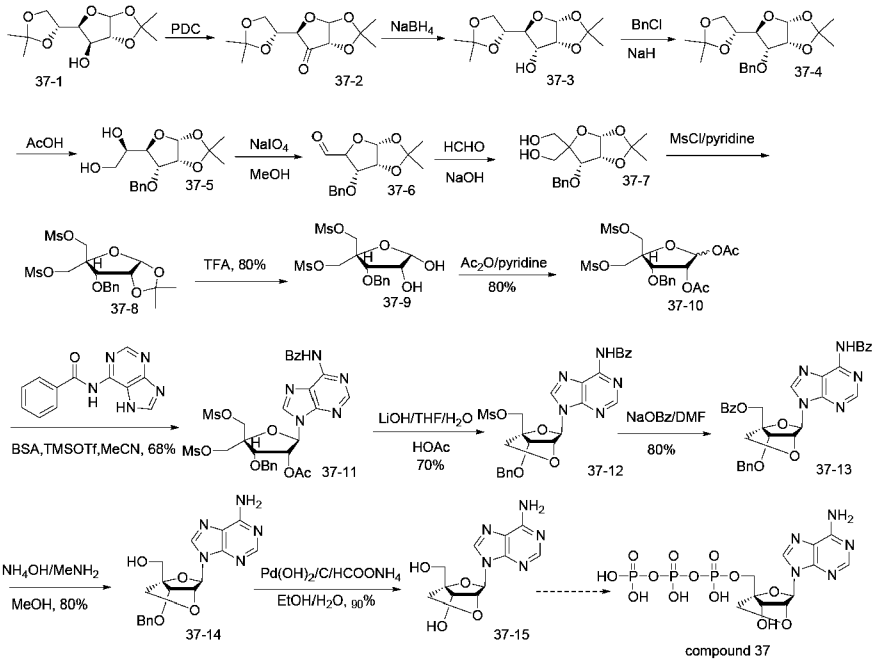
**[00918]** Desired compound 35 can be obtained by any useful method. For example, as shown above, compound 35-4 can be treated with subsequent protection, chlorination, and deprotection steps to provide compound 35. To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 43. Synthesis of 2'-iodo uridine (compound 36) and 2'-iodo UTP (NTP of said compound)**




**[00919]** $O^2$,2'-Cyclouridine 36-1 was protected to provide compound 36-2. Subsequent iodination, optionally mediated with selenium, provided compound 36. A triphosphate reaction was conducted to provide the NTP of compound 36. Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

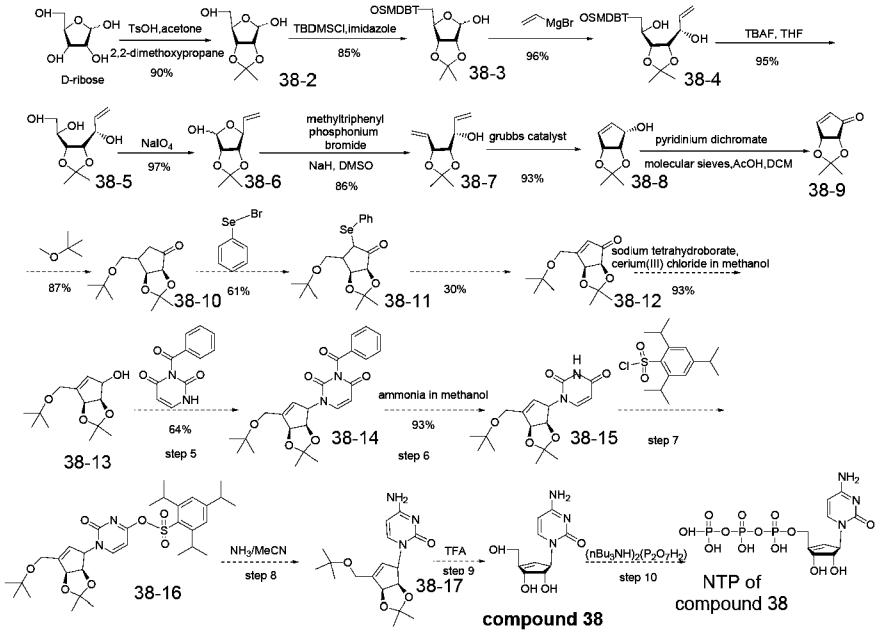
**Example 44. Synthesis of 2'-O,4'-C-methylene adenosine (compound 37) and 2'-O,4'-C-methylene ATP (NTP of said compound)**




**[00920]** Similar to the strategy provided above for compound 24, compound 37-7 was produced with compound 37-1. Subsequent mesylation, deprotection, and acetylation provided compound 37-10. Addition of uracil and subsequent internal cyclization provided compound 37-12. Various protection and deprotection steps provided compound 37.

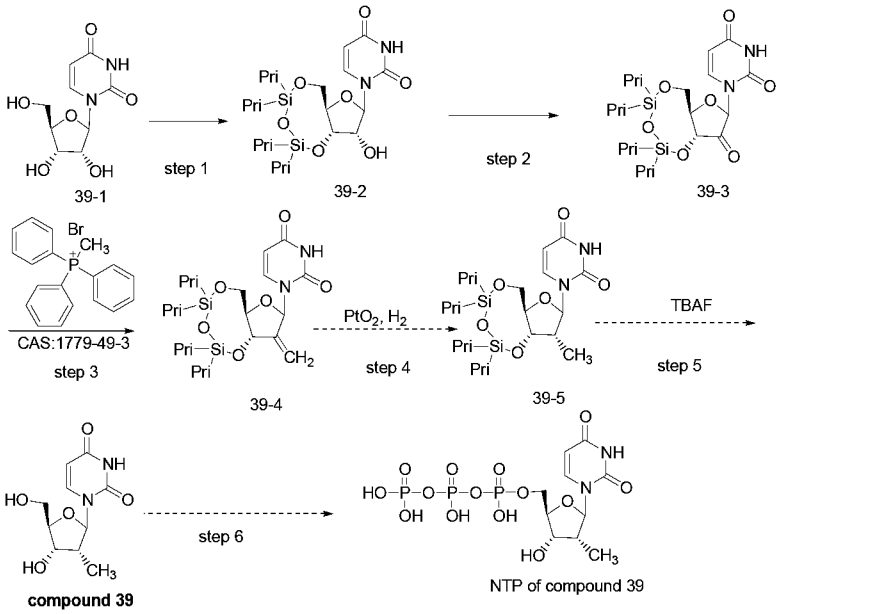
**[00921]** To obtain the corresponding NTP, a triphosphate reaction can be conducted (e.g., any described herein). Optionally, the NTP can be purified (e.g., using a Sephadex DEAE-A25 column), lyophilized, or evaporated (e.g., from EtOH).

**Example 45. Synthesis of cyclopentene diol cytidine (compound 38) and cyclopentene diol CTP (NTP of said compound)**



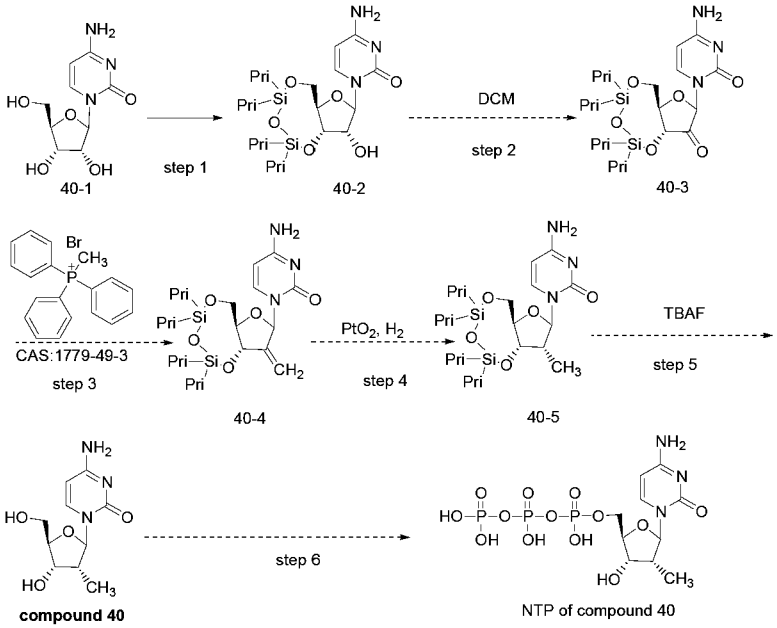

**[00922]** D-ribose was protected and then allylated to provide compound 38-4, which was subsequently cyclized and reduced to provide compound 38-7. Olefin metathesis and subsequent oxidation provided compound 38-9, and further reduction reactions and addition of N-benzoyluracil provided compound 38-14. Additional deprotection and protection reactions provided compound 38, and triphosphate reaction (e.g,. with any useful reaction condition, such as those described herein or in U.S. Pat. No. 7,893,227, incorporated herein by reference) provided the NTP of compound 38.

**Example 46. Synthesis of 2'-methyl uridine (compound 39) and 2'-methyl UTP (NTP of said compound)**



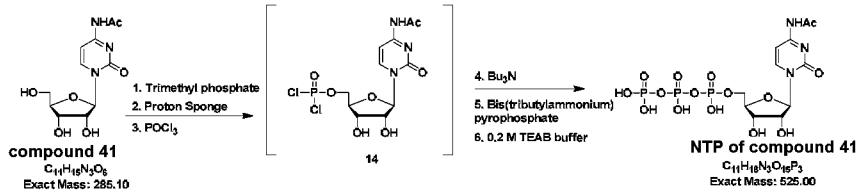

**[00923]** Uridine 39-1 was protected and then oxidized with 2 eq. of Dess-Martin periodane to provide compound 39-3. Subsequent Wittig reaction, hydrogenation, and deprotection steps provided compound 39.

**Example 47. Synthesis of 2'-methyl cytidine (compound 40) and 2'-methyl CTP (NTP of said compound)**




**[00924]** Cytidine 40-1 was protected and then oxidized to provide compound 40-3. Subsequent Wittig reaction, hydrogenation, and deprotection steps provided compound 40.

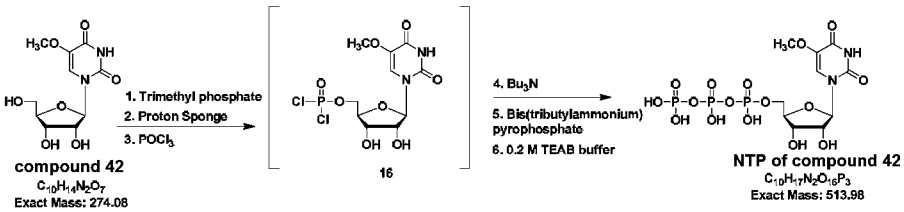
**Example 48. Synthesis of N-acetyl cytidine (compound 41) and N-acetyl CTP (NTP of said compound)**




**[00925]** A solution of N-acetyl-cytidine (compound 41) (103.0 mg, 0.36 mmol) was added to proton sponge (115.72mg, 0.54 mmol, 1.50 equiv) in 1.0 mL trimethylphosphate (TMP) and 1.0 mL of anhydrous tetrahydrofuran (THF). The solution was stirred for 10 minutes at 0°C. Phosphorous

oxychloride ($POCl_3$) (67.2 ul, 0.72 mmol, 2.0 eqiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (1.28 g, 2.34 mmol, 6.5 eqiv.) and tributylamine (350.0 ul, 1.45 mmol, 4.0 equiv.) in 2.5 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 24.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 16.81-17.80 min). Fractions containing the desired compound were pooled and lyophilized to produce the NTP of compound 41. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

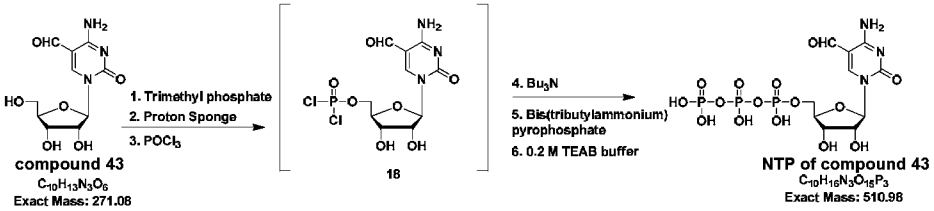
**Example 49. Synthesis of 5-methoxy uridine (compound 42) and 5-methoxy UTP (NTP of said compound)**




**[00926]** A solution of 5-methoxy uridine (compound 42) (69.0 mg, 0.25 mmol, plus heat to make it soluble) was added to proton sponge (80.36 mg, 0.375 mmol, 1.50 equiv.) in 0.7 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (46.7 ul, 0.50 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (894.60 mg, 1.63 mmol, 6.50 equiv.) and tributylamine (243.0 ul, 1.00 mmol, 4.0 equiv.) in 2.0 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 17.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC

(Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 16.57-17.51 min). Fractions containing the desired compound were pooled and lyophilized to produce the NTP of compound 42. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

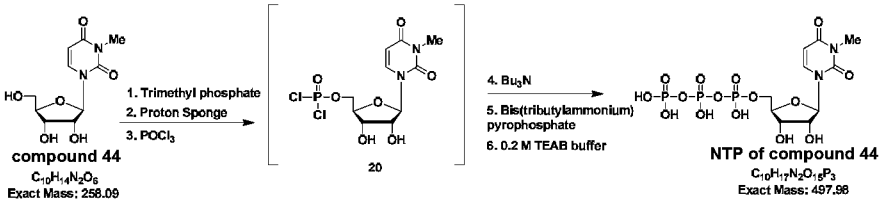
**Example 50. Synthesis of 5-formyl cytidine (compound 43) and 5-formyl CTP (NTP of said compound)**




**[00927]** A solution of 5-formyl cytidine (compound 43) ) (48.4 mg, 0.18 mmol, plus heat to make it soluble) was added to proton sponge (57.86 mg, 0.27 mmol, 1.50 equiv.) in 0.7 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (33.6 ul, 0.36 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (642.0 mg, 1.17 mmol, 6.50 equiv.) and tributylamine (175.0 ul, 0.72 mmol, 4.0 equiv.) in 1.7 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 12.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 17.04-17.87 min). Fractions containing the desired compound were pooled and lyophilized to provide the NTP of compound 43. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein

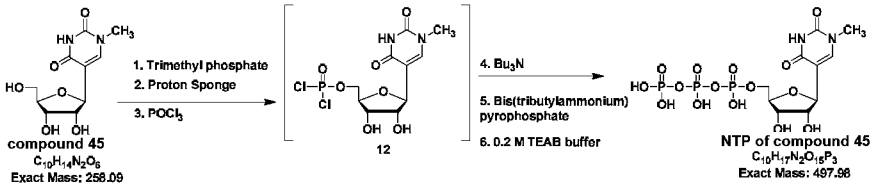
sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

**Example 51. Synthesis of 3-methyl uridine (compound 44) and 3-methyl UTP (NTP of said compound)**



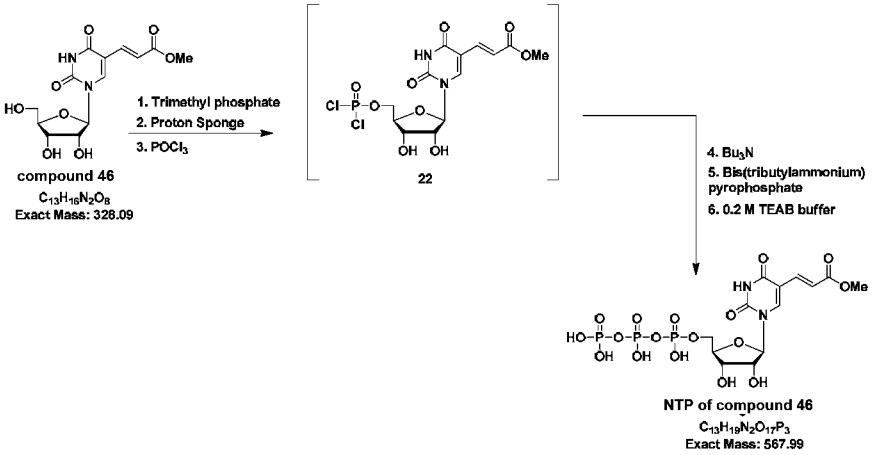

**[00928]** A solution of 3-methyl uridine (compound 44) (45.80 mg, 0.18 mmol) was added to proton sponge (57.86 mg, 0.27 mmol, 1.50 equiv.) in 0.5 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (33.6 ul, 0.36 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (652.0 mg, 1.19 mmol, 6.60 equiv.) and tributylamine (175.0 ul, 0.72 mmol, 4.0 equiv.) in 1.3 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 12.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 18.52-19.57 min). Fractions containing the desired compound were pooled and lyophilized to provide the NTP of compound 44. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

**Example 52. Synthesis of N1-methyl pseudouridine (compound 45) and N1-methyl pseudoUTP (NTP of said compound)**



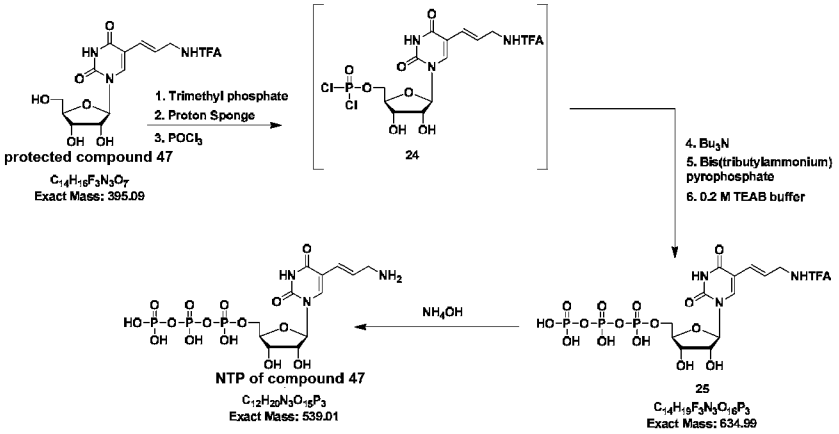

**[00929]** A solution of N1-methyl pseudouridine (compound 45) (96.6 mg, 0.374 mmol, plus heat to make it soluble) was added to proton sponge (120.0 mg, 0.56 mmol, 1.50 equiv.) in 0.8 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (70.0 ul, 0.75 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (1.36g, 2.47 mmol, 6.60 equiv.) and tributylamine (362.0 ul, 1.5 mmol, 4.0 equiv.) in 2.5 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 17.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 15.91-17.01 min). Fractions containing the desired compound were pooled and lyophilized was subjected to a triphosphorylation reaction to provide the NTP of compound 45. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

**Example 53. Synthesis of 5-methoxycarbonylethenyl uridine (compound 46) and 5-methoxycarbonylethenyl UTP (NTP of said compound)**



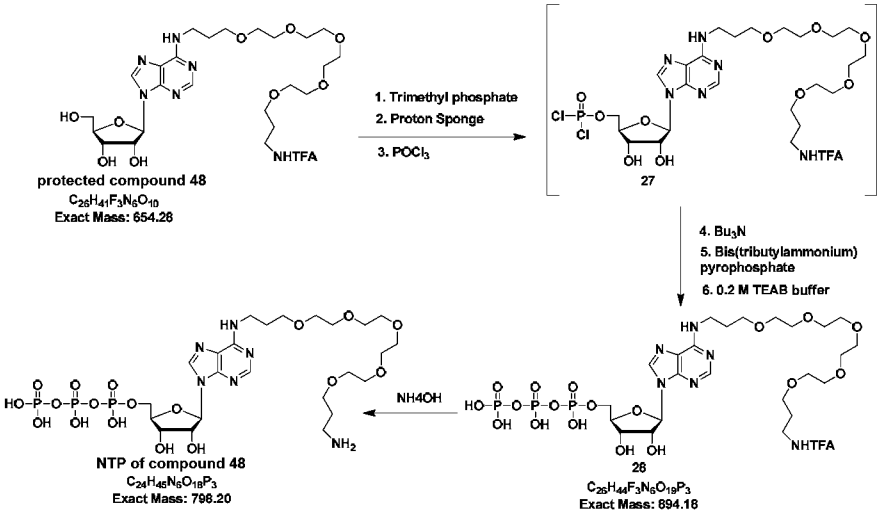

**[00930]** A solution of 5-methoxycarbonylethenyl uridine (compound 46) (102.0 mg, 0.31 mmol) was added to proton sponge (99.65 mg, 0.46 mmol, 1.50 equiv.) in 0.8 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (57.8 ul, 0.62 mmol, 2.0 equiv) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (1.12g, 2.05 mol, 6.60 equiv.) and tributylamine (300.0 ul, 1.24 mmol, 4.0 equiv.) in 2.5 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 20.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 21.56-23.21 min). Fractions containing the desired compound were pooled and lyophilized to provide the NTP of compound 46. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

**Example 54. Synthesis of 5-aminopropenyl uridine (compound 47) and 5-aminopropenyl UTP (NTP of said compound)**




**[00931]** 5-Aminopropenyl uridine 47 was protected and a solution of protected compound 47 (86.0 mg, 0.22 mmol) was added to proton sponge (70.7 mg, 0.33 mmol, 1.50 equiv.) in 0.7 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (41.1 ul, 0.44 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (784.6 mg, 1.43 mmol, 6.50 equiv.) and tributylamine (213.0 ul, 0.88 mmol, 4.0 equiv.) in 1.6 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 15.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. 18.0 ml of concentrated ammonium hydroxide was added to the reaction mixture to remove the trifluoroacetyl group. It was then stored stirring overnight. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 16.14-17.02 min). Fractions containing the desired compound were pooled and lyophilized to provide the NTP of compound 47. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

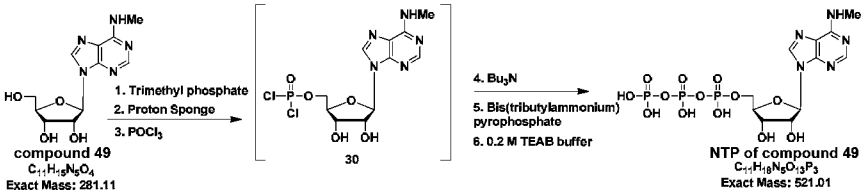
**Example 55. Synthesis of N-PEG adenosine (compound 48) and N-PEG ATP (NTP of said compound)**




**[00932]** N-PEG adenosine 48 was protected and a solution of the protected compound 48 (100.0 mg, 0.15 mmol) was added to proton sponge (49.3 mg, 0.23 mmol, 1.50 equiv.) in 0.65 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (28.0 ul, 0.3 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (537.7 mg, 0.98 mmol, 6.50 equiv.) and tributylamine (146.0 ul, 0.6 mmol, 4.0 equiv.) in 1.2 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 10.0 ml of 0.2M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. 18.0 ml of concentrated ammonium hydroxide was added to the reaction mixture to remove the trifluoroacetyl group. It was then stored stirring overnight. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 24.5-25.5 min). Fractions containing the desired compound were pooled and lyophilized to provide the NTP of compound 48. The triphosphorylation reactions were carried out in a two-neck flask flame-dried

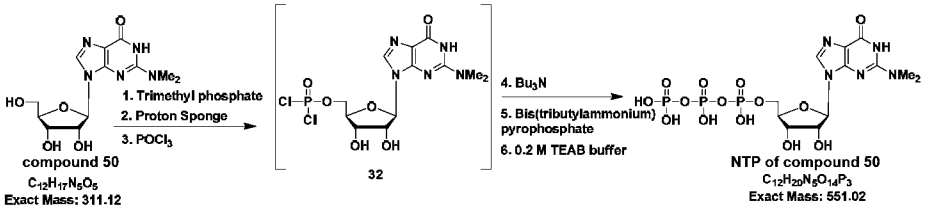
under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

**Example 56. Synthesis of N-methyl adenosine (compound 49) and N-methyl ATP (NTP of said compound)**



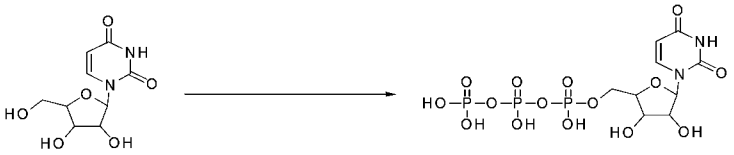

**[00933]** A solution of N-methyl adenosine (compound 49) (70.0 mg, 0.25 mmol) was added to proton sponge (79.29 mg, 0.37 mmol, 1.50 equiv.) in 0.7 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (46.66 ul, 0.50 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (888.85 mg, 1.62 mmol, 6.50 equiv.) and tributylamine (241.0 ul, 1.0 mmol, 4.0 equiv.) in 1.3 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 16.0 ml of 0.2 M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 19.62-20.14 min). Fractions containing the desired compound were pooled and lyophilized to provide the NTP of compound 49. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

**Example 57. Synthesis of N,N-dimethyl guanosine (compound 50) and N,N-dimethyl GTP (NTP of said compound)**




**[00934]** A solution of N,N-dimethyl guanosine (compound 50) (65.8 mg, 0.21 mmol) was added to proton sponge (68.58 mg, 0.32 mmol, 1.50 equiv) in 0.7 mL trimethylphosphate (TMP) and was stirred for 10 minutes at 0°C. Phosphorous oxychloride ($POCl_3$) (39.20 ul, 0.42 mmol, 2.0 equiv.) was added dropwise to the solution before being kept stirring for 2 hours under $N_2$ atmosphere. After 2 hours the solution was reacted with a mixture of bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) (751.67 mg, 1.37 mmol, 6.50 equiv.) and tributylamine (204.0 ul, 0.84 mmol, 4.0 equiv.) in 1.5 ml of dimethylformamide. After approximately 15 minutes, the reaction was quenched with 14.0 ml of 0.2 M triethylammonium bicarbonate (TEAB) and the clear solution was stirred at room temperature for an hour. The reaction mixture was lyophilized overnight and the crude reaction mixture was purified by HPLC (Shimadzu, Kyoto Japan, Phenomenex C18 preparative column, 250 x 21.20 mm, 10.0 micron; gradient: 100 % A for 3.0 min, then 1% B/min, A = 100 mM TEAB buffer, B = ACN; flow rate: 10.0 mL/min; retention time: 19.27-19.95 min). Fractions containing the desired compound were pooled and lyophilized to provide the NTP of compound 50. The triphosphorylation reactions were carried out in a two-neck flask flame-dried under $N_2$ atmosphere. Nucleosides and the protein sponge were dried over $P_2O_5$ under vacuum overnight prior to use. The formation of monophosphates was monitored by LCMS.

**Example 58. General methods for triphosphate synthesis of NTPS**



i ii

**[00935]** The nucleoside i can be phosphorylated by any useful method to provide a triphosphate compound ii. For example, the nucleoside can be added to proton sponge and trimethylphosphate (TMP) and cooled (e.g., to -40°C). Phosphorous oxychloride ($POCl_3$) can be added dropwise before reacting with bistributylammonium pyrophosphate (TBAPP or $(n\text{-}Bu_3NH)_2H_2P_2O_7$) and tributylamine. The reaction can then be quickly quenched with triethylammonium bicarbonate (TEAB). Exemplary conditions are provided in U.S. Pat. No. 7,893,227, which is incorporated herein by reference.

**[00936]** After the phosphorylation reaction, the reaction mixture can be optionally lyophilized, purified (e.g., by ion-exchange chromatography and/or HPLC), or converted to a sodium salt (e.g., by dissolving in MeOH and adding sodium perchlorate in acetone).

**Example 59: PCR for cDNA Production**

**[00937]** PCR procedures for the preparation of cDNA are performed using 2x KAPA HIFI™ HotStart ReadyMix by Kapa Biosystems (Woburn, MA). This system includes 2x KAPA ReadyMix12.5 µl; Forward Primer (10 uM) 0.75 µl; Reverse Primer (10 uM) 0.75 µl; Template cDNA 100 ng; and $dH_20$ diluted to 25.0 µl. The reaction conditions are at 95° C for 5 min. and 25 cycles of 98° C for 20 sec, then 58° C for 15 sec, then 72° C for 45 sec, then 72° C for 5 min. then 4° C to termination.

**[00938]** The reverse primer of the instant invention incorporates a poly-$T_{120}$ for a poly-$A_{120}$ 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.

**[00939]** The reaction is cleaned up using Invitrogen's PURELINK™ PCR Micro Kit (Carlsbad, CA) per manufacturer's instructions (up to 5 µg). Larger reactions will require a cleanup using a product with a larger capacity. Following the cleanup, the cDNA is quantified using the NanoDrop and analyzed by agarose gel electrophoresis to confirm the cDNA is the expected size. The cDNA is then submitted for sequencing analysis before proceeding to the *in vitro* transcription reaction.

**Example 60. In vitro Transcription (IVT)**

**[00940]** 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 un-natural NTPs.

**[00941]** A typical *in vitro* transcription reaction includes the following:

| | |
|---|---|
| Template cDNA | 1.0 μg |
| 10x transcription buffer (400 mM Tris-HCl pH 8.0, 190 mM MgCl2, 50 mM DTT, 10 mM Spermidine) | 2.0 μl |
| Custom NTPs (25mM each | 7.2 μl |
| RNase Inhibitor | 20 U |
| T7 RNA polymerase | 3000 U |
| $dH_20$ | up to 20.0 μl |

**[00942]** Incubation at 37° C for 3 hr-5 hrs.

**[00943]** 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 MEGACLEAR™ 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.

**[00944]** The T7 RNA polymerase may be selected from, T7 RNA polymerase, T3 RNA polymerase and mutant polymerases such as, but not limited to, the novel polymerases able to incorporate modified NTPs as well as those polymerases described by Liu (Esvelt *et al.* (Nature (2011) 472(7344):499-503 and U.S. Publication No. 20110177495) which recognize alternate promoters, Ellington (Chelliserrykattil and Ellington, Nature Biotechnology (2004) 22(9):1155-1160) describing a T7 RNA polymerase variant to transcribe 2'-O-methyl RNA and Sousa (Padilla and Sousa, Nucleic Acids Research (2002) 30(24): e128) describing a T7 RNA polymerase double mutant; herein incorporated by reference in their entireties.

**Example 61. Enzymatic Capping of mRNA**

**[00945]** Capping of the mRNA is performed as follows where the mixture includes: IVT RNA 60 μg-180μg and $dH_20$ up to 72 μl. The mixture is incubated at 65° C for 5 minutes to denature RNA, and then is transferred immediately to ice.

**[00946]** The protocol then involves the mixing of 10x Capping Buffer (0.5 M Tris-HCl (pH 8.0), 60 mM KCl, 12.5 mM $MgCl_2$) (10.0 μl); 20 mM GTP (5.0 μl); 20 mM S-Adenosyl Methionine (2.5

μl); RNase Inhibitor (100 U); 2′-O-Methyltransferase (400U); Vaccinia capping enzyme (Guanylyl transferase) (40 U); $dH_20$ (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.

**[00947]** The mRNA is then purified using Ambion's MEGACLEAR™ 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 62. PolyA Tailing Reaction**

**[00948]** 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); 10x Tailing Buffer (0.5 M Tris-HCl (pH 8.0), 2.5 M NaCl, 100 mM $MgCl_2$)(12.0 μl); 20 mM ATP (6.0 μl); Poly-A Polymerase (20 U); $dH_20$ 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 MEGACLEAR™ kit (Austin, TX) (up to 500 μg). Poly-A Polymerase is preferably a recombinant enzyme expressed in yeast.

**[00949]** For studies performed and described herein, the poly-A tail is encoded in the IVT template to comprise160 nucleotides in length. However, it should be understood that the processivity or integrity of the poly-A tailing reaction may not always result in exactly 160 nucleotides. Hence poly-A tails of approximately 160 nucleotides, e.g, about 150-165, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164 or 165 are within the scope of the invention.

**Example 63. Method of Screening for Protein Expression**

**A. Electrospray Ionization**

**[00950]** A biological sample which may contain proteins encoded by modified RNA administered to the subject is prepared and analyzed according to the manufacturer protocol for electrospray ionization (ESI) using 1, 2, 3 or 4 mass analyzers. A biologic sample may also be analyzed using a tandem ESI mass spectrometry system.

**[00951]** Patterns of protein fragments, or whole proteins, are compared to known controls for a given protein and identity is determined by comparison.

**B. Matrix-Assisted Laser Desorption/Ionization**

**[00952]** A biological sample which may contain proteins encoded by modified RNA administered to the subject is prepared and analyzed according to the manufacturer protocol for matrix-assisted laser desorption/ionization (MALDI).

**[00953]** Patterns of protein fragments, or whole proteins, are compared to known controls for a given protein and identity is determined by comparison.

**C. Liquid Chromatography-Mass spectrometry-Mass spectrometry**

**[00954]** A biological sample, which may contain proteins encoded by modified RNA, may be treated with a trypsin enzyme to digest the proteins contained within. The resulting peptides are analyzed by liquid chromatography-mass spectrometry-mass spectrometry (LC/MS/MS). The peptides are fragmented in the mass spectrometer to yield diagnostic patterns that can be matched to protein sequence databases via computer algorithms. The digested sample may be diluted to achieve 1 ng or less starting material for a given protein. Biological samples containing a simple buffer background (e.g. water or volatile salts) are amenable to direct in-solution digest; more complex backgrounds (e.g. detergent, non-volatile salts, glycerol) require an additional clean-up step to facilitate the sample analysis.

**[00955]** Patterns of protein fragments, or whole proteins, are compared to known controls for a given protein and identity is determined by comparison.

**Example 64. Cytokine Study: PBMC**

**A. PBMC isolation and Culture**

**[00956]** 50 mL of human blood from two donors was received from Research Blood Components (lots KP30928 and KP30931) in sodium heparin tubes. For each donor, the blood was pooled and diluted to 70 mL with DPBS (SAFC Bioscience 59331C, lot 071M8408) and split evenly between two 50 mL conical tubes. 10 mL of Ficoll Paque (GE Healthcare 17-5442-03, lot 10074400) was gently dispensed below the blood layer. The tubes were centrifuged at 2000 rpm for 30 minutes with low acceleration and braking. The tubes were removed and the buffy coat PBMC layers were gently transferred to a fresh 50 mL conical and washed with DPBS. The tubes were centrifuged at 1450 rpm for 10 minutes.

**[00957]** The supernatant was aspirated and the PBMC pellets were resuspended and washed in 50 mL of DPBS. The tubes were centrifuged at 1250 rpm for 10 minutes. This wash step was repeated, and the PBMC pellets were resuspended in 19 mL of Optimem I (Gibco 11058, lot 1072088) and counted. The cell suspensions were adjusted to a concentration of 3.0 x 10^6 cells / mL live cells.

**[00958]** These cells were then plated on five 96 well tissue culture treated round bottom plates (Costar 3799) per donor at 50 uL per well. Within 30 minutes, transfection mixtures were added to each well at a volume of 50 uL per well. After 4 hours post transfection, the media was supplemented with 10 uL of Fetal Bovine Serum (Gibco 10082, lot 1012368)

**B. Transfection Preparation**

**[00959]** Modified mRNA encoding human G-CSF (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) (containing either (1) natural NTPs, (2) 100% substitution with 5-methyl cytidine and pseudouridine, or (3) 100% substitution with 5-methyl cytidine and N1-methyl pseudouridine; mRNA encoding luciferase (IVT cDNA sequence shown in SEQ ID NO: 2; mRNA sequence shown in SEQ ID NO: 3, polyA tail of approximately 160 nucleotides not shown in sequence, 5'cap, Cap1, fully modified with 5-methylcytosine at each cytosine and pseudouridine replacement at each uridine site) (containing either (1) natural NTPs or (2) 100% substitution with 5-methyl cytidine and pseudouridine) and TLR agonist R848 (Invivogen tlrl-r848) were diluted to 38.4 ng / uL in a final volume of 2500 uL Optimem I.

**[00960]** Separately, 110 uL of Lipofectamine 2000 (Invitrogen 11668-027, lot 1070962) was diluted with 6.76 mL Optimem I. In a 96 well plate nine aliquots of 135 uL of each mRNA, positive control (R-848) or negative control (Optimem I) was added to 135 uL of the diluted Lipofectamine 2000. The plate containing the material to be transfected was incubated for 20 minutes. The transfection mixtures were then transferred to each of the human PBMC plates at 50 uL per well. The plates were then incubated at 37ºC. At 2, 4, 8, 20, and 44 hours each plate was removed from the incubator, and the supernatants were frozen.

**[00961]** After the last plate was removed, the supernatants were assayed using a human G-CSF ELISA kit (Invitrogen KHC2032) and human IFN-alpha ELISA kit (Thermo Scientific 41105-2). Each condition was done in duplicate.

**C. Protein and Innate Immune Response Analysis**

**[00962]** The ability of unmodified and modified mRNA to produce the encoded protein was assessed (G-CSF production) over time as was the ability of the mRNA to trigger innate immune recognition as measured by interferon-alpha production. Use of *in vitro* PBMC cultures is an accepted way to measure the immunostimulatory potential of oligonucleotides (Robbins et al., Oligonucleotides 2009 19:89-102).

**[00963]** Results were interpolated against the standard curve of each ELISA plate using a four parameter logistic curve fit. Shown in Tables 4 and 5 are the average from 3 separate PBMC donors of the G-CSF, interferon-alpha (IFN-alpha) and tumor necrosis factor alpha (TNF-alpha) production over time as measured by specific ELISA.

**[00964]** In the G-CSF ELISA, background signal from the Lipofectamine 2000 (LF2000) untreated condition was subtracted at each time point. The data demonstrated specific production of human G-CSF protein by human peripheral blood mononuclear is seen with G-CSF mRNA containing natural NTPs, 100% substitution with 5-methyl cytidine and pseudouridine, or 100% substitution with 5-methyl cytidine and N1-methyl pseudouridine. Production of G-CSF was significantly increased through the use of 5-methyl cytidine and N1-methyl pseudouridine modified mRNA relative to 5-methyl cytidine and pseudouridine modified mRNA.

**[00965]** With regards to innate immune recognition, while both modified mRNA chemistries largely prevented IFN-alpha and TNF-alpha production relative to positive controls (R848, p(I)p(C)), significant differences did exist between the chemistries. 5-methyl cytidine and pseudouridine modified mRNA resulted in low but detectable levels of IFN-alpha and TNF-alpha production, while 5-methyl cytidine and N1-methyl pseudouridine modified mRNA resulted in no detectable IFN-alpha and TNF-alpha production.

**[00966]** Consequently, it has been determined that, in addition to the need to review more than one cytokine marker of the activation of the innate immune response, it has surprisingly been found that combinations of modifications provide differing levels of cellular response (protein production and immune activation). The modification, N1-methyl pseudouridine, in this study has been shown to convey added protection over the standard combination of 5-methylcytidine/pseudouridine explored by others resulting in twice as much protein and almost 150 fold reduction in immune activation (TNF-alpha).

**[00967]** Given that PBMC contain a large array of innate immune RNA recognition sensors and are also capable of protein translation, it offers a useful system to test the interdependency of these two pathways. It is known that mRNA translation can be negatively affected by activation of such innate immune pathways (Kariko et al. Immunity (2005) 23:165-175; Warren et al. Cell Stem Cell (2010) 7:618-630). Using PBMC as an in vitro assay system it is possible to establish a correlation between translation (in this case G-CSF protein production) and cytokine production (in this case exemplified by IFN-alpha and TNF-alpha protein production). Better protein production is correlated with lower

induction of innate immune activation pathway, and new chemistries can be judged favorably based on this ratio (Table 6).

**[00968]** In this study, the PC Ratio for the two chemical modifications, pseudouridine and N1-methyl pseudouridine, both with 5-methy cytosine was 4742/141=34 as compared to 9944/1=9944 for the cytokine IFN-alpha. For the cytokine, TNF-alpha, the two chemistries had PC Ratios of 153 and 1243, respectively suggesting that for either cytokine, the N1-methylpseudouridine is the superior modification. In Tables 4 and 5, "NT" means not tested.

**Table 4. G-CSF**

| G-CSF: 3 Donor Average (pg/ml) | |
|---|---|
| G-CSF 5-methyl cytosine/ pseudouridine | 4742 |
| G-CSF 5-methylcytosine/ N1-methylpseudouridine | 9944 |
| Luciferase | 18 |
| LF2000 | 16 |

**Table 5. IFN-alpha and TNF-alpha**

| | IFN-alpha: 3 Donor Average (pg/ml) | TNF-alpha: 3 Donor Average (pg/ml) |
|---|---|---|
| G-CSF 5-methyl cytosine/ pseudouridine | 141 | 31 |
| G-CSF 5-methylcytosine/ N1-methylpseudouridine | 1 | 8 |
| P(I)P(C) | 1104 | NT |
| R-848 | NT | 1477 |
| LF2000 | 17 | 25 |

**Table 6. G-CSF to Cytokine Ratios**

| | G-CSF/ IFN-alpha (ratio) | | G-CSF/TNF-alpha (ratio) | |
|---|---|---|---|---|
| | 5-methyl cytosine/ pseudouridine | 5-methylcytosine/ N1-methylpseudouridine | 5-methyl cytosine/ pseudouridine | 5-methylcytosine/ N1-methylpseudouridine |
| PC Ratio | 34 | 9944 | 153 | 1243 |

**Example 65. Chemical Modification Ranges of Modified mRNA**

**[00969]** Modified nucleosides such as, but not limited to, the chemical modifications 5-methylcytosine and pseudouridine have been shown to lower the innate immune response and

increase expression of RNA in mammalian cells. Surprisingly and not previously known, the effects manifested by these chemical modifications can be titrated when the amount of chemical modification of a particular nucleotide is less than 100%. Previously, it was believed that the benefit of chemical modification could be derived using less than complete replacement of a modified nucleoside and published reports suggest no loss of benefit until the level of substitution with a modified nucleoside is less than 50% (Kariko et al., Immunity (2005) 23:165-175).

**[00970]** However, it has now been shown that the benefits of chemical modification are directly correlated with the degree of chemical modification and must be considered in view of more than a single measure of immune response. Such benefits include enhanced protein production or mRNA translation and reduced or avoidance of stimulating the innate immune response as measured by cytokine profiles and metrics of immune response triggers.

**[00971]** Enhanced mRNA translation and reduced or lack of innate immune stimulation are seen with 100% substitution with a modified nucleoside. Lesser percentages of substitution result in less mRNA translation and more innate immune stimulation, with unmodified mRNA showing the lowest translation and the highest innate immune stimulation.

***In Vitro* PBMC Studies: Percent modification**

**[00972]** 480 ng of G-CSF mRNA modified with 5-methylcytosine (5mC) and pseudouridine (pseudoU) or unmodified G-CSF mRNA was transfected with 0.4 uL of Lipofectamine 2000 into peripheral blood mononuclear cells (PBMC) from three normal blood donors (D1, D2, and D3). The G-CSF mRNA (SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) was completely modified with 5mC and pseudo (100% modification), not modified with 5mC and pseudo (0% modification) or was partially modified with 5mC and pseudoU so the mRNA would contain 75% modification, 50% modification or 25% modification. A control sample of Luciferase (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1; fully modified 5meC and pseudoU) was also analyzed for G-CSF expression. For TNF-alpha and IFN-alpha control samples of Lipofectamine2000, LPS, R-848, Luciferase (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1; fully modified 5mC and pseudo), and P(I)P(C) were also analyzed. The supernatant was harvested and run by ELISA 22 hours after transfection to determine the protein expression. The expression of G-CSF is shown in Table 7 and the expression of IFN-alpha and TNF-alpha is shown in Table 8. The expression of

IFN-alpha and TNF-alpha may be a secondary effect from the transfection of the G-CSF mRNA. Tables 7, 8 and FIG. 10 show that the amount of chemical modification of G-CSF, interferon alpha (IFN-alpha) and tumor necrosis factor-alpha (TNF-alpha) is titratable when the mRNA is not fully modified and the titratable trend is not the same for each target.

**[00973]** As mentioned above, using PBMC as an in vitro assay system it is possible to establish a correlation between translation (in this case G-CSF protein production) and cytokine production (in this case exemplified by IFN-alpha protein production). Better protein production is correlated with lower induction of innate immune activation pathway, and the percentage modification of a chemistry can be judged favorably based on this ratio (Table 9). As calculated from Tables 7 and 8 and shown in Table 9, full modification with 5-methylcytidine and pseudouridine shows a much better ratio of protein/cytokine production than without any modification (natural G-CSF mRNA) (100-fold for IFN-alpha and 27-fold for TNF-alpha). Partial modification shows a linear relationship with increasingly less modification resulting in a lower protein/cytokine ratio.

**Table 7. G-CSF Expression**

| | G-CSF Expression (pg/ml) | | |
|---|---|---|---|
| | **D1** | **D2** | **D3** |
| 100% modification | 1968.9 | 2595.6 | 2835.7 |
| 75% modification | 566.7 | 631.4 | 659.5 |
| 50% modification | 188.9 | 187.2 | 191.9 |
| 25% modification | 139.3 | 126.9 | 102.0 |
| 0% modification | 194.8 | 182.0 | 183.3 |
| Luciferase | 90.2 | 0.0 | 22.1 |

**Table 8. IFN-alpha and TNF-alpha Expression**

| | IFN-alpha Expression (pg/ml) | | | TNF-alpha Expression (pg/ml) | | |
|---|---|---|---|---|---|---|
| | **D1** | **D2** | **D3** | **D1** | **D2** | **D3** |
| 100% modification | 336.5 | 78.0 | 46.4 | 115.0 | 15.0 | 11.1 |
| 75% modification | 339.6 | 107.6 | 160.9 | 107.4 | 21.7 | 11.8 |
| 50% modification | 478.9 | 261.1 | 389.7 | 49.6 | 24.1 | 10.4 |
| 25% modification | 564.3 | 400.4 | 670.7 | 85.6 | 26.6 | 19.8 |
| 0% modification | 1421.6 | 810.5 | 1260.5 | 154.6 | 96.8 | 45.9 |
| LPS | 0.0 | 0.6 | 0.0 | 0.0 | 12.6 | 4.3 |
| R-848 | 0.5 | 3.0 | 14.1 | 655.2 | 989.9 | 420.4 |
| P(I)P(C) | 130.8 | 297.1 | 585.2 | 765.8 | 2362.7 | 1874.4 |
| Lipid only | 1952.2 | 866.6 | 855.8 | 248.5 | 82.0 | 60.7 |

**Table 9. PC Ratio and Effect of Percentage of Modification**

| % Modification | Average G-CSF (pg/ml) | Average IFN-a (pg/ml) | Average TNF-a (pg/ml) | G-CSF/ IFN-alpha (PC ratio) | G-CSF/TNF-alpha (PC ratio) |
|---|---|---|---|---|---|
| 100 | 2466 | 153 | 47 | 16 | 52 |
| 75 | 619 | 202 | 47 | 3.1 | 13 |
| 50 | 189 | 376 | 28 | 0.5 | 6.8 |
| 25 | 122 | 545 | 44 | 0.2 | 2.8 |
| 0 | 186 | 1164 | 99 | 0.16 | 1.9 |

**Example 66. Modified RNA transfected in PBMC**

**[00974]** 500 ng of G-CSF mRNA modified with 5-methylcytosine (5mC) and pseudouridine (pseudoU) or unmodified G-CSF mRNA was transfected with 0.4 uL of Lipofectamine 2000 into peripheral blood mononuclear cells (PBMC) from three normal blood donors (D1, D2, and D3). The G-CSF mRNA (SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) was completely modified with 5mC and pseudo (100% modification), not modified with 5mC and pseudo (0% modification) or was partially modified with 5mC and pseudoU so the mRNA would contain 50% modification, 25% modification, 10% modification, %5 modification, 1% modification or 0.1% modification. A control sample of mCherry (mRNA sequence shown in SEQ ID NO: 6; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1; fully modified 5meC and pseudouridine) and G-CSF fully modified with 5-methylcytosine and pseudouridine (Control G-CSF) was also analyzed for G-CSF expression. For tumor necrosis factor-alpha (TNF-alpha) and interferon-alpha (IFN-alpha) control samples of Lipofectamine2000, LPS, R-848, Luciferase (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1; fully modified 5mC and pseudo), and P(I)P(C) were also analyzed. The supernatant was harvested 6 hours and 18 hours after transfection and run by ELISA to determine the protein expression. The expression of G-CSF, IFN-alpha, and TNF-alpha for Donor 1 is shown in Table 10, Donor 2 is shown in Table 11 and Donor 3 is shown in Table 12.

**[00975]** Full 100% modification with 5-methylcytidine and pseudouridine resulted in the most protein translation (G-CSF) and the least amount of cytokine produced across all three human PBMC donors. Decreasing amounts of modification results in more cytokine production (IFN-alpha and TNF-alpha), thus further highlighting the importance of fully modification to reduce cytokines and to improve protein translation (as evidenced here by G-CSF production).

**Table 10. Donor 1**

| | G-CSF (pg/mL) | IFN-alpha (pg/mL) | TNF-alpha (pg/mL) |
|---|---|---|---|

| | 6 hours | 18 hours | 6 hours | 18 hours | 6 hours | 18 hours |
|---|---|---|---|---|---|---|
| 100% Mod | 1815 | 2224 | 1 | 13 | 0 | 0 |
| 75% Mod | 591 | 614 | 0 | 89 | 0 | 0 |
| 50% Mod | 172 | 147 | 0 | 193 | 0 | 0 |
| 25% Mod | 111 | 92 | 2 | 219 | 0 | 0 |
| 10% Mod | 138 | 138 | 7 | 536 | 18 | 0 |
| 1% Mod | 199 | 214 | 9 | 660 | 18 | 3 |
| 0.1% Mod | 222 | 208 | 10 | 597 | 0 | 6 |
| 0 % Mod | 273 | 299 | 10 | 501 | 10 | 0 |
| Control G-CSF | 957 | 1274 | 3 | 123 | 18633 | 1620 |
| mCherry | 0 | 0 | 0 | 10 | 0 | 0 |
| Untreated | N/A | N/A | 0 | 0 | 1 | 1 |

**Table 11. Donor 2**

| | **G-CSF (pg/mL)** | | **IFN-alpha (pg/mL)** | | **TNF-alpha (pg/mL)** | |
|---|---|---|---|---|---|---|
| | 6 hours | 18 hours | 6 hours | 18 hours | 6 hours | 18 hours |
| 100% Mod | 2184 | 2432 | 0 | 7 | 0 | 11 |
| 75% Mod | 935 | 958 | 3 | 130 | 0 | 0 |
| 50% Mod | 192 | 253 | 2 | 625 | 7 | 23 |
| 25% Mod | 153 | 158 | 7 | 464 | 6 | 6 |
| 10% Mod | 203 | 223 | 25 | 700 | 22 | 39 |
| 1% Mod | 288 | 275 | 27 | 962 | 51 | 66 |
| 0.1% Mod | 318 | 288 | 33 | 635 | 28 | 5 |
| 0 % Mod | 389 | 413 | 26 | 748 | 1 | 253 |
| Control G-CSF | 1461 | 1634 | 1 | 59 | 481 | 814 |
| mCherry | 0 | 7 | 0 | 1 | 0 | 0 |
| Untreated | N/A | N/A | 1 | 0 | 0 | 0 |

**Table 12. Donor 3**

| | **G-CSF (pg/mL)** | | **IFN-alpha (pg/mL)** | | **TNF-alpha (pg/mL)** | |
|---|---|---|---|---|---|---|
| | 6 hours | 18 hours | 6 hours | 18 hours | 6 hours | 18 hours |
| 100% Mod | 6086 | 7549 | 7 | 658 | 11 | 11 |
| 75% Mod | 2479 | 2378 | 23 | 752 | 4 | 35 |
| 50% Mod | 667 | 774 | 24 | 896 | 22 | 18 |
| 25% Mod | 480 | 541 | 57 | 1557 | 43 | 115 |
| 10% Mod | 838 | 956 | 159 | 2755 | 144 | 123 |
| 1% Mod | 1108 | 1197 | 235 | 3415 | 88 | 270 |
| 0.1% Mod | 1338 | 1177 | 191 | 2873 | 37 | 363 |
| 0 % Mod | 1463 | 1666 | 215 | 3793 | 74 | 429 |
| Control G-CSF | 3272 | 3603 | 16 | 1557 | 731 | 9066 |
| mCherry | 0 | 0 | 2 | 645 | 0 | 0 |
| Untreated | N/A | N/A | 1 | 1 | 0 | 8 |

**Example 67. Microames Reverse Mutation Screen of Modifications**

**Background and Methods**

**[00976]** The microames screen is a version of the full Ames preincubation assay. It detects both frameshift and base-pair substitution mutations using four *Salmonella* tester strains (TA97a, TA98, TA100 and TA1535) and one *Escherichia coli* strain (WP2 *uvrA* pKM101). Strains TA97a and TA98 detect frameshift mutations, and TA100, TA1535 and WP2 *uvrA* pKM101 detect base-pair substitution mutations. This scaled-down Ames test uses minimal compound, is conducted with and without metabolic activation (S9 fraction), and uses multiwell plates. This teste is a microbial assay to detect the mutagenic potential of test compounds.

**[00977]** The microAmes screen for 5-Methylcytidine, Pseudouridine or N'-methylpseudouridine test article was tested in duplicate with strains TA97a, TA98, TA100, TA1535 and WP2 *uvrA* pKM101 in the presence and absence of a metabolic activation system (AROCLOR™ 1254 induced rat liver S9 microsomal fraction) at 0.25, 2.5, 12.5, 25, 75, and 250 ug/well. Positive control compounds were used at 4 different concentrations to ensure the assay system was sensitive to known mutagenic compounds. DMSO was used as the vehicle control. Positive and vehicle controls yielded the expected results, demonstrating that the microAmes screen is sufficiently sensitive to detect mutagens.

**Results**

**[00978]** For 5-methylcytosine, precipitates were not observed with any tester strain either with or without metabolic activation. Cytotoxicity (reduction in the background lawn and/or number of revertants) was not observed in any strain either with or without metabolic activation. There was no increase in the number of revertant colonies as compared with the vehicle control in any strain with or without metabolic activation. Therefore, 5-Methylcytidine was not mutagenic up to 250 ug/well in strains TA97a, TA98, TA100, TA1535 and WP2 uvrA pKM101 with or without metabolic activation under the conditions of the microAmes screen.

**[00979]** Precipitates were not observed with any tester strain either with or without metabolic activation for pseudouridine. Cytotoxicity (reduction in the number of revertants) was observed with strain TA100 without metabolic activation. Cytotoxicity (reduction in the background lawn and/or number of revertants) was not observed in any other strain either with or without metabolic activation. There was no increase in the number of revertant colonies as compared with the vehicle control in any strain with or without metabolic activation. Therefore, pseudouridine was not mutagenic up to 75 ug/well in strain TA100 without metabolic activation and up to 250 µg/well in

strains TA97a, TA98, TA1535 and WP2 uvrA pKM101 with or without metabolic activation and strain TA100 without metabolic activation under the conditions of this microAmes screen.

**[00980]** For the modification, N1-methylpseudouridine precipitates were not observed with any tester strain either with or without metabolic activation. Cytotoxicity (reduction in the background lawn and/or number of revertants) was not observed in any strain either with or without metabolic activation. There was no increase in the number of revertant colonies as compared with the vehicle control in any strain with or without metabolic activation. N1-methylpseudouridine was not mutagenic up to 250 μg/well in strains TA97a, TA98, TA100, TA1535 and WP2 uvrA pKM101 with or without metabolic activation under the conditions of this microAmes screen. N1-methylpseudouridine was found less mutagenic than pseudouridine.

**[00981]** The comparison in this microAMES test of 5 methyl cytidine, pseudouridine, and N1-methylpseudouridine reveal them to be generally non-mutagenic. Of particular note, however, was the difference between pseudouridine and N1-methylpseudouridine, where pseudouridine did show a cytotoxic response in one bacterial strain where N1-methylpseudouridine did not. These microAMES tests are routinely used as part of the pre-clinical assessment of compound safety and highlight an important difference between N1-methylpseudouridine and pseudouridine.

**Example 68. Toxicity of Nucleoside Triphosphates (NTPs)**

**[00982]** The cytotoxicity of natural and modified nucleoside triphosphates (NTPs) alone or in combination with other bases, was analyzed in human embryonic kidney 293 (HEK293) cells in the absence of transfection reagent. HEK293 cells were seeded on 96-well plates at a density of 30,000 cells per well having 0.75ul of RNAiMAX™ (Invitrogen, Carlsbad, CA) per well at a total well volume of 100ul. 10 ul of the NTPs outlined in Table 12 were combined with 10 ul of lipid dilution and incubated for 30 minutes to form a complex before 80 ul of the HEK293 cell suspension was added to the NTP complex.

**[00983]** Natural and modified NTPs were transfected at a concentration of 2.1 nM, 21 nM, 210 nM, 2.1 um, 21 uM, 210 um or 2.1 mM. NTPs in combination were transfected at a total concentration of NTPs of 8.4 nM, 84 nM, 840 nM, 8.4 uM, 84 uM, 840 uM and 8.4 mM. As a control modified G-CSF mRNA (SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1; fully modified 5-methylcytosine and pseudouridine) was transfected in HEK293 cells at a concentration of 8.4 nM. The cytotoxicity of the NTPs and the modified G-CSF mRNA was assayed at 4, 24, 48 and 72 hours post addition to the HEK293 cells

using a CYTO TOX-GLO™ assay from Promega (Madison, WI) following the manufacturer protocol except pippeting was used for lysing the cells instead of shaking the plates.

**[00984]** Table 13 and 14 show the percent of viable cells for each of the NTPs, NTP combinations and controls tested. There was no toxicity seen with the individual NTPs as compared to the untreated cells. These data demonstrate that introduction of individual NTPs, including 5-methylcytidine, pseudouridine, and N1-methylpseudouridine, into mammalian cells is not toxic at doses 1,000,000 times an effective dose when introduced as a modified mRNA.

**Table 13. Cytotoxicity of Individual NTPs**

| Individual NTP Cytotoxicity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Time | Dose | | | | | | |
| | | 2.1 mM | 210 uM | 21 uM | 2.1 uM | 210 nM | 21 nM | 2.1 nM |
| **Adenine** | 4 hr | 90.03 | 85.97 | 91.20 | 90.23 | 90.36 | 93.21 | 93.48 |
| | 24 hr | 88.42 | 87.31 | 86.86 | 86.81 | 86.94 | 87.19 | 86.44 |
| | 48 hr | 93.71 | 90.55 | 89.94 | 89.80 | 89.17 | 91.13 | 92.12 |
| | 72 hr | 97.49 | 94.81 | 93.83 | 94.58 | 92.22 | 93.88 | 95.74 |
| **Cytosine** | 4 hr | 90.51 | 89.88 | 91.41 | 90.49 | 88.95 | 93.11 | 93.34 |
| | 24 hr | 86.92 | 86.33 | 85.72 | 86.70 | 86.12 | 86.16 | 85.78 |
| | 48 hr | 94.23 | 87.81 | 87.28 | 87.73 | 85.36 | 88.95 | 88.99 |
| | 72 hr | 97.15 | 92.34 | 92.22 | 88.93 | 88.22 | 91.80 | 94.22 |
| **Guanine** | 4 hr | 90.96 | 90.14 | 91.36 | 90.60 | 90.00 | 92.84 | 93.33 |
| | 24 hr | 86.37 | 85.86 | 85.93 | 86.13 | 86.35 | 85.50 | 85.41 |
| | 48 hr | 93.83 | 87.05 | 88.18 | 87.89 | 85.31 | 87.92 | 89.57 |
| | 72 hr | 97.04 | 91.41 | 92.39 | 92.30 | 92.19 | 92.55 | 93.72 |
| **Uracil** | 4 hr | 90.97 | 89.60 | 91.95 | 90.90 | 91.05 | 92.90 | 93.15 |
| | 24 hr | 87.68 | 86.48 | 85.89 | 86.75 | 86.52 | 87.23 | 87.63 |
| | 48 hr | 94.39 | 88.98 | 89.11 | 89.44 | 88.33 | 88.89 | 91.28 |
| | 72 hr | 96.82 | 93.45 | 93.63 | 94.60 | 94.50 | 94.53 | 95.51 |
| **Pseudouridine** | 4 hr | 92.09 | 92.37 | 91.35 | 92.02 | 92.84 | 91.96 | 92.26 |
| | 24 hr | 88.38 | 86.68 | 86.05 | 86.75 | 85.91 | 87.59 | 87.31 |
| | 48 hr | 88.62 | 87.79 | 87.73 | 87.66 | 87.82 | 89.03 | 91.99 |
| | 72 hr | 96.87 | 89.82 | 94.23 | 93.54 | 92.37 | 94.26 | 94.25 |
| **5-methyl cytosine** | 4 hr | 92.01 | 91.54 | 91.16 | 91.31 | 92.31 | 91.40 | 92.23 |
| | 24 hr | 87.97 | 85.76 | 84.72 | 85.14 | 84.71 | 86.37 | 86.35 |
| | 48 hr | 87.29 | 85.94 | 85.74 | 86.18 | 86.44 | 87.10 | 88.18 |
| | 72 hr | 96.08 | 88.10 | 92.26 | 90.92 | 89.97 | 92.10 | 91.93 |
| **N1-methyl pseudouridine** | 4 hr | 92.45 | 91.43 | 91.48 | 90.41 | 92.15 | 91.44 | 91.89 |
| | 24 hr | 88.92 | 86.48 | 85.17 | 85.72 | 85.89 | 86.85 | 87.79 |
| | 48 hr | 89.84 | 86.02 | 87.52 | 85.85 | 87.38 | 86.72 | 87.81 |
| | 72 hr | 96.80 | 93.03 | 93.83 | 92.25 | 92.40 | 92.84 | 92.98 |
| **Untreated** | 4 hr | 92.77 | -- | -- | -- | -- | -- | -- |
| | 24 hr | 87.52 | -- | -- | -- | -- | -- | -- |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 48 hr | 92.95 | -- | -- | -- | -- | -- | -- |
| | 72 hr | 96.97 | -- | -- | -- | -- | -- | -- |

**Table 14. Cytotoxicity of NTPs in Combination**

| NTP Combination Cytotoxicity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Time** | **Dose** | | | | | | |
| | | **8.4 mM** | **840 uM** | **84 uM** | **8.4 uM** | **840 nM** | **84 nM** | **8.4 nM** |
| **Pseudouridine/ 5-methylcytosine / Adenine/ Guanine** | 4 hr | 92.27 | 92.04 | 91.47 | 90.86 | 90.87 | 91.10 | 91.50 |
| | 24 hr | 88.51 | 86.90 | 86.43 | 88.15 | 88.46 | 86.28 | 87.51 |
| | 48 hr | 88.30 | 87.36 | 88.58 | 88.13 | 87.39 | 88.72 | 90.55 |
| | 72 hr | 96.53 | 94.42 | 94.31 | 94.53 | 94.38 | 94.36 | 93.65 |
| **N1-methyl pseudouridine/ 5-methylcytosine / Adenine/ Guanine** | 4 hr | 92.31 | 91.71 | 91.36 | 91.15 | 91.30 | 90.86 | 91.38 |
| | 24 hr | 88.19 | 87.07 | 86.46 | 87.70 | 88.13 | 85.30 | 87.21 |
| | 48 hr | 87.17 | 86.53 | 87.51 | 85.85 | 84.69 | 87.73 | 86.79 |
| | 72 hr | 96.40 | 94.88 | 94.40 | 93.65 | 94.82 | 92.72 | 93.10 |
| **G-CSF modified mRNA** | 4 hr | na | na | na | na | na | na | 92.63 |
| | 24 hr | na | na | na | na | na | na | 87.53 |
| | 48 hr | na | na | na | na | na | na | 91.70 |
| | 72 hr | na | na | na | na | na | na | 96.36 |

**Example 69. Innate Immune Response Study in BJ Fibroblasts**

**[00985]** Human primary foreskin fibroblasts (BJ fibroblasts) were obtained from American Type Culture Collection (ATCC) (catalog # CRL-2522) and grown in Eagle's Minimum Essential Medium (ATCC, catalog # 30-2003) supplemented with 10% fetal bovine serum at 37ºC, under 5% $CO_2$. BJ fibroblasts were seeded on a 24-well plate at a density of 300,000 cells per well in 0.5 ml of culture medium. 250 ng of modified G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) fully modified with 5-methylcytosine and pseudouridine (Gen1) or fully modified with 5-methylcytosine and N1-methylpseudouridine (Gen2) having Cap0, Cap1 or no cap was transfected using Lipofectamine 2000 (Invitrogen, catalog # 11668-019), following manufacturer's protocol. Control samples of poly I:C (PIC), Lipofectamine 2000 (Lipo), natural luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) and natural G-CSF mRNA were also transfected. The cells were harvested after 18 hours, the total RNA was isolated and DNASE® treated using the RNeasy micro kit (catalog #74004) following the manufacturer's protocol. 100 ng of total RNA was used for cDNA synthesis using High Capacity

cDNA Reverse Transcription kit (catalog #4368814) following the manufacturer's protocol. The cDNA was then analyzed for the expression of innate immune response genes by quantitative real time PCR using SybrGreen in a Biorad CFX 384 instrument following manufacturer's protocol. Table 15 shows the expression level of innate immune response transcripts relative to house-keeping gene HPRT (hypoxanthine phosphoribosytransferase) and is expressed as fold-induction relative to HPRT. In the table, the panel of standard metrics includes: RIG-I is retinoic acid inducible gene 1, IL6 is interleukin-6, OAS-1 is oligoadenylate synthetase 1, IFNb is interferon-beta, AIM2 is absent in melanoma-2, IFIT-1 is interferon-induced protein with tetratricopeptide repeats 1, PKR is protein kinase R, TNFa is tumor necrosis factor alpha and IFNa is interferon alpha.

**Table 15. Innate Immune Response Transcript Levels**

| Formulation | RIG-I | IL6 | OAS-1 | IFNb | AIM2 | IFIT-1 | PKR | TNFa | IFNa |
|---|---|---|---|---|---|---|---|---|---|
| Natural Luciferase | 71.5 | 20.6 | 20.778 | 11.404 | 0.251 | 151.218 | 16.001 | 0.526 | 0.067 |
| Natural G-CSF | 73.3 | 47.1 | 19.359 | 13.615 | 0.264 | 142.011 | 11.667 | 1.185 | 0.153 |
| PIC | 30.0 | 2.8 | 8.628 | 1.523 | 0.100 | 71.914 | 10.326 | 0.264 | 0.063 |
| G-CSF Gen1-UC | 0.81 | 0.22 | 0.080 | 0.009 | 0.008 | 2.220 | 1.592 | 0.090 | 0.027 |
| G-CSF Gen1-Cap0 | 0.54 | 0.26 | 0.042 | 0.005 | 0.008 | 1.314 | 1.568 | 0.088 | 0.038 |
| G-CSF Gen1-Cap1 | 0.58 | 0.30 | 0.035 | 0.007 | 0.006 | 1.510 | 1.371 | 0.090 | 0.040 |
| G-CSF Gen2-UC | 0.21 | 0.20 | 0.002 | 0.007 | 0.007 | 0.603 | 0.969 | 0.129 | 0.005 |
| G-CSF Gen2-Cap0 | 0.23 | 0.21 | 0.002 | 0.0014 | 0.007 | 0.648 | 1.547 | 0.121 | 0.035 |
| G-CSF Gen2-Cap1 | 0.27 | 0.26 | 0.011 | 0.004 | 0.005 | 0.678 | 1.557 | 0.099 | 0.037 |
| Lipo | 0.27 | 0.53 | 0.001 | 0 | 0.007 | 0.954 | 1.536 | 0.158 | 0.064 |

**Example 70. *In vivo* detection of Innate Immune Response**

**[00986]** In an effort to distringuish the importance of different chemical modification of mRNA on *in vivo* protein production and cytokine response *in vivo*, female BALB/C mice (n=5) are injected intramuscularly with G-CSF mRNA (GCSF mRNA unmod) (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence;) with a 5'cap of Cap1 , G-CSF mRNA fully modified with 5-methylcytosine and pseudouridine (GCSF mRNA 5mc/pU), G-CSF mRNA fully modified with 5-methylcytosine and N1-methylpseudouridine with (GCSF mRNA

5mc/N1pU) or without a 5' cap (GCSF mRNA 5mc/N1 pU no cap) or a control of either R848 or 5% sucrose as described in Table 16.

**Table 16. Dosing Chart**

| Formulation | Route | Dose (ug/mouse) | Dose (ul) |
|---|---|---|---|
| GCSF mRNA unmod | I.M. | 200 | 50 |
| GCSF mRNA 5mc/pU | I.M. | 200 | 50 |
| GCSF mRNA 5mc/N1pU | I.M. | 200 | 50 |
| GCSF mRNA 5mc/N1pU no cap | I.M. | 200 | 50 |
| R848 | I.M. | 75 | 50 |
| 5% sucrose | I.M. | - | 50 |
| Untreated | I.M. | - | - |

**[00987]** Blood is collected at 8 hours after dosing. Using ELISA the protein levels of G-CSF, TNF-alpha and IFN-alpha is determined by ELISA. 8 hours after dosing, muscle is collected from the injection site and quantitative real time polymerase chain reaction (QPCR) is used to determine the mRNA levels of RIG-I, PKR, AIM-2, IFIT-1, OAS-2, MDA-5, IFN-beta, TNF-alpha, IL-6, G-CSF, CD45 in the muscle.

**Example 71. *In vivo* detection of Innate Immune Response Study**

**[00988]** Female BALB/C mice (n=5) were injected intramuscularly with G-CSF mRNA (GCSF mRNA unmod) (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence;) with a 5'cap of Cap1 , G-CSF mRNA fully modified with 5-methylcytosine and pseudouridine (GCSF mRNA 5mc/pU), G-CSF mRNA fully modified with 5-methylcytosine and N1-methylpseudouridine with (GCSF mRNA 5mc/N1pU) or without a 5' cap (GCSF mRNA 5mc/N1 pU no cap) or a control of either R848 or 5% sucrose as described in Table 17. Blood is collected at 8 hours after dosing and using ELISA the protein levels of G-CSF and interferon-alpha (IFN-alpha) is determined by ELISA and are shown in Table 17.

**[00989]** As shown in Table 17, unmodified, 5mc/pU, and 5mc/N1pU modified G-CSF mRNA resulted in human G-CSF expression in mouse serum. The uncapped 5mC/N1pU modified G-CSF mRNA showed no human G-CSF expression in serum, highlighting the importance of having a 5' cap structure for protein translation.

**[00990]** As expected, no human G-CSF protein was expressed in the R848, 5% sucrose only, and untreated groups. Importantly, significant differences were seen in cytokine production as measured by mouse IFN-alpha in the serum. As expected, unmodified G-CSF mRNA demonstrated a robust

cytokine response *in vivo* (greater than the R848 positive control). The 5mc/pU modified G-CSF mRNA did show a low but detectable cytokine response *in vivo*, while the 5mc/N1pU modified mRNA showed no detectable IFN-alpha in the serum (and same as vehicle or untreated animals).

**[00991]** Also, the response of 5mc/N1pU modified mRNA was the same regardless of whether it was capped or not. These *in vivo* results reinforce the conclusion that 1) that unmodified mRNA produce a robust innate immune response, 2) that this is reduced, but not abolished, through 100% incorporation of 5mc/pU modification, and 3) that incorporation of 5mc/N1pU modifications results in no detectable cytokine response.

**[00992]** Lastly, given that these injections are in 5% sucrose (which has no effect by itself), these result should accurately reflect the immunostimulatory potential of these modifications.

**[00993]** From the data it is evident that N1pU modified molecules produce more protein while concomitantly having little or no effect on IFN-alpha expression. It is also evident that capping is required for protein production for this chemical modification. The Protein: Cytokine Ratio of 748 as compared to the PC Ratio for the unmodified mRNA (PC=9) means that this chemical modification is far superior as related to the effects or biological implications associated with IFN-alpha.

**Table 17. Human G-CSF and Mouse IFN-alpha in serum**

| Formulation | Route | Dose (ug/mouse) | Dose (ul) | G-CSF protein (pg/ml) | IFN-alpha expression (pg/ml) | PC Ratio |
|---|---|---|---|---|---|---|
| GCSF mRNA unmod | I.M. | 200 | 50 | 605.6 | 67.01 | 9 |
| GCSF mRNA 5mc/pU | I.M. | 200 | 50 | 356.5 | 8.87 | 40 |
| GCSF mRNA5mc/N1pU | I.M. | 200 | 50 | 748.1 | 0 | 748 |
| GCSF mRNA5mc/N1pU no cap | I.M. | 200 | 50 | 6.5 | 0 | 6.5 |
| R848 | I.M. | 75 | 50 | 3.4 | 40.97 | .08 |
| 5% sucrose | I.M. | - | 50 | 0 | 1.49 | 0 |
| Untreated | I.M. | - | - | 0 | 0 | 0 |

**Example 72: *In Vivo* Delivery Using Lipoplexes**

**A. Human G-CSF Modified RNA**

**[00994]** A formulation containing 100 µg of one of two versions of modified human G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) (G-CSF fully modified with 5-methylcytosine and pseudouridine (G-CSF) or G-CSF fully modified with 5-methylcytosine and N1-methyl-pseudouridine (G-CSF-N1) lipoplexed with 30% by volume of RNAIMAX™ and delivered in 150 uL intramuscularly (I.M.) and in 225uL intravenously (I.V.) to C57/BL6 mice.

**[00995]** Three control groups were administered either 100 µg of modified luciferase mRNA (IVT cDNA sequence shown in SEQ ID NO: 2; mRNA sequence shown in SEQ ID NO: 3, polyA tail of approximately 160 nucleotides not shown in sequence, 5'cap, Cap1, fully modified with 5-methylcytosine at each cytosine and pseudouridine replacement at each uridine site) intramuscularly (Luc-unsp I.M.) or 150 µg of modified luciferase mRNA intravenously (Luc-unsp I.V.) or 150 uL of the formulation buffer intramuscularly (Buffer I.M.). 6 hours after administration of a formulation, serum was collected to measure the amount of human G-CSF protein in the mouse serum by human G-CSF ELISA and the results are shown in Table 18.

**[00996]** These results demonstrate that both 5-methylcytosine/pseudouridine and 5-methylcytosine/N1-methylpseudouridine modified human G-CSF mRNA can result in specific human G-CSF protein expression in serum when delivered via I.V. or I.M. route of administration in a lipoplex formulation.

**Table 18. Human G-CSF in Serum (I.M. and I.V. Injection Route)**

| Formulation | Route | G-CSF (pg/ml) |
|---|---|---|
| G-CSF | I.M. | 85.6 |
| G-CSF-N1 | I.M. | 40.1 |
| G-CSF | I.V. | 31.0 |
| G-CSF-N1 | I.V. | 6.1 |
| Luc-unsp | I.M. | 0.0 |
| Luc-unsp | I.V. | 0.0 |
| Buffer | I.M. | 0.0 |

**B. Human G-CSF Modified RNA Comparison**

**[00997]** A formulation containing 100 µg of either modified human G-CSF mRNA lipoplexed with 30% by volume of RNAIMAX™ with a 5-methylcytosine (5mc) and a pseudouridine (ψ) modification (G-CSF-Gen1-Lipoplex), modified human G-CSF mRNA with a 5mc and ψ modification in saline (G-CSF-Gen1-Saline), modified human G-CSF mRNA with a N1-5-methylcytosine (N1-5mc) and a ψ modification lipoplexed with 30% by volume of RNAIMAX™ (G-CSF-Gen2-Lipoplex), modified human G-CSF mRNA with a N1-5mc and ψ modification in saline (G-CSF-Gen2-Saline), modified luciferase with a 5mc and ψ modification lipoplexed with 30% by volume of RNAIMAX™ (Luc-Lipoplex), or luciferase mRNA fully modified with 5mc and ψ modifications in saline (Luc-Saline) was delivered intramuscularly (I.M.) or subcutaneously (S.C.) and a control group for each method of administration was giving a dose of 80uL of the formulation buffer (F. Buffer) to C57/BL6 mice. 13 hours post injection serum and tissue from the site of

injection were collected from each mouse and analyzed by G-CSF ELISA to compare human G-CSF protein levels. The results of the human G-CSF protein in mouse serum from the intramuscular administration and the subcutaneous administration results are shown in Table 19.

**[00998]** These results demonstrate that 5-methylcytosine/pseudouridine and 5-methylcytosine/N1-methylpseudouridine modified human G-CSF mRNA can result in specific human G-CSF protein expression in serum when delivered via I.M. or S.C. route of administration whether in a saline formulation or in a lipoplex formulation. As shown in Table 19, 5-methylcytosine/N1-methylpseudouridine modified human G-CSF mRNA generally demonstrates increased human G-CSF protein production relative to 5-methylcytosine/pseudouridine modified human G-CSF mRNA.

**Table 19. Human G-CSF Protein in Mouse Serum**

| Formulation | G-CSF (pg/ml) | |
|---|---|---|
| | I.M. Injection Route | S.C. Injection Route |
| G-CSF-Gen1-Lipoplex | 13.988 | 42.855 |
| GCSF-Gen1-saline | 9.375 | 4.614 |
| GCSF-Gen2-lipoplex | 75.572 | 32.107 |
| GCSF-Gen2-saline | 20.190 | 45.024 |
| Luc lipoplex | 0 | 3.754 |
| Luc saline | 0.0748 | 0 |
| F. Buffer | 4.977 | 2.156 |

**Example 73. Multi-Site Administration: Intramuscular and Subcutaneous**

**[00999]** Human G-CSF modified mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) modified as either Gen1 or Gen2 (5-methylcytosine (5mc) and a pseudouridine (ψ) modification, G-CSF-Gen1; or N1-5-methylcytosine (N1-5mc) and a ψ modification, G-CSF-Gen2) and formulated in saline were delivered to mice via intramuscular (IM) or subcutaneous (SC) injection. Injection of four doses or 2x 50ug (two sites) daily for three days (24 hrs interval) was performed. The fourth dose was administered 6 hrs before blood collection and CBC analysis. Controls included Luciferase (cDNA sequence for IVT shown in SEQ ID NO: 2; mRNA sequence shown in SEQ ID NO: 3, polyA tail of approximately 160 nucleotides not shown in sequence, 5'cap, Cap1, fully modified with 5-methylcytosine at each cytosine and pseudouridine replacement at each uridine site) or the formulation buffer (F.Buffer). The mice were bled at 72 hours after the first mRNA injection (6 hours after the last mRNA dose) to determine the effect of mRNA-encoded human G-CSF on the neutrophil count. The dosing regimen is shown in Table 20 as are the resulting neutrophil counts (thousands/uL). In Table 20, an asterisks (*) indicate statistical significance at $p<0.05$.

**[001000]** For intramuscular administration, the data reveal a four fold increase in neutrophil count above control at day 3 for the Gen1 G-CSF mRNA and a two fold increase for the Gen2 G-CSF mRNA. For subcutaneous administration, the data reveal a two fold increase in neutrophil count above control at day 3 for the Gen2 G-CSF mRNA.

**[001001]** These data demonstrate that both 5-methylcytidine/pseudouridine and 5-methylcytidine/N1-methylpseudouridine-modified mRNA can be biologically active, as evidenced by specific increases in blood neutrophil counts.

**Table 20. Dosing Regimen**

| Gr. | Treatment | Route | N = | Dose (µg/mouse) | Dose Vol. (µl/mouse) | Dosing Vehicle | Neutrophil Thous/uL |
|---|---|---|---|---|---|---|---|
| **1** | G-CSF (Gen1) | I.M | 5 | 2x50ug (four doses) | 50 | F. buffer | 840* |
| **2** | G-CSF (Gen1) | S.C | 5 | 2x50ug (four doses) | 50 | F. buffer | 430 |
| **3** | G-CSF (Gen2) | I.M | 5 | 2x50ug (four doses) | 50 | F. buffer | 746* |
| **4** | G-CSF (Gen2) | S.C | 5 | 2x50ug (four doses) | 50 | F. buffer | 683 |
| **5** | Luc (Gen1) | I.M. | 5 | 2x50ug (four doses) | 50 | F. buffer | 201 |
| **6** | Luc (Gen1) | S.C. | 5 | 2x50ug (four doses) | 50 | F. buffer | 307 |
| **7** | Luc (Gen2) | I.M | 5 | 2x50ug (four doses) | 50 | F. buffer | 336 |
| **8** | Luc (Gen2) | S.C | 5 | 2x50ug (four doses) | 50 | F. buffer | 357 |
| **9** | F. Buffer | I.M | 4 | 0 (four doses) | 50 | F. buffer | 245 |
| **10** | F. Buffer | S.C. | 4 | 0 (four doses) | 50 | F. buffer | 509 |
| **11** | Untreated | - | 4 | | | - | 312 |

**Example 74. Intravenous Administration**

**[001002]** Human G-CSF modified mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) modified with 5-methylcytosine (5mc) and a pseudouridine (ψ) modification (Gen1); or having no modifications and formulated in 10% lipoplex (RNAIMAX™) were delivered to mice at a dose of 50 ug RNA and in a volume of 100 ul via intravenous (IV) injection at days 0, 2 and 4. Neutrophils were measured at days 1, 5 and 8. Controls included non-specific mammalian RNA or the formulation buffer alone (F.Buffer). The mice were bled at days 1, 5 and 8 to determine the effect of mRNA-encoded human G-CSF to increase neutrophil count. The dosing regimen is shown in Table 21 as are the resulting neutrophil counts (thousands/uL; K/uL).

**[001003]** For intravenous administration, the data reveal a four to five fold increase in neutrophil count above control at day 5 with G-CSF modified mRNA but not with unmodified G-CSF mRNA

or non-specific controls. Blood count returned to baseline four days after the final injection. No other changes in leukocyte populations were observed.

**[001004]** In Table 21, an asterisk (*) indicates statistical significance at $p<0.001$ compared to buffer.

**[001005]** These data demonstrate that lipoplex-formulated 5-methylcytidine/pseudouridine-modified mRNA can be biologically active, when delivered through an I.V. route of administration as evidenced by specific increases in blood neutrophil counts. No other cell subsets were significantly altered. Unmodified G-CSF mRNA similarly administered showed no pharmacologic effect on neutrophil counts.

**Table 21. Dosing Regimen**

| Gr. | Treatment | N | Dose Vol. (µl/mouse) | Dosing Vehicle | Neutrophil K/uL |
|---|---|---|---|---|---|
| **1** | G-CSF (Gen1) Day 1 | 5 | 100 | 10% lipoplex | 2.91 |
| **2** | G-CSF (Gen1) Day 5 | 5 | 100 | 10% lipoplex | 5.32* |
| **3** | G-CSF (Gen1) Day 8 | 5 | 100 | 10% lipoplex | 2.06 |
| **4** | G-CSF (no modification) Day 1 | 5 | 100 | 10% lipoplex | 1.88 |
| **5** | G-CSF (no modification) Day 5 | 5 | 100 | 10% lipoplex | 1.95 |
| **6** | G-CSF (no modification) Day 8 | 5 | 100 | 10% lipoplex | 2.09 |
| **7** | RNA control Day 1 | 5 | 100 | 10% lipoplex | 2.90 |
| **8** | RNA control Day 5 | 5 | 100 | 10% lipoplex | 1.68 |
| **9** | RNA control Day 8 | 4 | 100 | 10% lipoplex | 1.72 |
| **10** | F. Buffer Day 1 | 4 | 100 | 10% lipoplex | 2.51 |
| **11** | F. Buffer Day 5 | 4 | 100 | 10% lipoplex | 1.31 |
| **12** | F. Buffer Day 8 | 4 | 100 | 10% lipoplex | 1.92 |

**Example 75: Routes of Administration**

**[001006]** Studies were performed to investigate split dosing using different routes of administration. Studies utilizing multiple subcutaneous or intramuscular injection sites at one time point were designed and performed to investigate ways to increase modified mRNA drug exposure and improve protein production. In addition to detection of the expressed protein product, an assessment of the physiological function of proteins was also determined through analyzing samples from the animal tested.

**[001007]** Surprisingly, it has been determined that split dosing of modified mRNA produces greater protein production and phenotypic responses than those produced by single unit dosing or multi-dosing schemes.

**[001008]** The design of a split dose experiment involved using human erythropoietin (EPO) modified mRNA (mRNA sequence shown in SEQ ID NO: 5; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) or luciferase modified mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) administered in buffer alone or formulated with 30% lipoplex (RNAIMAX$^{TM}$). The dosing vehicle (buffer) consisted of 150mM NaCl, 2 mM $CaCl_2$, 2 mM $Na^+$-phosphate (1.4mM monobasic sodium phosphate; 0.6mM dibasic sodium phosphate), and 0.5 mM EDTA, pH 6.5. The pH was adjusted using sodium hydroxide and the final solution was filter sterilized. The mRNA was modified with 5methylC (5meC) at each cytosine and pseudouridine replacement at each uridine site.

**[001009]** 4 mice per group were dosed intramuscularly (I.M.), intravenously (I.V.) or subcutaneously (S.C.) by the dosing chart outlined in Table 22. Serum was collected 13 hours post injection from all mice, tissue was collected from the site of injection from the intramuscular and subcutaneous group and the spleen, liver and kidneys were collected from the intravenous group. The results from the intramuscular group and the subcutaneous group results are shown in Table 23.

**Table 22. Dosing Chart**

| Group | Treatment | Route | Dose of modified mRNA | Total Dose | Dosing Vehicle |
|---|---|---|---|---|---|
| 1 | Lipoplex-human EPO modified mRNA | I.M. | 4 x 100 ug + 30% Lipoplex | 4x70 ul | Lipoplex |
| 2 | Lipoplex-human EPO modified mRNA | I.M. | 4 x 100 ug | 4x70 ul | Buffer |
| 3 | Lipoplex-human EPO modified mRNA | S.C. | 4 x 100 ug + 30% Lipoplex | 4x70 ul | Lipoplex |
| 4 | Lipoplex-human EPO modified mRNA | S.C. | 4 x 100 ug | 4x70 ul | Buffer |
| 5 | Lipoplex-human EPO modified mRNA | I.V. | 200 ug + 30% Lipoplex | 140 ul | Lipoplex |
| 6 | Lipoplexed-Luciferase modified mRNA | I.M. | 100 ug + 30% Lipoplex | 4x70 ul | Lipoplex |
| 7 | Lipoplexed-Luciferase modified mRNA | I.M. | 100 ug | 4x70 ul | Buffer |
| 8 | Lipoplexed-Luciferase modified mRNA | S.C. | 100 ug + 30% Lipoplex | 4x70 ul | Lipoplex |
| 9 | Lipoplexed-Luciferase modified mRNA | S.C. | 100 ug | 4x70 ul | Buffer |
| 10 | Lipoplexed-human EPO modified mRNA | I.V. | 200 ug + 30% Lipoplex | 140 ul | Lipoplex |
| 11 | Formulation Buffer | I.M. | 4x multi dosing | 4x70 ul | Buffer |

**Table 23. Human EPO Protein in Mouse Serum (I.M. Injection Route)**

| Formulation | EPO (pg/ml) | |
|---|---|---|
| | I.M. Injection Route | S.C. Injection Route |
| Epo-Lipoplex | 67.1 | 2.2 |
| Luc-Lipoplex | 0 | 0 |
| Epo-Saline | 100.9 | 11.4 |
| Luc-Saline | 0 | 0 |
| Formulation Buffer | 0 | 0 |

**Example 76: *In Vivo* Delivery Using Varying Lipid Ratios**

**[001010]** Modified mRNA was delivered to C57/BL6 mice to evaluate varying lipid ratios and the resulting protein expression. Formulations of 100µg modified human EPO mRNA (mRNA sequence shown in SEQ ID NO: 5; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1; fully modified with 5-methylcytosine and pseudouridine) lipoplexed with 10%, 30% or 50% RNAIMAX™, 100µg modified luciferase mRNA (IVT cDNA sequence shown in SEQ ID NO: 2; mRNA sequence shown in SEQ ID NO: 3, polyA tail of approximately 160 nucleotides not shown in sequence, 5'cap, Cap1, fully modified with 5-methylcytosine at each cytosine and pseudouridine replacement at each uridine site) lipoplexed with 10%, 30% or 50% RNAIMAX™ or a formulation buffer were administered intramuscularly to mice in a single 70 µl dose. Serum was collected 13 hours post injection to undergo a human EPO ELISA to determine the human EPO protein level in each mouse. The results of the human EPO ELISA, shown in Table 24, show that modified human EPO expressed in the muscle is secreted into the serum for each of the different percentage of RNAIMAX™.

**Table 24. Human EPO Protein in Mouse Serum (IM Injection Route)**

| Formulation | EPO (pg/ml) |
|---|---|
| Epo + 10% RNAiMAX | 11.4 |
| Luc + 10% RNAiMAX | 0 |
| Epo + 30% RNAiMAX | 27.1 |
| Luc + 30% RNAiMAX | 0 |
| Epo + 50% RNAiMAX | 19.7 |
| Luc + 50% RNAiMAX | 0 |
| F. Buffer | 0 |

**Example 77: *In Vivo* Delivery of Modified RNA in Rats**

**[001011]** Protein production of modified mRNA was evaluated by delivering modified G-CSF mRNA or modified Factor IX mRNA to female Sprague Dawley rats (n=6). Rats were injected with 400 ug in 100 ul of G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of

approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) fully modified with 5-methylcytosine and pseudouridine (G-CSF Gen1), G-CSF mRNA fully modified with 5-methylcytosine and N1-methylpseudouridine (G-CSF Gen2) or Factor IX mRNA (mRNA sequence shown in SEQ ID NO: 6; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) fully modified with 5-methylcytosine and pseudouridine (Factor IX Gen1) reconstituted from the lyophilized form in 5% sucrose. Blood was collected 8 hours after injection and the G-CSF protein level in serum was measured by ELISA. Table 25 shows the G-CSF protein levels in serum after 8 hours.

**[001012]** These results demonstrate that both G-CSF Gen 1 and G-CSF Gen 2 modified mRNA can produce human G-CSF protein in a rat following a single intramuscular injection, and that human G-CSF protein production is improved when using Gen 2 chemistry over Gen 1 chemistry.

**Table 25. G-CSF Protein in Rat Serum (I.M. Injection Route)**

| Formulation | G-CSF protein (pg/ml) |
|---|---|
| G-CSF Gen1 | 19.37 |
| G-CSF Gen2 | 64.72 |
| Factor IX Gen 1 | 2.25 |

**Example 78. Chemical Modification: *In vitro* studies**

**A. *In vitro* Screening in PBMC**

**[001013]** 500 ng of G-CSF (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) mRNA fully modified with the chemical modification outlined Tables 26 and 27 was transfected with 0.4 uL Lipofectamine 2000 into peripheral blood mononuclear cells (PBMC) from three normal blood donors. Control samples of LPS, R848, P(I)P(C) and mCherry (mRNA sequence shown in SEQ ID NO: 4; polyA tail of approximately 160 nucleotides not shown in sequence, 5'cap, Cap1; fully modified with 5-methylcytosine and pseudouridine) were also analyzed. The supernatant was harvested and stored frozen until analyzed by ELISA to determine the G-CSF protein expression, and the induction of the cytokines interferon-alpha (IFN-α) and tumor necrosis factor alpha (TNF-α). The protein expression of G-CSF is shown in Table 26, the expression of IFN-α and TNF-α is shown in Table 27.

**[001014]** The data in Table 26 demonstrates that many, but not all, chemical modifications can be used to productively produce human G-CSF in PBMC. Of note, 100% N1-methylpseudouridine substitution demonstrates the highest level of human G-CSF production (almost 10-fold higher than pseudouridine itself). When N1-methylpseudouridine is used in combination with 5-methylcytidine

a high level of human G-CSF protein is also produced (this is also higher than when pseudouridine is used in combination with 5 methylcytidine).

**[001015]** Given the inverse relationship between protein production and cytokine production in PBMC, a similar trend is also seen in Table 27, where 100% substitution with N1-methylpseudouridine results no cytokine induction (similar to transfection only controls) and pseudouridine shows detectable cytokine induction which is above background.

**[001016]** Other modifications such as N6-methyladenosine and α-thiocytidine appear to increase cytokine stimulation.

**Table 26. Chemical Modifications and G-CSF Protein Expression**

| Chemical Modifications | G-CSF Protein Expression (pg/ml) | | |
|---|---|---|---|
| | Donor 1 | Donor 2 | Donor 3 |
| Pseudouridine | 2477 | 1,909 | 1,498 |
| 5-methyluridine | 318 | 359 | 345 |
| N1-methylpseudouridine | 21,495 | 16,550 | 12,441 |
| 2-thiouridine | 932 | 1,000 | 600 |
| 4-thiouridine | 5 | 391 | 218 |
| 5-methoxyuridine | 2,964 | 1,832 | 1,800 |
| 5-methylcytosine and pseudouridine (1st set) | 2,632 | 1,955 | 1,373 |
| 5-methylcytosine and N1-methylpseudouridine (1st set) | 10,232 | 7,245 | 6,214 |
| 2'Fluoroguanosine | 59 | 186 | 177 |
| 2'Fluorouridine | 118 | 209 | 191 |
| 5-methylcytosine and pseudouridine (2nd set) | 1,682 | 1,382 | 1,036 |
| 5-methylcytosine and N1-methylpseudouridine (2nd set) | 9,564 | 8,509 | 7,141 |
| 5-bromouridine | 314 | 482 | 291 |
| 5-(2-carbomethoxyvinyl)uridine | 77 | 286 | 177 |
| 5-[3(1-E-propenylamino)uridine | 541 | 491 | 550 |
| α-thiocytidine | 105 | 264 | 245 |
| 5-methylcytosine and pseudouridine (3rd set) | 1,595 | 1,432 | 955 |
| N1-methyladenosine | 182 | 177 | 191 |
| N6-methyladenosine | 100 | 168 | 200 |
| 5-methylcytidine | 291 | 277 | 359 |
| N4-acetylcytidine | 50 | 136 | 36 |
| 5-formylcytidine | 18 | 205 | 23 |
| 5-methylcytosine and pseudouridine (4th set) | 264 | 350 | 182 |
| 5-methylcytosine and N1-methylpseudouridine (4th set) | 9,505 | 6,927 | 5,405 |
| LPS | 1,209 | 786 | 636 |
| mCherry | 5 | 168 | 164 |
| R848 | 709 | 732 | 636 |
| P(I)P(C) | 5 | 186 | 182 |

**Table 27. Chemical Modifications and Cytokine Expression**

| Chemical Modifications | IFN-α Expression (pg/ml) | TNF-α Expression (pg/ml) |
|---|---|---|

| | Donor 1 | Donor 2 | Donor 3 | Donor 1 | Donor 2 | Donor 3 |
|---|---|---|---|---|---|---|
| Pseudouridine | 120 | 77 | 171 | 36 | 81 | 126 |
| 5-methyluridine | 245 | 135 | 334 | 94 | 100 | 157 |
| N1-methylpseudouridine | 26 | 75 | 138 | 101 | 106 | 134 |
| 2-thiouridine | 100 | 108 | 154 | 133 | 133 | 141 |
| 4-thiouridine | 463 | 258 | 659 | 169 | 126 | 254 |
| 5-methoxyuridine | 0 | 64 | 133 | 39 | 74 | 111 |
| 5-methylcytosine and pseudouridine (1st set) | 88 | 94 | 148 | 64 | 89 | 121 |
| 5-methylcytosine and N1-methylpseudouridine (1st set) | 0 | 60 | 136 | 54 | 79 | 126 |
| 2'Fluoroguanosine | 107 | 97 | 194 | 91 | 94 | 141 |
| 2'Fluorouridine | 158 | 103 | 178 | 164 | 121 | 156 |
| 5-methylcytosine and pseudouridine (2nd set) | 133 | 92 | 167 | 99 | 111 | 150 |
| 5-methylcytosine and N1-methylpseudouridine (2nd set) | 0 | 66 | 140 | 54 | 97 | 149 |
| 5-bromouridine | 95 | 86 | 181 | 87 | 106 | 157 |
| 5-(2-carbomethoxyvinyl)uridine | 0 | 61 | 130 | 40 | 81 | 116 |
| 5-[3(1-E-propenylamino)uridine | 0 | 58 | 132 | 71 | 90 | 119 |
| α-thiocytidine | 1,138 | 565 | 695 | 300 | 273 | 277 |
| 5-methylcytosine and pseudouridine (3rd set) | 88 | 75 | 150 | 84 | 89 | 130 |
| N1-methyladenosine | 322 | 255 | 377 | 256 | 157 | 294 |
| N6-methyladenosine | 1,935 | 1,065 | 1,492 | 1,080 | 630 | 857 |
| 5-methylcytidine | 643 | 359 | 529 | 176 | 136 | 193 |
| N4-acetylcytidine | 789 | 593 | 431 | 263 | 67 | 207 |
| 5-formylcytidine | 180 | 93 | 88 | 136 | 30 | 40 |
| 5-methylcytosine and pseudouridine (4th set) | 131 | 28 | 18 | 53 | 24 | 29 |
| 5-methylcytosine and N1-methylpseudouridine (4th set) | 0 | 0 | 0 | 36 | 14 | 13 |
| LPS | 0 | 67 | 146 | 7,004 | 3,974 | 4,020 |
| mCherry | 100 | 75 | 143 | 67 | 100 | 133 |
| R848 | 674 | 619 | 562 | 11,179 | 8,546 | 9,907 |
| P(I)P(C) | 470 | 117 | 362 | 249 | 177 | 197 |

**B. *In vitro* Screening in HeLa Cells**

**[001017]** The day before transfection, 20,000 HeLa cells (ATCC no. CCL-2; Manassas, VA) were harvested by treatment with Trypsin-EDTA solution (LifeTechnologies, Grand Island, NY) and seeded in a total volume of 100ul EMEM medium (supplemented with 10%FCS and 1x Glutamax) per well in a 96-well cell culture plate (Corning, Manassas, VA). The cells were grown at 37oG in

5% $CO_2$ atmosphere overnight. Next day, 83 ng of Luciferase modified RNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) with the chemical modification described in Table 28, were diluted in 10ul final volume of OPTI-MEM (LifeTechnologies, Grand Island, NY). Lipofectamine 2000 (LifeTechnologies, Grand Island, NY) was used as transfection reagent and 0.2 ul were diluted in 10 ul final volume of OPTI-MEM. After 5 minutes of incubation at room temperature, both solutions were combined and incubated an additional 15 minute at room temperature. Then the 20ul combined solution was added to the 100ul cell culture medium containing the HeLa cells and incubated at room temperature.

**[001018]** After 18 to 22 hours of incubation cells expressing luciferase were lysed with 100 ul of Passive Lysis Buffer (Promega, Madison, WI) according to manufacturer instructions. Aliquots of the lysates were transferred to white opaque polystyrene 96-well plates (Corning, Manassas, VA) and combined with 100 ul complete luciferase assay solution (Promega, Madison, WI). The lysate volumes were adjusted or diluted until no more than 2 mio relative light units (RLU) per well were detected for the strongest signal producing samples and the RLUs for each chemistry tested are shown in Table 28. The plate reader was a BioTek Synergy H1 (BioTek, Winooski, VT).The background signal of the plates without reagent was about 200 relative light units per well.

**[001019]** These results demonstrate that many, but not all, chemical modifications can be used to productively produce human G-CSF in HeLa cells. Of note, 100% N1-methylpseudouridine substitution demonstrates the highest level of human G-CSF production.

**Table 28. Relative Light Units of Luciferase**

| Chemical Modification | RLU |
|---|---|
| N6-methyladenosine (m6a) | 534 |
| 5-methylcytidine (m5c) | 138,428 |
| N4-acetylcytidine (ac4c) | 235,412 |
| 5-formylcytidine (f5c) | 436 |
| 5-methylcytosine/pseudouridine, test A1 | 48,659 |
| 5-methylcytosine/N1-methylpseudouridine, test A1 | 190,924 |
| Pseudouridine | 655,632 |
| 1-methylpseudouridine (m1u) | 1,517,998 |
| 2-thiouridine (s2u) | 3387 |
| 5-methoxyuridine (mo5u) | 253,719 |
| 5-methylcytosine/pseudouridine, test B1 | 317,744 |
| 5-methylcytosine/N1-methylpseudouridine, test B1 | 265,871 |
| 5-Bromo-uridine | 43,276 |

| | |
|---|---|
| 5 (2 carbovinyl) uridine | 531 |
| 5 (3-1E propenyl Amino) uridine | 446 |
| 5-methylcytosine/pseudouridine, test A2 | 295,824 |
| 5-methylcytosine/N1-methylpseudouridine, test A2 | 233,921 |
| 5-methyluridine | 50,932 |
| α-Thio-cytidine | 26,358 |
| 5-methylcytosine/pseudouridine, test B2 | 481,477 |
| 5-methylcytosine/N1-methylpseudouridine, test B2 | 271,989 |
| 5-methylcytosine/pseudouridine, test A3 | 438,831 |
| 5-methylcytosine/N1-methylpseudouridine, test A3 | 277,499 |
| Unmodified Luciferase | 234,802 |

**C. *In vitro* Screening in Rabbit Reticulocyte Lysates**

**[001020]** Luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) was modified with the chemical modification listed in Table 29 and were diluted in sterile nuclease-free water to a final amount of 250 ng in 10 ul. The diluted luciferase was added to 40 ul of freshly prepared Rabbit Reticulocyte Lysate and the *in vitro* translation reaction was done in a standard 1.5 mL polypropylene reaction tube (Thermo Fisher Scientific, Waltham, MA) at 30ºC in a dry heating block. The translation assay was done with the Rabbit Reticulocyte Lysate (nuclease-treated) kit (Promega, Madison, WI) according to the manufacturer's instructions. The reaction buffer was supplemented with a one-to-one blend of provided amino acid stock solutions devoid of either Leucine or Methionine resulting in a reaction mix containing sufficient amounts of both amino acids to allow effective *in vitro* translation.

**[001021]** After 60 minutes of incubation, the reaction was stopped by placing the reaction tubes on ice. Aliquots of the *in vitro* translation reaction containing luciferase modified RNA were transferred to white opaque polystyrene 96-well plates (Corning, Manassas, VA) and combined with 100ul complete luciferase assay solution (Promega, Madison, WI). The volumes of the *in vitro* translation reactions were adjusted or diluted until no more than 2 mio relative light units (RLUs) per well were detected for the strongest signal producing samples and the RLUs for each chemistry tested are shown in Table 29. The plate reader was a BioTek Synergy H1 (BioTek, Winooski, VT).The background signal of the plates without reagent was about 200 relative light units per well.

**[001022]** These cell-free translation results very nicely correlate with the protein production results in HeLa, with the same modifications generally working or not working in both systems. One notable exception is 5-formylcytidine modified luciferase mRNA which worked in the cell-free translation system, but not in the HeLa cell-based transfection system. A similar difference between the two assays was also seen with 5-formylcytidine modified G-CSF mRNA.

**Table 29. Relative Light Units of Luciferase**

| Chemical Modification | RLU |
|---|---|
| N6-methyladenosine (m6a) | 398 |
| 5-methylcytidine (m5c) | 152,989 |
| N4-acetylcytidine (ac4c) | 60,879 |
| 5-formylcytidine (f5c) | 55,208 |
| 5-methylcytosine/pseudouridine, test A1 | 349,398 |
| 5-methylcytosine/N1-methylpseudouridine, test A1 | 205,465 |
| Pseudouridine | 587,795 |
| 1-methylpseudouridine (m1u) | 589,758 |
| 2-thiouridine (s2u) | 708 |
| 5-methoxyuridine (mo5u) | 288,647 |
| 5-methylcytosine/pseudouridine, test B1 | 454,662 |
| 5-methylcytosine/N1-methylpseudouridine, test B1 | 223,732 |
| 5-Bromo-uridine | 221,879 |
| 5 (2 carbovinyl) uridine | 225 |
| 5 (3-1E propenyl Amino) uridine | 211 |
| 5-methylcytosine/pseudouridine, test A2 | 558,779 |
| 5-methylcytosine/N1-methylpseudouridine, test A2 | 333,082 |
| 5-methyluridine | 214,680 |
| α-Thio-cytidine | 123,878 |
| 5-methylcytosine/pseudouridine, test B2 | 487,805 |
| 5-methylcytosine/N1-methylpseudouridine, test B2 | 154,096 |
| 5-methylcytosine/pseudouridine, test A3 | 413,535 |
| 5-methylcytosine/N1-methylpseudouridine, test A3 | 292,954 |
| Unmodified Luciferase | 225,986 |

**Example 79. Chemical Modification: *In vivo* studies**

**A. *In vivo* Screening of G-CSF Modified mRNA**

**[001023]** Balb-C mice (n=4) are intramuscularly injected in each leg with modified G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown

in sequence; 5'cap, Cap1), fully modified with the chemical modifications outlined in Table 30, is formulated in 1xPBS. A control of luciferase modified mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1; fully modified with pseudouridine and 5-methylcytosine) and a control of PBS are also tested. After 8 hours serum is collected to determine G-CSF protein levels cytokine levels by ELISA.

**Table 30. G-CSF**

| mRNA | Chemical Modifications |
|---|---|
| G-CSF | Pseudouridine |
| G-CSF | 5-methyluridine |
| G-CSF | 2-thiouridine |
| G-CSF | 4-thiouridine |
| G-CSF | 5-methoxyuridine |
| G-CSF | 2'-fluorouridine |
| G-CSF | 5-bromouridine |
| G-CSF | 5-[3(1-E-propenylamino)uridine) |
| G-CSF | alpha-thio-cytidine |
| G-CSF | 5-methylcytidine |
| G-CSF | N4-acetylcytidine |
| G-CSF | Pseudouridine and 5-methylcytosine |
| G-CSF | N1-methylpseudouridine and 5-methylcytosine |
| Luciferase | Pseudouridine and 5-methylcytosine |
| PBS | None |

**B. *In vivo* Screening of Luciferase Modified mRNA**

**[001024]** Balb-C mice (n=4) were subcutaneously injected with 200ul containing 42 to 103 ug of modified luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1), fully modified with the chemical modifications outlined in Table 31, was formulated in 1xPBS. A control of PBS was also tested. The dosages of the modified luciferase mRNA is also outlined in Table 31. 8 hours after dosing the mice were imaged to determine luciferase expression. Twenty minutes prior to imaging, mice were injected intraperitoneally with a D-luciferin solution at 150 mg/kg. Animals were then anesthetized and images were acquired with an IVIS Lumina II imaging system (Perkin Elmer). Bioluminescence was measured as total flux (photons/second) of the entire mouse.

**[001025]** As demonstrated in Table 31, all luciferase mRNA modified chemistries demonstrated *in vivo* activity, with the exception of 2'-fluorouridine. In addition 1-methylpseudouridine modified mRNA demonstrated very high expression of luciferase (5-fold greater expression than pseudouridine containing mRNA).

[001026]

**Table 31. Luciferase Screening**

| mRNA | Chemical Modifications | Dose (ug) of mRNA | Dose volume (ml) | Luciferase expression (photon/second) |
|---|---|---|---|---|
| Luciferase | 5-methylcytidine | 83 | 0.72 | 1.94E+07 |
| Luciferase | N4-acetylcytidine | 76 | 0.72 | 1.11E07 |
| Luciferase | Pseudouridine | 95 | 1.20 | 1.36E+07 |
| Luciferase | 1-methylpseudouridine | 103 | 0.72 | 7.40E+07 |
| Luciferase | 5-methoxyuridine | 95 | 1.22 | 3.32+07 |
| Luciferase | 5-methyluridine | 94 | 0.86 | 7.42E+06 |
| Luciferase | 5-bromouridine | 89 | 1.49 | 3.75E+07 |
| Luciferase | 2'-fluoroguanosine | 42 | 0.72 | 5.88E+05 |
| Luciferase | 2'-fluorocytidine | 47 | 0.72 | 4.21E+05 |
| Luciferase | 2'-flurorouridine | 59 | 0.72 | 3.47E+05 |
| PBS | None | - | 0.72 | 3.16E+05 |

**Example 80. *In vivo* Screening of Combination Luciferase Modified mRNA**

**[001027]** Balb-C mice (n=4) were subcutaneously injected with 200ul of 100 ug of modified luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1), fully modified with the chemical modifications outlined in Table 32, was formulated in 1x PBS. A control of PBS was also tested. The dosages of the modified luciferase mRNA is also outlined in Table 29. 8 hours after dosing the mice were imaged to determine luciferase expression. Twenty minutes prior to imaging, mice were injected intraperitoneally with a D-luciferin solution at 150 mg/kg. Animals were then anesthetized and images were acquired with an IVIS Lumina II imaging system (Perkin Elmer). Bioluminescence was measured as total flux (photons/second) of the entire mouse.

**[001028]** As demonstrated in Table 32, all luciferase mRNA modified chemistries (in combination) demonstrated *in vivo* activity. In addition the presence of N1-methylpseudouridine in the modified mRNA (with N4-acetylcytidine or 5 methylcytidine) demonstrated higher expression than when the same combinations where tested using with pseudouridine. Taken together, these data demonstrate that N1-methylpseudouridine containing luciferase mRNA results in improved protein expression in vivo whether used alone (Table 31) or when used in combination with other modified nulceotides (Table 32).

**Table 32. Luciferase Screening Combinations**

| mRNA | Chemical Modifications | Luciferase expression (photon/second) |
|---|---|---|

| | | |
|---|---|---|
| Luciferase | N4-acetylcytidine/pseudouridine | 4.18E+06 |
| Luciferase | N4-acetylcytidine/N1-methylpseudouridine | 2.88E+07 |
| Luciferase | 5-methylcytidine/5-methoxyuridine | 3.48E+07 |
| Luciferase | 5-methylcytidine/5-methyluridine | 1.44E+07 |
| Luciferase | 5-methylcytidine/where 50% of the uridine is replaced with 2-thiouridine | 2.39E+06 |
| Luciferase | 5-methylcytidine/pseudouridine | 2.36E+07 |
| Luciferase | 5-methylcytidine/N1-methyl-pseudouridine | 4.15E+07 |
| PBS | None | 3.59E+05 |

**Example 81. Stability of Modified RNA**

**A. Storage of Modified RNA**

**[001029]** Stability experiments were conducted to obtain a better understanding of storage conditions to retain the integrity of modified RNA. Unmodified G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1), G-CSF mRNA fully modified with 5-methylcytosine and pseudouridine and G-CSF mRNA fully modified with 5-methylcytosine and pseudouridine lipoplexed with 0.75% by volume of RNAIMAX™ was stored at 50ºC, 40ºC, 37ºC, 25ºC, 4ºC or -20ºC. After the mRNA had been stored for 0 hours, 2 hours, 6 hours, 24 hours, 48 hours, 5 days and 14 days, the mRNA was analyzed by gel electrophoresis using a Bio-Rad EXPERION™ system. The modified, unmodified and lipoplexed G-CSF mRNA was also stored in RNASTABLE® (Biomatrica, Inc. San Diego, CA) at 40ºC or water at -80 ºC or 40ºC for 35 days before being analyzed by gel electrophoresis.

**[001030]** All mRNA samples without stabilizer were stable after 2 weeks after storage at 4ºC or -20ºC. Modified G-CSF mRNA, with or without lipoplex, was more stable than unmodified G-CSF when stored at 25ºC (stable out to 5 days versus 48 hours), 37ºC (stable out to 24 hours versus 6 hours) and 50ºC (stable out to 6 hours versus 2 hours). Unmodified G-CSF mRNA, modified G-CSF mRNA with or without lipoplex tolerated 12 freeze/thaw cycles.

**[001031]** mRNA samples stored in stabilizer at 40ºC showed similar stability to the mRNA samples stored in water at -80ºC after 35 days whereas the mRNA stored in water at 40ºC showed heavy degradation after 18 days.

**Example 82. Cell viability in BJ Fibroblasts**

**[001032]** Human primary foreskin fibroblasts (BJ fibroblasts) were obtained from American Type Culture Collection (ATCC) (catalog #CRL-2522) and grown in Eagle's Minimum Essential Medium (ATCC, cat# 30-2003) supplemented with 10% fetal bovine serum at 37ºC, under 5% $CO_2$. BJ fibroblasts were seeded on a 24-well plate at a density of 130,000 cells per well in 0.5 ml of

culture medium. 250 ng of modified G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) fully modified with 5-methylcytosine and pseudouridine (Gen1) or fully modified with 5-methylcytosine and N1-methylpseudouridine (Gen2) was transfected using Lipofectamine 2000 (Invitrogen, cat# 11668-019), following manufacturer's protocol. Control samples of Lipofectamine 2000 (LF2000) and unmodified G-CSF mRNA were also transfected. The modified mRNA or control samples were transfected daily for 4 days. The viability of the cells after transfection was evaluated 6 hours and 24 hours after the first transfection (T1, 6 hours or T1, 24 hours), and 24 hours after the second (T2, 24 hours) and fourth transfection (T4, 24 hours).

**[001033]** To determine cell viability, the culture medium was completely removed and the cells were washed once with 600ul of sterile PBS without Ca2+/Mg2+ (Gibco/Life Technologies, Manassas, VA) in order to rinse-off loosely attached cells. PBS was removed and discarded. The cleaned fibroblasts in each well were treated with 220ul of a diluted CELL TITER GLO® (Promega, catalog #G7570) stock solution (the CELL TITER GLO® stock solution was further diluted 1:1 with an equal amount of sterile PBS). A sterile pipet tip was used to scratch the cells off the plate and accelerate the lysis process.

**[001034]** For two time intervals, T1, 24 hours and T2, 24 hours, an alternative protocol was applied. Cells were washed with PBS, as described above, and subsequently trypsinized with Trypsin/EDTA solution (Gibco/Life Technologies, Manassas, VA). Cells were detached and collected in 500ul of medium containing trypsin inhibitor. Cells were harvested by centrifugation at 1200 rcf for 5 minutes. The cell pellet was resuspended in 500ul PBS. This cell suspension was kept on ice, and 100ul of this was combined with 100ul of undiluted Cell Titer Glo solution.

**[001035]** All of the CELL TITER GLO® lysates were then incubated at room temperature for 20 minutes. 20 ul of the lysates were transferred to a white opaque polystyrene 96-well plate (Corning, Manassas, VA) and combined with 100 ul diluted CELL TITER GLO® solution. The plate reader used was from BioTek Synergy H1 (BioTek, Winooski, VT) and the absolute values were normalized to signal of the untreated BJ Fibroblasts to 100% cell vitality. The percent viability for the BJ fibroblasts are shown in Table 33.

**[001036]** Importantly, all of these experiments are conducted in the absence of any interferon or other cytokine inhibitors and thus represent an accurate measure of the cytotoxicity of the different mRNA.

**[001037]** These results demonstrate that repeated transfection of BJ fibroblasts with unmodified mRNA results in loss of cell viability that is apparent as early as 24 hrs after the first transfection (T1, 24 hours) and continues to be apparent and more pronounced at subsequent time points.

**[001038]** There is also a loss of viability with repeated transfection of 5methylcytidine and pseudouridine modified mRNA that is apparent 24 hours after the fourth daily transfection (T4, 24 hours). No loss of cell viability over the course of this experiment is seen using 5methylcytidine and N1-methylpseudouridine modified mRNA. These results demonstrate that 5methylcytidine and N1-methylpseudouridine containing mRNA have improved cell viability when analyzed under repeated transfection. The ability to repeatedly administer modified mRNA is important in most therapeutic applications, and as such the ability to do so without cytotoxicity is also important. While not wishing to be bound by theory, it is believed that response genes following a single transfection may lead to a decrease in protein production, cytokine induction, and eventually loss of cell viability. These results are consistent with N1-methylpseudouridine-containing mRNA showing an improved profile in this respect relative to both unmodified mRNA and pseudouridine-modified mRNA.

**Table 33. Percent Viability**

| | T1, 6 hours | T1, 24 hours | T2, 24 hours | T4, 24 hours |
|---|---|---|---|---|
| Gen 1 G-CSF | 81 | 108 | 91 | 65 |
| Gen 2 G-CSF | 99 | 102 | 128 | 87 |
| Unmodified G-CSF | 101 | 72 | 74 | 42 |
| LF2000 | 99 | 80 | 114 | 106 |
| Untreated | 100 | 100 | 100 | 100 |

**Example 83. Innate Immune Response in BJ Fibroblasts**

**[001039]** Human primary foreskin fibroblasts (BJ fibroblasts) are obtained from American Type Culture Collection (ATCC) (catalog #CRL-2522) and grown in Eagle's Minimum Essential Medium (ATCC, cat# 30-2003) supplemented with 10% fetal bovine serum at 37ºC, under 5% $CO_2$. BJ fibroblasts are seeded on a 24-well plate at a density of 130,000 cells per well in 0.5 ml of culture medium. 250 ng of modified G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) fully modified with 5-methylcytosine and pseudouridine (Gen1) or fully modified with 5-methylcytosine and N1-methylpseudouridine (Gen2) is transfected using Lipofectamine 2000 (Invitrogen, cat# 11668-019), following manufacturer's protocol. Control samples of Lipofectamine 2000 and unmodified G-CSF

mRNA (natural G-CSF) are also transfected. The cells are transfected for five consecutive days. The transfection complexes are removed four hours after each round of transfection.

**[001040]** The culture supernatant is assayed for secreted GCSF (R&D Systems, catalog #DCS50), tumor necrosis factor-alpha (TNF-alpha ) and interferon alpha (IFN-alpha) by ELISA every day after transfection following manufacturer's protocols. The cells are analyzed for viability using CELL TITER GLO® (Promega, catalog #G7570) 6 hrs and 18 hrs after the first round of transfection and every alternate day following that. At the same time from the harvested cells, total RNA is isolated and treated with DNASE® using the RNAEASY micro kit (catalog #74004) following the manufacturer's protocol. 100 ng of total RNA is used for cDNA synthesis using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, cat #4368814) following the manufacturer's protocol. The cDNA is then analyzed for the expression of innate immune response genes by quantitative real time PCR using SybrGreen in a Biorad CFX 384 instrument following the manufacturer's protocol.

**Example 84. *In vitro* Transcription with wild-type T7 polymerase**

**[001041]** Luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) and G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) were fully modified with different chemistries and chemistry combinations listed in Tables 34-37 using wild-type T7 polymerase as previously described.

**[001042]** The yield of the translation reactions was determined by spectrophometric measurement (OD260) and the yield for Luciferase is shown in Table 34 and G-CSF is shown in Table 36.

**[001043]** The luciferase and G-CSF modified mRNA were also subjected to an enzymatic capping reaction and each modified mRNA capping reaction was evaluated for yield by spectrophometic measurement (OD260) and correct size assessed using bioanalyzer. The yield from the capping reaction for luciferase is shown in Table 35 and G-CSF is shown in Table 37.

**Table 34. *In vitro* transcription chemistry for Luciferase**

| Chemical Modification | Yield (mg) |
|---|---|
| N6-methyladenosine | 0.99 |
| 5-methylcytidine | 1.29 |
| N4-acetylcytidine | 1.0 |
| 5-formylcytidine | 0.55 |
| Pseudouridine | 2.0 |

| | |
|---|---|
| N1-methylpseudouridine | 1.43 |
| 2-thiouridine | 1.56 |
| 5-methoxyuridine | 2.35 |
| 5-methyluridine | 1.01 |
| α-Thio-cytidine | 0.83 |
| 5-Br-uridine (5Bru) | 1.96 |
| 5 (2 carbomethoxyvinyl) uridine | 0.89 |
| 5 (3-1E propenyl Amino) uridine | 2.01 |
| N4-acetylcytidine/pseudouridine | 1.34 |
| N4-acetylcytidine/N1-methylpseudouridine | 1.26 |
| 5-methylcytidine/5-methoxyuridine | 1.38 |
| 5-methylcytidine/5-bromouridine | 0.12 |
| 5-methylcytidine/5-methyluridine | 2.97 |
| 5-methylcytidine/ half of the uridines are modified with 2-thiouridine | 1.59 |
| 5-methylcytidine/2-thiouridine | 0.90 |
| 5-methylcytidine/pseudouridine | 1.83 |
| 5-methylcytidine/N1 methyl pseudouridine | 1.33 |

**Table 35. Capping chemistry and yield for Luciferase modified mRNA**

| **Chemical Modification** | **Yield (mg)** |
|---|---|
| 5-methylcytidine | 1.02 |
| N4-acetylcytidine | 0.93 |
| 5-formylcytidine | 0.55 |
| Pseudouridine | 2.07 |
| N1-methylpseudouridine | 1.27 |
| 2-thiouridine | 1.44 |
| 5-methoxyuridine | 2 |
| 5-methyluridine | 0.8 |
| α-Thio-cytidine | 0.74 |
| 5-Br-uridine (5Bru) | 1.29 |
| 5 (2 carbomethoxyvinyl) uridine | 0.54 |
| 5 (3-1E propenyl Amino) uridine | 1.39 |
| N4-acetylcytidine/pseudouridine | 0.99 |
| N4-acetylcytidine/N1-methylpseudouridine | 1.08 |
| 5-methylcytidine/5-methoxyuridine | 1.13 |
| 5-methylcytidine/5-methyluridine | 1.08 |
| 5-methylcytidine/ half of the uridines are modified with 2-thiouridine | 1.2 |
| 5-methylcytidine/2-thiouridine | 1.27 |
| 5-methylcytidine/pseudouridine | 1.19 |
| 5-methylcytidine/N1 methyl pseudouridine | 1.04 |

**Table 36. *In vitro* transcription chemistry and yield for G-CSF modified mRNA**

| **Chemical Modification** | **Yield** |
|---|---|

| | (mg) |
|---|---|
| N6-methyladenosine | 1.57 |
| 5-methylcytidine | 2.05 |
| N4-acetylcytidine | 3.13 |
| 5-formylcytidine | 1.41 |
| Pseudouridine | 4.1 |
| N1-methylpseudouridine | 3.24 |
| 2-thiouridine | 3.46 |
| 5-methoxyuridine | 2.57 |
| 5-methyluridine | 4.27 |
| 4-thiouridine | 1.45 |
| 2'-F-uridine | 0.96 |
| α-Thio-cytidine | 2.29 |
| 2'-F-guanosine | 0.6 |
| N-1-methyladenosine | 0.63 |
| 5-Br-uridine (5Bru) | 1.08 |
| 5 (2 carbomethoxyvinyl) uridine | 1.8 |
| 5 (3-1E propenyl Amino) uridine | 2.09 |
| N4-acetylcytidine/pseudouridine | 1.72 |
| N4-acetylcytidine/N1-methylpseudouridine | 1.37 |
| 5-methylcytidine/5-methoxyuridine | 1.85 |
| 5-methylcytidine/5-methyluridine | 1.56 |
| 5-methylcytidine/ half of the uridines are modified with 2-thiouridine | 1.84 |
| 5-methylcytidine/2-thiouridine | 2.53 |
| 5-methylcytidine/pseudouridine | 0.63 |
| N4-acetylcytidine/2-thiouridine | 1.3 |
| N4-acetylcytidine/5-bromouridine | 1.37 |
| 5-methylcytidine/N1 methyl pseudouridine | 1.25 |
| N4-acetylcytidine/pseudouridine | 2.24 |

**Table 37. Capping chemistry and yield for G-CSF modified mRNA**

| Chemical Modification | Yield (mg) |
|---|---|
| N6-methyladenosine | 1.04 |
| 5-methylcytidine | 1.08 |
| N4-acetylcytidine | 2.73 |
| 5-formylcytidine | 0.95 |
| Pseudouridine | 3.88 |
| N1-methylpseudouridine | 2.58 |
| 2-thiouridine | 2.57 |
| 5-methoxyuridine | 2.05 |
| 5-methyluridine | 3.56 |
| 4-thiouridine | 0.91 |
| 2'-F-uridine | 0.54 |
| α-Thio-cytidine | 1.79 |
| 2'-F-guanosine | 0.14 |
| 5-Br-uridine (5Bru) | 0.79 |
| 5 (2 carbomethoxyvinyl) uridine | 1.28 |

| | |
|---|---|
| 5 (3-1E propenyl Amino) uridine | 1.78 |
| N4-acetylcytidine/pseudouridine | 0.29 |
| N4-acetylcytidine/N1-methylpseudouridine | 0.33 |
| 5-methylcytidine/5-methoxyuridine | 0.91 |
| 5-methylcytidine/5-methyluridine | 0.61 |
| 5-methylcytidine/ half of the uridines are modified with 2-thiouridine | 1.24 |
| 5-methylcytidine/pseudouridine | 1.08 |
| N4-acetylcytidine/2-thiouridine | 1.34 |
| N4-acetylcytidine/5-bromouridine | 1.22 |
| 5-methylcytidine/N1 methyl pseudouridine | 1.56 |

**Example 85. *In vitro* Transcription with mutant T7 polymerase**

**[001044]** Luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) and G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) were fully modified with different chemistries and chemistry combinations listed in Tables 38-41 using a mutant T7 polymerase (Durascribe® T7 Transcription kit (Cat. No. DS010925) (Epicentre®, Madison, WI).

**[001045]** The yield of the translation reactions was determined by spectrophometric measurement (OD260) and the yield for Luciferase is shown in Table 38 and G-CSF is shown in Table 40.

**[001046]** The luciferase and G-CSF modified mRNA were also subjected to an enzymatic capping reaction and each modified mRNA capping reaction was evaluated for yield by spectrophometic measurement (OD260) and correct size assessed using bioanalyzer. The yield from the capping reaction for luciferase is shown in Table 39 and G-CSF is shown in Table 41.

**Table 38. *In vitro* transcription chemistry and yield for Luciferase modified mRNA**

| Chemical Modification | Yield (ug) |
|---|---|
| 2'Fluorocytosine | 71.4 |
| 2'Fluorouridine | 57.5 |
| 5-methylcytosine/pseudouridine, test A | 26.4 |
| 5-methylcytosine/N1-methylpseudouridine, test A | 73.3 |
| N1-acetylcytidine/2-fluorouridine | 202.2 |
| 5-methylcytidine/2-fluorouridine | 131.9 |
| 2-fluorocytosine/pseudouridine | 119.3 |
| 2-fluorocytosine/N1-methylpseudouridine | 107.0 |
| 2-fluorocytosine/2-thiouridine | 34.7 |
| 2-fluorocytosine/5-bromouridine | 81.0 |
| 2-fluorocytosine/2-fluorouridine | 80.4 |
| 2-fluoroguanine/5-methylcytosine | 61.2 |
| 2-fluoroguanine/5-methylcytosine/pseudouridine | 65.0 |
| 2-fluoroguanine/5-methylcytidine/N1-methylpseudouridine | 41.2 |

| | |
|---|---|
| 2-fluoroguanine/pseudouridine | 79.1 |
| 2-fluoroguanine/N1-methylpseudouridine | 74.6 |
| 5-methylcytidine/pseudouridine, test B | 91.8 |
| 5-methylcytidine/N1-methylpseudouridine, test B | 72.4 |
| 2'fluoroadenosine | 190.98 |

**Table 39. Capping chemistry and yield for Luciferase modified mRNA**

| Chemical Modification | Yield (ug) |
|---|---|
| 2'Fluorocytosine | 19.2 |
| 2'Fluorouridine | 16.7 |
| 5-methylcytosine/pseudouridine, test A | 7.0 |
| 5-methylcytosine/N1-methylpseudouridine, test A | 21.5 |
| N1-acetylcytidine/2-fluorouridine | 47.5 |
| 5-methylcytidine/2-fluorouridine | 53.2 |
| 2-fluorocytosine/pseudouridine | 58.4 |
| 2-fluorocytosine/N1-methylpseudouridine | 26.2 |
| 2-fluorocytosine/2-thiouridine | 12.9 |
| 2-fluorocytosine/5-bromouridine | 26.5 |
| 2-fluorocytosine/2-fluorouridine | 35.7 |
| 2-fluoroguanine/5-methylcytosine | 24.7 |
| 2-fluoroguanine/5-methylcytosine/pseudouridine | 32.3 |
| 2-fluoroguanine/5-methylcytidine/N1-methylpseudouridine | 31.3 |
| 2-fluoroguanine/pseudouridine | 20.9 |
| 2-fluoroguanine/N1-methylpseudouridine | 29.8 |
| 5-methylcytidine/pseudouridine, test B | 58.2 |
| 5-methylcytidine/N1-methylpseudouridine, test B | 44.4 |

**Table 40. *In vitro* transcription chemistry and yield for G-CSF modified mRNA**

| Chemical Modification | Yield (ug) |
|---|---|
| 2'Fluorocytosine | 56.5 |
| 2'Fluorouridine | 79.4 |
| 5-methylcytosine/pseudouridine, test A | 21.2 |
| 5-methylcytosine/N1-methylpseudouridine, test A | 77.1 |
| N1-acetylcytidine/2-fluorouridine | 168.6 |
| 5-methylcytidine/2-fluorouridine | 134.7 |
| 2-fluorocytosine/pseudouridine | 97.8 |
| 2-fluorocytosine/N1-methylpseudouridine | 103.1 |
| 2-fluorocytosine/2-thiouridine | 58.8 |
| 2-fluorocytosine/5-bromouridine | 88.8 |
| 2-fluorocytosine/2-fluorouridine | 93.9 |
| 2-fluoroguanine/5-methylcytosine | 97.3 |
| 2-fluoroguanine/5-methylcytosine/pseudouridine | 96.0 |
| 2-fluoroguanine/5-methylcytidine/N1-methylpseudouridine | 82.0 |
| 2-fluoroguanine/pseudouridine | 68.0 |

| | |
|---|---|
| 2-fluoroguanine/N1-methylpseudouridine | 59.3 |
| 5-methylcytidine/pseudouridine, test B | 58,7 |
| 5-methylcytidine/N1-methylpseudouridine, test B | 78.0 |

**Table 41. Capping chemistry and yield for G-CSF modified mRNA**

| Chemical Modification | Yield (ug) |
|---|---|
| 2'Fluorocytosine | 16.9 |
| 2'Fluorouridine | 17.0 |
| 5-methylcytosine/pseudouridine, test A | 10.6 |
| 5-methylcytosine/N1-methylpseudouridine, test A | 22.7 |
| N1-acetylcytidine/2-fluorouridine | 19.9 |
| 5-methylcytidine/2-fluorouridine | 21.3 |
| 2-fluorocytosine/pseudouridine | 65.2 |
| 2-fluorocytosine/N1-methylpseudouridine | 58.9 |
| 2-fluorocytosine/2-thiouridine | 41.2 |
| 2-fluorocytosine/5-bromouridine | 35.8 |
| 2-fluorocytosine/2-fluorouridine | 36.7 |
| 2-fluoroguanine/5-methylcytosine | 36.6 |
| 2-fluoroguanine/5-methylcytosine/pseudouridine | 37.3 |
| 2-fluoroguanine/5-methylcytidine/N1-methylpseudouridine | 30.7 |
| 2-fluoroguanine/pseudouridine | 29.0 |
| 2-fluoroguanine/N1-methylpseudouridine | 22.7 |
| 5-methylcytidine/pseudouridine, test B | 60.4 |
| 5-methylcytidine/N1-methylpseudouridine, test B | 33.0 |

**Example 86. 2'O-methyl and 2'Fluoro compounds**

**[001047]** Luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) were produced as fully modified versions with the chemistries in Table 42 and transcribed using mutant T7 polymerase (Durascribe® T7 Transcription kit (Cat. No. DS010925) (Epicentre®, Madison, WI). 2' fluoro-containing mRNA were made using Durascribe T7, however, 2'Omethyl-containing mRNA could not be transcribed using Durascribe T7.

**[001048]** Incorporation of 2'Omethyl modified mRNA might possibly be accomplished using other mutant T7 polymerases (Nat Biotechnol. (2004) 22:1155-1160; Nucleic Acids Res. (2002) 30:e138). Alternatively, 2'OMe modifications could be introduced post-transcriptionally using enzymatic means.

**[001049]** Introduction of modifications on the 2' group of the sugar has many potential advantages. 2'OMe substitutions, like 2' fluoro substitutions are known to protect against nucleases and also

have been shown to abolish innate immune recognition when incorporated into other nucleic acids such as siRNA and anti-sense (incorporated in its entirety, Crooke, ed. Antisense Drug Technology, 2nd edition; Boca Raton: CRC press).

**[001050]** The 2'Fluoro-modified mRNA were then transfected into HeLa cells to assess protein production in a cell context and the same mRNA were also assessed in a cell-free rabbit reticulocyte system. A control of unmodified luciferase (natural luciferase) was used for both transcription experiments, a control of untreated and mock transfected (Lipofectamine 2000 alone) were also analyzed for the HeLa transfection and a control of no RNA was analyzed for the rabbit reticulysates.

**[001051]** For the HeLa transfection experiments, the day before transfection, 20,000 HeLa cells (ATCC no. CCL-2; Manassas, VA) were harvested by treatment with Trypsin-EDTA solution (LifeTechnologies, Grand Island, NY) and seeded in a total volume of 100ul EMEM medium (supplemented with 10%FCS and 1x Glutamax) per well in a 96-well cell culture plate (Corning, Manassas, VA). The cells were grown at 37oG in 5% $CO_2$ atmosphere overnight. Next day, 83 ng of the 2'fluoro-containing luciferase modified RNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) with the chemical modification described in Table 42, were diluted in 10ul final volume of OPTI-MEM (LifeTechnologies, Grand Island, NY). Lipofectamine 2000 (LifeTechnologies, Grand Island, NY) was used as transfection reagent and 0.2 ul were diluted in 10 ul final volume of OPTI-MEM. After 5 minutes of incubation at room temperature, both solutions were combined and incubated an additional 15 minute at room temperature. Then the 20ul combined solution was added to the 100ul cell culture medium containing the HeLa cells and incubated at room temperature. After 18 to 22 hours of incubation cells expressing luciferase were lysed with 100 ul of Passive Lysis Buffer (Promega, Madison, WI) according to manufacturer instructions. Aliquots of the lysates were transferred to white opaque polystyrene 96-well plates (Corning, Manassas, VA) and combined with 100 ul complete luciferase assay solution (Promega, Madison, WI). The lysate volumes were adjusted or diluted until no more than 2 mio relative light units (RLU) per well were detected for the strongest signal producing samples and the RLUs for each chemistry tested are shown in Table 42. The plate reader was a BioTek Synergy H1 (BioTek, Winooski, VT).The background signal of the plates without reagent was about 200 relative light units per well.

**[001052]** For the rabbit reticulocyte lysate assay, 2'-fluoro-containing luciferase mRNA were diluted in sterile nuclease-free water to a final amount of 250 ng in 10 ul and added to 40 ul of freshly prepared Rabbit Reticulocyte Lysate and the *in vitro* translation reaction was done in a standard 1.5 mL polypropylene reaction tube (Thermo Fisher Scientific, Waltham, MA) at 30ºC in a dry heating block. The translation assay was done with the Rabbit Reticulocyte Lysate (nuclease-treated) kit (Promega, Madison, WI) according to the manufacturer's instructions. The reaction buffer was supplemented with a one-to-one blend of provided amino acid stock solutions devoid of either Leucine or Methionine resulting in a reaction mix containing sufficient amounts of both amino acids to allow effective *in vitro* translation. After 60 minutes of incubation, the reaction was stopped by placing the reaction tubes on ice.

**[001053]** Aliquots of the *in vitro* translation reaction containing luciferase modified RNA were transferred to white opaque polystyrene 96-well plates (Corning, Manassas, VA) and combined with 100ul complete luciferase assay solution (Promega, Madison, WI). The volumes of the *in vitro* translation reactions were adjusted or diluted until no more than 2 mio relative light units (RLUs) per well were detected for the strongest signal producing samples and the RLUs for each chemistry tested are shown in Table 43. The plate reader was a BioTek Synergy H1 (BioTek, Winooski, VT). The background signal of the plates without reagent was about 160 relative light units per well.

**[001054]** As can be seen in Table 42 and 43, multiple 2'Fluoro-containing compounds are active in vitro and produce luciferase protein.

**Table 42. HeLa Cells**

| Chemical Modification | Concentration (ug/ml) | Volume (ul) | Yield (ug) | RLU |
|---|---|---|---|---|
| 2'Fluoroadenosine | 381.96 | 500 | 190.98 | 388.5 |
| 2'Fluorocytosine | 654.56 | 500 | 327.28 | 2420 |
| 2'Fluoroguanine | 541,795 | 500 | 270.90 | 11,705.5 |
| 2'Flurorouridine | 944.005 | 500 | 472.00 | 6767.5 |
| Natural luciferase | N/A | N/A | N/A | 133,853.5 |
| Mock | N/A | N/A | N/A | 340 |
| Untreated | N/A | N/A | N/A | 238 |

**Table 43. Rabbit Reticulysates**

| Chemical Modification | RLU |
|---|---|
| 2'Fluoroadenosine | 162 |
| 2'Fluorocytosine | 208 |
| 2'Fluoroguanine | 371,509 |
| 2'Flurorouridine | 258 |
| Natural luciferase | 2,159,968 |
| No RNA | 156 |

**Example 87. Luciferase in HeLa Cells using a combination of modifications**

**[001055]** To evaluate using of 2'fluoro-modified mRNA in combination with other modification a series of mRNA were transcribed using either wild-type T7 polymerase (non-fluoro-containing compounds) or using mutant T7 polymerases (fluyoro-containing compounds) as described in Example 86. All modified mRNA were tested by in vitro transfection in HeLa cells.

**[001056]** The day before transfection, 20,000 HeLa cells (ATCC no. CCL-2; Manassas, VA) were harvested by treatment with Trypsin-EDTA solution (LifeTechnologies, Grand Island, NY) and seeded in a total volume of 100ul EMEM medium (supplemented with 10%FCS and 1x Glutamax) per well in a 96-well cell culture plate (Corning, Manassas, VA). The cells were grown at 37oG in 5% $CO_2$ atmosphere overnight. Next day, 83 ng of Luciferase modified RNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) with the chemical modification described in Table 44, were diluted in 10ul final volume of OPTI-MEM (LifeTechnologies, Grand Island, NY). Lipofectamine 2000 (LifeTechnologies, Grand Island, NY) was used as transfection reagent and 0.2 ul were diluted in 10 ul final volume of OPTI-MEM. After 5 minutes of incubation at room temperature, both solutions were combined and incubated an additional 15 minute at room temperature. Then the 20ul combined solution was added to the 100ul cell culture medium containing the HeLa cells and incubated at room temperature.

**[001057]** After 18 to 22 hours of incubation cells expressing luciferase were lysed with 100 ul of Passive Lysis Buffer (Promega, Madison, WI) according to manufacturer instructions. Aliquots of the lysates were transferred to white opaque polystyrene 96-well plates (Corning, Manassas, VA) and combined with 100 ul complete luciferase assay solution (Promega, Madison, WI). The lysate volumes were adjusted or diluted until no more than 2 mio relative light units (RLU) per well were detected for the strongest signal producing samples and the RLUs for each chemistry tested are shown in Table 44. The plate reader was a BioTek Synergy H1 (BioTek, Winooski, VT).The background signal of the plates without reagent was about 200 relative light units per well.

**[001058]** As evidenced in Table 44, most combinations of modifications resulted in mRNA which produced functional luciferase protein, including all the non-flouro containing compounds and many of the combinations containing 2'fluro modifications.

**Table 44. Luciferase**

| Chemical Modification | RLU |
|---|---|

| | |
|---|---|
| N4-acetylcytidine/pseudouridine | 113,796 |
| N4-acetylcytidine/N1-methylpseudouridine | 316,326 |
| 5-methylcytidine/5-methoxyuridine | 24,948 |
| 5-methylcytidine/5-methyluridine | 43,675 |
| 5-methylcytidine/half of the uridines modified with 50% 2-thiouridine | 41,601 |
| 5-methylcytidine/2-thiouridine | 1,102 |
| 5-methylcytidine/pseudouridine | 51,035 |
| 5-methylcytidine/N1 methyl pseudouridine | 152,151 |
| N4-acetylcytidine/2'Fluorouridine triphosphate | 288 |
| 5-methylcytidine/2'Fluorouridine triphosphate | 269 |
| 2'Fluorocytosine triphosphate /pseudouridine | 260 |
| 2'Fluorocytosine triphosphate /N1-methylpseudouridine | 412 |
| 2'Fluorocytosine triphosphate/2-thiouridine | 427 |
| 2'Fluorocytosine triphosphate/5-bromouridine | 253 |
| 2'Fluorocytosine triphosphate /2'Fluorouridine triphosphate | 184 |
| 2'Fluoroguanine triphosphate/5-methylcytidine | 321 |
| 2'Fluoroguanine triphosphate/5-methylcytidine/Pseudouridine | 207 |
| 2'Fluoroguanine /5-methylcytidine/N1 methylpsuedouridine | 235 |
| 2'Fluoroguanine/pseudouridine | 218 |
| 2'Fluoroguanine/N1-methylpsuedouridine | 247 |
| 5-methylcytidine/pseudouridine, test A | 13,833 |
| 5-methylcytidine/N-methylpseudouridine, test A | 598 |
| 2'Fluorocytosine triphosphate | 201 |
| 2'F luorouridine triphosphate | 305 |
| 5-methylcytidine/pseudouridine, test B | 115,401 |
| 5-methylcytidine/N-methylpseudouridine, test B | 21,034 |
| Natural luciferase | 30,801 |
| Untreated | 344 |
| Mock | 262 |

**Example 88. G-CSF *In Vitro* Transcription**

**[001059]** To assess the activity of all our different chemical modifications in the context of a second open reading frame, we replicated experiments previously conducted using luciferase mRNA, with human G-CSF mRNA. G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) were fully modified with the chemistries in Tables 45 and 46 using wild-type T7 polymerase (for all non-fluoro-containing compounds) or mutant T7 polymerase (for all fluoro-containing compounds). The mutant T7 polymerase was obtained commercially (Durascribe® T7 Transcription kit (Cat. No. DS010925) (Epicentre®, Madison, WI).

**[001060]** The modified RNA in Tables 45 and 46 were transfected *in vitro* in HeLa cells or added to rabbit reticulysates (250ng of modified mRNA) as indicated. A control of untreated, mock transfected (transfection reagent alone), G-CSF fully modified with 5-methylcytosine and N1-

methylpseudouridine or luciferase control (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) fully modified with 5-methylcytosine and N1-methylpseudouridine were also analyzed. The expression of G-CSF protein was determined by ELISA and the values are shown in Tables 45 and 46. In Table 45, "NT" means not tested.

**[001061]** As shown in Table 45, many, but not all, chemical modifications resulted in human G-CSF protein production. These results from cell-based and cell-free translation systems correlate very nicely with the same modifications generally working or not working in both systems. One notable exception is 5-formylcytidine modified G-CSF mRNA which worked in the cell-free translation system, but not in the HeLa cell-based transfection system. A similar difference between the two assays was also seen with 5-formylcytidine modified luciferase mRNA.

**[001062]** As demonstrated in Table 46, many, but not all, G-CSF mRNA modified chemistries (when used in combination) demonstrated *in vivo* activity. In addition the presence of N1-methylpseudouridine in the modified mRNA (with N4-acetylcytidine or 5 methylcytidine) demonstrated higher expression than when the same combinations where tested using with pseudouridine. Taken together, these data demonstrate that N1-methylpseudouridine containing G-CSF mRNA results in improved protein expression *in vitro*.

**Table 45. G-CSF Expression**

| **Chemical Modification** | **G-CSF protein (pg/ml) HeLa cells** | **G-CSF protein (pg/ml) Rabbit reticulysates cells** |
|---|---|---|
| Pseudouridine | 1,150,909 | 147,875 |
| 5-methyluridine | 347,045 | 147,250 |
| 2-thiouridine | 417,273 | 18,375 |
| N1-methylpseudouridine | NT | 230,000 |
| 4-thiouridine | 107,273 | 52,375 |
| 5-methoxyuridine | 1,715,909 | 201,750 |
| 5-methylcytosine/pseudouridine, Test A | 609,545 | 119,750 |
| 5-methylcytosine/N1-methylpseudouridine , Test A | 1,534,318 | 110,500 |
| 2'-Fluoro-guanosine | 11,818 | 0 |
| 2'-Fluoro-uridine | 60,455 | 0 |
| 5-methylcytosine/pseudouridine, Test B | 358,182 | 57,875 |
| 5-methylcytosine/N1-methylpseudouridine , Test B | 1,568,636 | 76,750 |
| 5-Bromo-uridine | 186,591 | 72,000 |
| 5-(2carbomethoxyvinyl) uridine | 1,364 | 0 |
| 5-[3(1-E-propenylamino) uridine | 27,955 | 32,625 |

| | | |
|---|---|---|
| α-thio-cytidine | 120,455 | 42,625 |
| 5-methylcytosine/pseudouridine, Test C | 882,500 | 49,250 |
| N1-methyl-adenosine | 4,773 | 0 |
| N6-methyl-adenosine | 1,591 | 0 |
| 5-methyl-cytidine | 646,591 | 79,375 |
| N4-acetylcytidine | 39,545 | 8,000 |
| 5-formyl-cytidine | 0 | 24,000 |
| 5-methylcytosine/pseudouridine, Test D | 87,045 | 47,750 |
| 5-methylcytosine/N1-methylpseudouridine , Test D | 1,168,864 | 97,125 |
| Mock | 909 | 682 |
| Untreated | 0 | 0 |
| 5-methylcytosine/N1-methylpseudouridine , Control | 1,106,591 | NT |
| Luciferase control | NT | 0 |

**Table 46. Combination Chemistries in HeLa cells**

| **Chemical Modification** | **G-CSF protein (pg/ml) HeLa cells** |
|---|---|
| N4-acetylcytidine/pseudouridine | 537,273 |
| N4-acetylcytidine/N1-methylpseudouridine | 1,091,818 |
| 5-methylcytidine/5-methoxyuridine | 516,136 |
| 5-methylcytidine/5-bromouridine | 48,864 |
| 5-methylcytidine/5-methyluridine | 207,500 |
| 5-methylcytidine/2-thiouridine | 33,409 |
| N4-acetylcytidine/5-bromouridine | 211,591 |
| N4-acetylcytidine/2-thiouridine | 46,136 |
| 5-methylcytosine/pseudouridine | 301,364 |
| 5-methylcytosine/N1-methylpseudouridine | 1,017,727 |
| N4-acetylcytidine/2'Fluorouridine triphosphate | 62,273 |
| 5-methylcytidine/2'Fluorouridine triphosphate | 49,318 |
| 2'Fluorocytosine triphosphate/pseudouridine | 7,955 |
| 2'Fluorocytosine triphosphate/N1-methylpseudouridine | 1,364 |
| 2'Fluorocytosine triphosphate/2-thiouridine | 0 |
| 2'Fluorocytosine triphosphate/5-bromouridine | 1,818 |
| 2'Fluorocytosine triphosphate/2'Fluorouridine triphosphate | 909 |
| 2'Fluoroguanine triphosphate/5-methylcytidine | 0 |
| 2'Fluoroguanine triphosphate/5-methylcytidine/pseudouridine | 0 |
| 2'Fluoroguanine triphosphate /5-methylcytidine/N1 methylpseudouridine | 1,818 |
| 2'Fluoroguanine triphosphate/pseudouridine | 1,136 |
| 2'Fluoroguanine triphosphate/2'Fluorocytosine triphosphate/N1-methylpseudouridine | 0 |
| 5-methylcytidine/pseudouridine | 617,727 |
| 5-methylcytidine/N1-methylpseudouridine | 747,045 |
| 5-methylcytidine/pseudouridine | 475,455 |
| 5-methylcytidine/N1-methylpseudouridine | 689,091 |

| | |
|---|---|
| 5-methylcytosine/N1-methylpseudouridine, Control 1 | 848,409 |
| 5-methylcytosine/N1-methylpseudouridine, Control 2 | 581,818 |
| Mock | 682 |
| Untreated | 0 |
| Luciferase 2'Fluorocytosine triphosphate | 0 |
| Luciferase 2'Fluorouridine triphosphate | 0 |

**Example 89. Screening of Chemistries**

**[001063]** The tables listed in below (Tables 47-49) summarize much of the *in vitro* and *in vitro* screening data with the different compounds presented in the previous examples. A good correlation exists between cell-based and cell-free translation assays. The same chemistry substitutions generally show good concordance whether tested in the context of luciferase or G-CSF mRNA. Lastly, N1-methylpseudouridine containing mRNA show a very high level of protein expression with little to no detectable cytokine stimulation in vitro and in vivo, and is superior to mRNA containing pseudouridine both *in vitro* and *in vivo*.

**[001064]** Luciferase mRNA (mRNA sequence shown in SEQ ID NO: 3; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) and G-CSF mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) were modified with naturally and non-naturally occurring chemistries described in Tables 47 and 48 or combination chemistries described in Table 48 and tested using methods described herein.

**[001065]** In Tables 47 and 48, "*" refers to *in vitro* transcription reaction using a mutant T7 polymerase (Durascribe® T7 Transcription kit (Cat. No. DS010925) (Epicentre®, Madison, WI); "**" refers to the second result *in vitro* transcription reaction using a mutant T7 polymerase (Durascribe® T7 Transcription kit (Cat. No. DS010925) (Epicentre®, Madison, WI); "***" refers to production seen in cell free translations (rabbit reticulocyte lysates); the protein production of HeLa is judged by "+," "+/-" and "-"; when referring to G-CSF PBMC "++++" means greater than 6,000 pg/ml G-CSF, "+++" means greater than 3,000 pg/ml G-CSF, "++" means greater than 1,500 pg/ml G-CSF, "+" means greater than 300 pg/ml G-CSF, "+/-" means 150-300 pg/ml G-CSF and the background was about 110 pg/ml; when referring to cytokine PBMC "++++" means greater than 1,000 pg/ml interferon-alpha (IFN-alpha), "+++" means greater than 600 pg/ml IFN-alpha, "++" means greater than 300 pg/ml IFN-alpha, "+" means greater than 100 pg/ml IFN-alpha, "-" means less than 100 pg/ml and the background was about 70 pg/ml; and "NT" means not tested. In Table

48, the protein production was evaluated using a mutant T7 polymerase (Durascribe® T7 Transcription kit (Cat. No. DS010925) (Epicentre®, Madison, WI).

**Table 47. Naturally Occurring**

| Common Name (symbol) | IVT (Luc) | IVT (G-CSF) | Protein (Luc; HeLa) | Protein (G-CSF; HeLa) | Protein (G-CSF; PBMC) | Cytokines (G-CSF; PBMC) | In Vivo Protein (Luc) | In Vivo Protein (G-CSF) |
|---|---|---|---|---|---|---|---|---|
| 1-methyladenosine ($m^1A$) | Fail | Pass | NT | - | +/- | ++ | NT | NT |
| $N^6$-methyladenosine ($m^6A$) | Pass | Pass | - | - | +/- | ++++ | NT | NT |
| 2'-O-methyladenosine (Am) | Fail* | Not Done | NT | NT | NT | NT | NT | NT |
| 5-methylcytidine ($m^5C$) | Pass | Pass | + | + | + | ++ | + | NT |
| 2'-O-methylcytidine (Cm) | Fail* | Not Done | NT | NT | NT | NT | NT | NT |
| 2-thiocytidine ($s^2C$) | Fail | Fail | NT | NT | NT | NT | NT | NT |
| $N^4$-acetylcytidine ($ac^4C$) | Pass | Pass | + | + | +/- | +++ | + | NT |
| 5-formylcytidine ($f^5C$) | Pass | Pass | -*** | -*** | - | + | NT | NT |
| 2'-O-methylguanosine (Gm) | Fail* | Not Done | NT | NT | NT | NT | NT | NT |
| inosine (I) | Fail | Fail | NT | NT | NT | NT | NT | NT |
| pseudouridine (Y) | Pass | Pass | + | + | ++ | + | + | NT |
| 5-methyluridine ($m^5U$) | Pass | Pass | + | + | +/- | + | NT | NT |
| 2'-O-methyluridine (Um) | Fail* | Not Done | NT | NT | NT | NT | NT | NT |
| 1-methylpseudouridine ($m^1Y$) | Pass | Pass | + | Not Done | ++++ | - | + | NT |
| 2-thiouridine ($s^2U$) | Pass | Pass | - | + | + | + | NT | NT |
| 4-thiouridine ($s^4U$) | Fail | Pass | | + | +/- | ++ | NT | NT |
| 5-methoxyuridine ($mo^5U$) | Pass | Pass | + | + | ++ | - | + | NT |
| 3-methyluridine ($m^3U$) | Fail | Fail | NT | NT | NT | NT | NT | NT |

**Table 48. Non-Naturally Occurring**

| Common Name | IVT (Luc) | IVT (G- | Protein (Luc; | Protein (G- | Protein (G-CSF; | Cytokines (G- | In Vivo Protein | In Vivo Protein |
|---|---|---|---|---|---|---|---|---|

| | | CSF) | HeLa) | CSF; HeLa) | PBMC) | CSF; PBMC) | (Luc) | (G-CSF) |
|---|---|---|---|---|---|---|---|---|
| 2'-F-ara-guanosine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 2'-F-ara-adenosine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 2'-F-ara-cytidine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 2'-F-ara-uridine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 2'-F-guanosine | Fail/ Pass** | Pass/Fail** | +** | +/- | - | + | + | NT |
| 2'-F-adenosine | Fail/ Pass** | Fail/Fail** | -** | NT | NT | NT | NT | NT |
| 2'-F-cytidine | Fail/ Pass** | Fail/Pass** | +** | NT | NT | NT | + | NT |
| 2'-F-uridine | Fail/ Pass** | Pass/Pass** | +** | + | +/- | + | - | NT |
| 2'-OH-ara-guanosine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 2'-OH-ara-adenosine | Not Done | Not Done | NT | NT | NT | NT | NT | NT |
| 2'-OH-ara-cytidine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 2'-OH-ara-uridine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 5-Br-Uridine | Pass | Pass | + | + | + | + | + | |
| 5-(2-carbomethoxyvinyl) Uridine | Pass | Pass | - | - | +/- | - | | |
| 5-[3-(1-E-Propenylamino) Uridine (aka Chem 5) | Pass | Pass | - | + | + | - | | |
| N6-(19-Amino-pentaoxanonadecyl) A | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 2-Dimethylamino guanosine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| 6-Aza-cytidine | Fail | Fail | NT | NT | NT | NT | NT | NT |
| a-Thio-cytidine | Pass | Pass | + | + | +/- | +++ | NT | NT |
| Pseudo-isocytidine | NT | NT | NT | NT | NT | NT | NT | NT |
| 5-Iodo-uridine | NT | NT | NT | NT | NT | NT | NT | NT |
| a-Thio-uridine | NT | NT | NT | NT | NT | NT | NT | NT |
| 6-Aza-uridine | NT | NT | NT | NT | NT | NT | NT | NT |
| Deoxy-thymidine | NT | NT | NT | NT | NT | NT | NT | NT |
| a-Thio guanosine | NT | NT | NT | NT | NT | NT | NT | NT |
| 8-Oxo-guanosine | NT | NT | NT | NT | NT | NT | NT | NT |
| O6-Methyl-guanosine | NT | NT | NT | NT | NT | NT | NT | NT |
| 7-Deaza-guanosine | NT | NT | NT | NT | NT | NT | NT | NT |
| 6-Chloro-purine | NT | NT | NT | NT | NT | NT | NT | NT |
| a-Thio-adenosine | NT | NT | NT | NT | NT | NT | NT | NT |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 7-Deaza-adenosine | NT | NT | NT | NT | NT | NT | NT | NT |
| 5-iodo-cytidine | NT | NT | NT | NT | NT | NT | NT | NT |

**[001066]** In Table 49, the protein production of HeLa is judged by "+," "+/-" and "-"; when referring to G-CSF PBMC "++++" means greater than 6,000 pg/ml G-CSF, "+++" means greater than 3,000 pg/ml G-CSF, "++" means greater than 1,500 pg/ml G-CSF, "+" means greater than 300 pg/ml G-CSF, "+/-" means 150-300 pg/ml G-CSF and the background was about 110 pg/ml; when referring to cytokine PBMC "++++" means greater than 1,000 pg/ml interferon-alpha (IFN-alpha), "+++" means greater than 600 pg/ml IFN-alpha, "++" means greater than 300 pg/ml IFN-alpha, "+" means greater than 100 pg/ml IFN-alpha, "-" means less than 100 pg/ml and the background was about 70 pg/ml; "WT" refers to the wild type T7 polymerase, "MT" refers to mutant T7 polymerase (Durascribe® T7 Transcription kit (Cat. No. DS010925) (Epicentre®, Madison, WI) and "NT" means not tested.

**Table 49. Combination Chemistry**

| **Cytidine analog** | **Uridine analog** | **Purine** | **IVT Luc** | **IVT (G-CSF)** | **Protein (Luc; HeLa)** | **Protein (G-CSF; HeLa)** | **Protein (G-CSF; PBMC)** | **Cytokines (G-CSF; PBMC)** | **In Vivo Protein (Luc)** |
|---|---|---|---|---|---|---|---|---|---|
| N4-acetylcytidine | pseudouridine | A,G | Pass WT | Pass WT | + | + | NT | NT | + |
| N4-acetylcytidine | N1-methylpseudouridine | A,G | Pass WT | Pass WT | + | + | NT | NT | + |
| 5-methylcytidine | 5-methoxyuridine | A,G | Pass WT | Pass WT | + | + | NT | NT | + |
| 5-methylcytidine | 5-bromouridine | A,G | Pass WT | Pass WT | Not Done | + | NT | NT | |
| 5-methylcytidine | 5-methyluridine | A,G | Pass WT | Pass WT | + | + | NT | NT | + |
| 5-methylcytidine | 50% 2-thiouridine; 50% uridine | A,G | Pass WT | Pass WT | + | NT | NT | NT | + |
| 5-methylcytidine | 100% 2-thiouridine | A,G | Pass WT | Pass WT | - | + | NT | NT | |
| 5-methylcytidine | pseudouridine | A,G | Pass WT | Pass WT | + | + | ++ | + | + |
| 5-methylcytidine | N1-methylpseudouridine | A,G | Pass WT | Pass WT | + | + | ++++ | - | + |
| N4-acetylcytidine | 2-thiouridine | A,G | Not Done | Pass WT | Not Done | + | NT | NT | NT |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| N4-acetylcytidine | 5-bromouridine | A, G | Not Done | Pass WT | Not Done | + | NT | NT | NT |
| N4-acetylcytidine | 2 Fluorouridine triphosphate | A,G | Pass | Pass | - | + | NT | NT | NT |
| 5-methylcytidine | 2 Fluorouridine triphosphate | A,G | Pass | Pass | - | + | NT | NT | NT |
| 2 Fluorocytosine triphosphate | pseudouridine | A,G | Pass | Pass | - | + | NT | NT | NT |
| 2 Fluorocytosine triphosphate | N1-methylpseudouridine | A,G | Pass | Pass | - | +/- | NT | NT | NT |
| 2 Fluorocytosine triphosphate | 2-thiouridine | A,G | Pass | Pass | - | - | NT | NT | NT |
| 2 Fluorocytosine triphosphate | 5-bromouridine | A,G | Pass | Pass | - | +/- | NT | NT | NT |
| 2 Fluorocytosine triphosphate | 2 Fluorouridine triphosphate | A,G | Pass | Pass | - | +/- | NT | NT | NT |
| 5-methylcytidine | uridine | A,2 Fluoro GTP | Pass | Pass | - | - | NT | NT | NT |
| 5-methylcytidine | pseudouridine | A,2 Fluoro GTP | Pass | Pass | - | - | NT | NT | NT |
| 5-methylcytidine | N1-methylpseudouridine | A,2 Fluoro GTP | Pass | Pass | - | +/- | NT | NT | NT |
| 2 Fluorocytosine triphosphate | pseudouridine | A,2 Fluoro GTP | Pass | Pass | - | +/- | NT | NT | NT |
| 2 Fluorocytosine triphosphate | N1-methylpseudouridine | A,2 Fluoro GTP | Pass | Pass | - | - | NT | NT | NT |

**Example 90. 2'Fluoro Chemistries in PBMC**

**[001067]** The ability of G-CSF modified mRNA (mRNA sequence shown in SEQ ID NO: 1; polyA tail of approximately 160 nucleotides not shown in sequence; 5'cap, Cap1) to trigger innate an immune response was determined by measuring interferon-alpha (IFN-alpha) and tumor necrosis factor-alpha (TNF-alpha) production. Use of *in vitro* PBMC cultures is an accepted way to measure the immunostimulatory potential of oligonucleotides (Robbins et al., Oligonucleotides 2009 19:89-

102) and transfection methods are described herein. Shown in Table 50 are the average from 2 or 3 separate PBMC donors of the interferon-alpha (IFN-alpha) and tumor necrosis factor alpha (TNF-alpha) production over time as measured by specific ELISA. Controls of R848, P(I)P(C), LPS and Lipofectamine 2000 (L2000) were also analyzed.

**[001068]** With regards to innate immune recognition, while both modified mRNA chemistries largely prevented IFN-alpha and TNF-alpha production relative to positive controls (R848, P(I)P(C)), 2'fluoro compounds reduce IFN-alpha and TNF-alpha production even lower than other combinations and N4-acetylcytidine combinations raised the cytokine profile.

**Table 50. IFN-alpha and TNF-alpha**

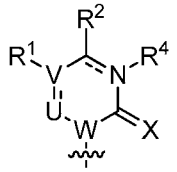
| | IFN-alpha: 3 Donor Average (pg/ml) | TNF-alpha: 2 Donor Average (pg/ml) |
|---|---|---|
| L2000 | 1 | 361 |
| P(I)P(C) | 482 | 544 |
| R848 | 45 | 8,235 |
| LPS | 0 | 6,889 |
| N4-acetylcytidine/pseudouridine | 694 | 528 |
| N4-acetylcytidine/N1-methylpseudouridine | 307 | 283 |
| 5-methylcytidine/5-methoxyuridine | 0 | 411 |
| 5-methylcytidine/5-bromouridine | 0 | 270 |
| 5-methylcytidine/5-methyluridine | 456 | 428 |
| 5-methylcytidine/2-thiouridine | 274 | 277 |
| N4-acetylcytidine/2-thiouridine | 0 | 285 |
| N4-acetylcytidine/5-bromouridine | 44 | 403 |
| 5-methylcytidine/pseudouridine | 73 | 332 |
| 5-methylcytidine/N1-methylpseudouridine | 31 | 280 |
| N4-acetylcytidine/2'fluorouridine triphosphate | 35 | 32 |
| 5-methylcytodine /2'fluorouridine triphosphate | 24 | 0 |
| 2'fluorocytidine triphosphate/N1-methylpseudouridine | 0 | 11 |
| 2'fluorocytidine triphosphate/2-thiouridine | 0 | 0 |
| 2'fluorocytidine/ triphosphate5-bromouridine | 12 | 2 |
| 2'fluorocytidine triphosphate/2'fluorouridine triphosphate | 11 | 0 |
| 2'fluorocytidine triphosphate/5-methylcytidine | 14 | 23 |
| 2'fluorocytidine triphosphate/5-methylcytidine/pseudouridine | 6 | 21 |
| 2'fluorocytidine triphosphate/5-methylcytidine/N1-methylpseudouridine | 3 | 15 |
| 2'fluorocytidine triphosphate/pseudouridine | 0 | 4 |
| 2'fluorocytidine triphosphate/N1-methylpseudouridine | 6 | 20 |
| 5-methylcytidine/pseudouridine | 82 | 18 |
| 5-methylcytidien/N1-methylpseudouridine | 35 | 3 |

**OTHER EMBODIMENTS**

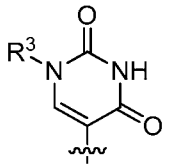
**[001069]** It is to be understood that while the present disclosure has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the present disclosure, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

**WHAT IS CLAIMED IS:**

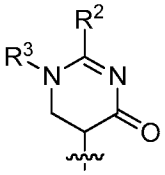
1. An isolated polynucleotide encoding a polypeptide of interest, said isolated polynucleotide comprising:
   (a) a sequence of n number of linked nucleosides or nucleotides comprising at least one modified nucleoside or nucleotide as compared to the chemical structure of an A, G, U or C nucleoside or nucleotide,
   (b) a 5' UTR comprising at least one Kozak sequence,
   (c) a 3' UTR, and
   (d) at least one 5′ cap structure.
2. The isolated polynucleotide of claim 1, further comprising a poly-A tail.
3. The isolated polynucleotide of claim 2 which is purified.
4. The isolated polynucleotide of claim 3, wherein the at least one 5′ cap structure is selected from the group consisting of Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2′fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.
5. The isolated polynucleotide of claim 4, wherein the modification of the nucleoside is located in the nucleoside base and/or sugar portion of the nucleoside.
6. The isolated polynucleotide of claim 5, wherein the modification is located in the nucleotide base and the nucleoside base has the formula:



XII-a,



XII-b, or



XII-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, $C_{1-6}$ alkyl, $C_{1-6}$ alkenyl, $C_{1-6}$ alkynyl, halo, or $-OR^c$, wherein $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl are each optionally substituted with –OH, $-NR^aR^b$, -SH, $-C(O)R^c$, $-C(O)OR^c$, $-NHC(O)R^c$, or $-NHC(O)OR^c$;

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

$R^2$ is H, -$OR^c$, -$SR^c$, -$NR^aR^b$, 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, -$NR^aR^b$, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, $C_{2-20}$ alkynyl, $C_{1-20}$ alkoxy, or $C_{1-20}$ thioalkyl;

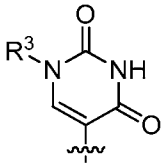
$R^3$ is H or $C_{1-20}$ alkyl;

$R^4$ is H or $C_{1-20}$ alkyl; wherein when $\backsimeq$ denotes a double bond then $R^4$ is absent, or N-$R^4$, taken together, forms a positively charged N substituted with $C_{1-20}$ alkyl;

$R^a$ 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

$R^c$ is H, $C_{1-20}$ alkyl, $C_{2-20}$ alkenyl, phenyl, benzyl, a polyethylene glycol group, or an amino-polyethylene glycol group.

7. The isolated polynucleotide of claim 6, wherein the modification is located in the nucleoside base and the nucleoside base has the formula:



XII-b

wherein:

$R^3$ is $C_{1-20}$ alkyl.

8. The isolated polynucleotide of claim 7, wherein $R^3$ is $C_{1-4}$ alkyl.

9. The isolated polynucleotide of claim 7, wherein $R^3$ is $CH_3$.

10. The isolated polynucleotide of claim 1, wherein the modified nucleoside is not pseudouridine (ψ) or 5-methyl-cytidine ($m^5C$).

11. A pharmaceutical composition comprising the isolated polynucleotide of claim 6 and a pharmaceutically acceptable excipient.

12. The pharmaceutical composition of claim 11, wherein the excipient is selected from a solvent, aqueous solvent, non-aqueous solvent, dispersion media, diluent, dispersion, suspension aid, surface active agent, isotonic agent, thickening or emulsifying agent, preservative, lipid, lipidoids liposome, lipid nanoparticle, core-shell nanoparticles, polymer, lipoplexe peptide, protein, cell, hyaluronidase, and mixtures thereof.

13. A method of increasing the level of a polypeptide of interest in a mammalian subject comprising administering to said subject the isolated polynucleotide of claim 6.
14. The method of claim 13, wherein the polynucleotide is formulated.
15. The method of claim 13, wherein isolated polynucleotide has a Protein:Cytokine Ratio of greater than 100 for either TNF-alpha or IFN-alpha.
16. The method of claim 13, wherein the isolated polynucleotide is administered at a total daily dose of between 1 ug and 150 ug.
17. The method of claim 16, wherein administration is by injection.
18. The method of claim 16, wherein administration is intradermal or subcutaneous or intramuscular.
19. The method of claim 13, wherein levels of the polypeptide of interest in the serum of the mammal are at least 50 pg/mL at least two hours after administration.
20. The method of claim 19, wherein the levels of the polypeptide of interest in the serum of the mammal remain above 50 pg/mL for at least 72 hours after administration.
21. The method of claim 20, wherein the levels of the polypeptide of interest in the serum of the mammal remain above 60 pg/mL for at least 72 hours after administration.
22. The method of claim 13, wherein administration is in two or more equal or unequal split doses.
23. The method of claim 22, wherein the level of the polypeptide produced by the subject by administering split doses of the polynucleotide is greater than the levels produced by administering the same total daily dose of polynucleotide as a single administration.
24. The method of claim 13, wherein the mammalian subject is a human patient in need of an increased level of the polypeptide of interest.
25. The method of claim 24, wherein the increased level of the polypeptide of interest is detectable in a bodily fluid of said patient.
26. The method of claim 53, wherein the bodily fluid is selected from the group consisting of peripheral blood, serum, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, broncheoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions,

mucosal secretion, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyl cavity fluid, and umbilical cord blood.

27. The method of claim 24, wherein administration is according to a dosing regimen which occurs over the course of hours, days, weeks, months, or years.

28. The method of claim 17, wherein injection is achieved by using one or more devices selected from multi-needle injection systems, catheter or lumen systems, and ultrasound, electrical or radiation based systems.

29. The method of claim 22, wherein the amount polynucleotide administered in any dose is substantially equal.

30. The method of claim 22, wherein a first dose, a second dose or any of a plurality of doses are administered at substantially the same time.

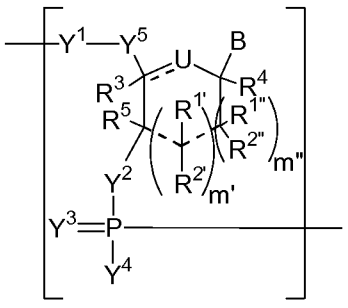
31. The method of claim 13, wherein administration comprises a single unit dose between about 10 mg/kg and about 500 mg/kg.

32. The method of claim 13, wherein administration comprises a single unit dose between about 1.0 mg/kg and about 10 mg/kg.

33. The method of claim 13, wherein administration comprises a single unit dose between about 0.001 mg/kg and about 1.0 mg/kg.

34. The isolated polynucleotide of claim 4, wherein the modification of the nucleoside includes replacing or substituting an atom of a pyrimidine nucleobase with an amine, an SH, a methyl or ethyl, or a chloro or fluoro.

35 The isolated polynucleotide of claim 1, wherein the n number of linked nucleosides comprise the Formula (Ia):



(Ia), or a pharmaceutically acceptable salt or stereoisomer thereof,

wherein

U is O, S, $N(R^U)_{nu}$, or $C(R^U)_{nu}$, wherein nu is an integer from 0 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

 is a single or double bond;

--- is a single bond or absent;

each of $R^{1'}$, $R^{2'}$, $R^{1''}$, $R^{2''}$, $R^1$, $R^2$, $R^3$, $R^4$, and $R^5$, if present, is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent; wherein the combination of $R^3$ with one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, $R^{2''}$, or $R^5$ can join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl; wherein the combination of $R^5$ with one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, or $R^{2''}$ can join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl; and wherein the combination of $R^4$ and one or more of $R^{1'}$, $R^{1''}$, $R^{2'}$, $R^{2''}$, $R^3$, or $R^5$ can join together to form optionally substituted alkylene or optionally substituted heteroalkylene and, taken together with the carbons to which they are attached, provide an optionally substituted heterocyclyl;

each of m' and m" is, independently, an integer from 0 to 3;

each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, $-NR^{N1}-$, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, or absent;

each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino;

each $Y^5$ is, independently, O, S, Se, optionally substituted alkylene, or optionally substituted heteroalkylene;

n is an integer from 1 to 100,000; and

B is a nucleobase.

36. The isolated polynucleotide of claim 35, wherein B is not pseudouridine (ψ) or 5-methyl-cytidine ($m^5C$).

37. The isolated polynucleotide of claim 35 or 36, wherein:

U is O or $C(R^U)_{nu}$, wherein nu is an integer from 1 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

each of $R^1$, $R^{1'}$, $R^{1''}$, $R^2$, $R^{2'}$, and $R^{2''}$, if present, is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl;

each of $R^3$ and $R^4$ is, independently, H, halo, hydroxy, optionally substituted alkyl, or optionally substituted alkoxyalkoxy;

each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, $-NR^{N1}-$, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl;

each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, or optionally substituted amino;

each $Y^5$ is, independently, O or optionally substituted alkylene; and

n is an integer from 10 to 10,000.

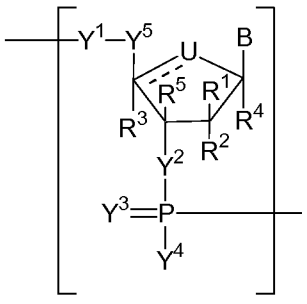
38. The isolated polynucleotide of claim 37, wherein each of $R^1$, $R^{1'}$, and $R^{1''}$, if present, is H.

39. The isolated polynucleotide of claim 38, wherein each of $R^2$, $R^{2'}$, and $R^{2''}$, if present, is, independently, H, halo, hydroxy, optionally substituted alkoxy, or optionally substituted alkoxyalkoxy.

40. The isolated polynucleotide of claim 39, wherein each of $R^2$, $R^{2'}$, and $R^{2''}$, if present, is H.

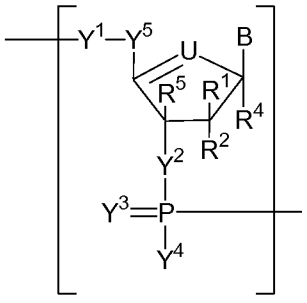
41. The isolated polynucleotide of claim 40, wherein each of $R^1$, $R^{1'}$, and $R^{1''}$, if present, is, independently, H, halo, hydroxy, optionally substituted alkoxy, or optionally substituted alkoxyalkoxy.

42. The isolated polynucleotide of claim 35, wherein the n number of linked nucleotides comprise the Formula (IIa):

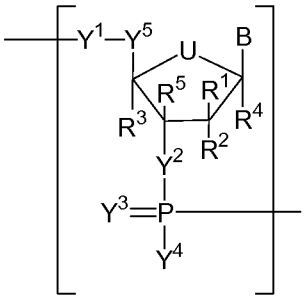


(IIa), or a pharmaceutically acceptable salt or stereoisomer thereof.

43. The isolated polynucleotide of claim 42, wherein the n number of linked nucleosides comprise the Formula (IIb) or (IIc):



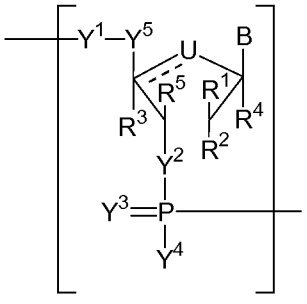
(IIb), or



(IIc), or a pharmaceutically acceptable salt thereof.

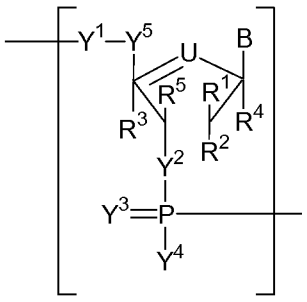
44. The isolated polynucleotide of claim 43, wherein the n number of linked nucleosides comprise the Formula (IIb-1), (IIb-2), or (IIc-1)-(IIc-4).

45. The isolated polynucleotide of claim 35, wherein the n number of linked nucleosides comprise the Formula (IId):

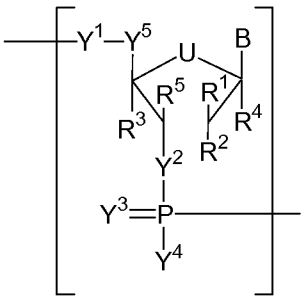


(IId), or a pharmaceutically acceptable salt or stereoisomer thereof.

46. The isolated polynucleotide of claim 44, wherein the n number of linked nucleosides comprise the Formula (IIe) or (IIf):

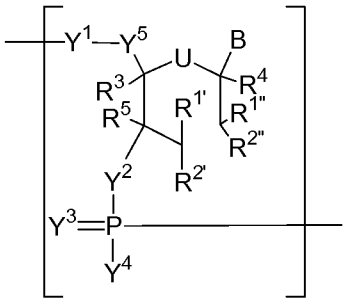


(IIe) or

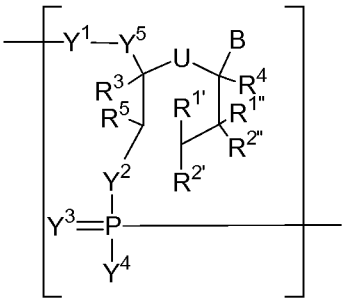


(IIf), or a pharmaceutically acceptable salt thereof.

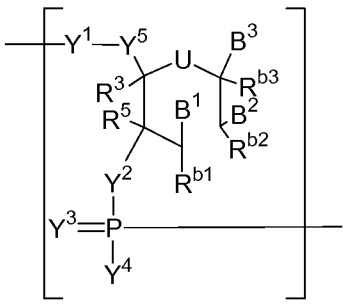
47. The isolated polynucleotide of claim 35, wherein each of said linked nucleotides independently have one of Formulas (IIg)-(IIj):



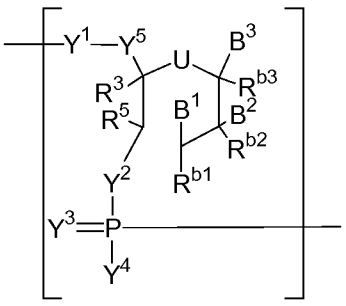
(IIg),



(IIh),

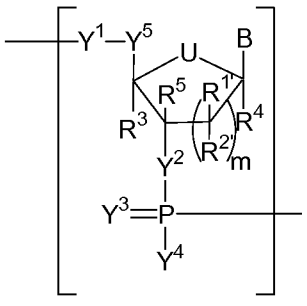


(IIi), or



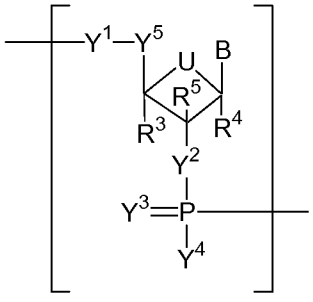
(IIj),or a pharmaceutically acceptable salt or stereoisomer thereof.

48. The isolated polynucleotide of claim 35, wherein the n number of linked nucleosides comprise the Formula (IIk):



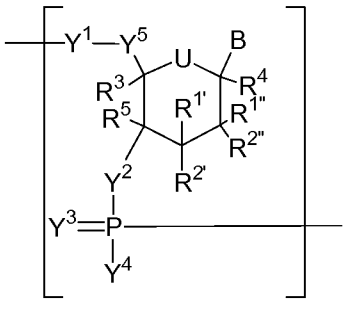
(IIk) , or a pharmaceutically acceptable salt or stereoisomer thereof.

49. The isolated polynucleotide of claim 48, wherein the n number of linked nucleosides comprise the Formula (III):



(III), or a pharmaceutically acceptable salt or stereoisomer thereof.

50. The isolated polynucleotide of claim 48, wherein the n number of linked nucleosides comprise the Formula (IIm):



(IIm), or a pharmaceutically acceptable salt or stereoisomer thereof,

wherein

each of $R^{1'}$, $R^{1''}$, $R^{2'}$, and $R^{2''}$ is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, or absent; and wherein the combination of $R^{2'}$ and $R^3$ or the combination of $R^{2''}$ and $R^3$ can be

taken together to form optionally substituted alkylene or optionally substituted heteroalkylene.

51. The isolated polynucleotide of any one of claims 38-50, wherein

U is O or $C(R^U)_{nu}$, wherein nu is an integer from 1 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

each of $R^1$ and $R^2$ is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl;

each of $R^3$ and $R^4$ is, independently, H or optionally substituted alkyl;

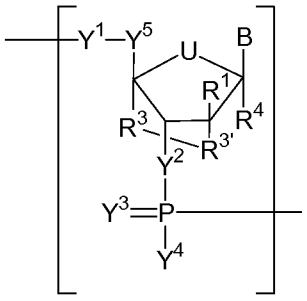
each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, $-NR^{N1}-$, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl;

each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, or optionally substituted amino;

each $Y^5$ is, independently, O or optionally substituted alkylene; and

n is an integer from 10 to 10,000.

52. The isolated polynucleotide of claim 35, wherein the n number of linked nucleosides comprise the Formula (IIn):




(IIn), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein

U is O or $C(R^U)_{nu}$, wherein nu is an integer from 1 to 2 and each $R^U$ is, independently, H, halo, or optionally substituted alkyl;

each of $R^1$ and $R^4$ is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl;

$R^{3'}$ is O, S, or $-NR^{N1}-$, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl;

$R^{3"}$ is optionally substituted alkylene or optionally substituted heteroalkylene;

each of $Y^1$, $Y^2$, and $Y^3$, is, independently, O, S, Se, $-NR^{N1}-$, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein $R^{N1}$ is H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl;

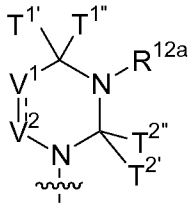
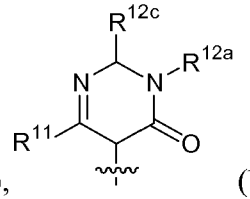
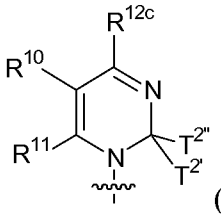
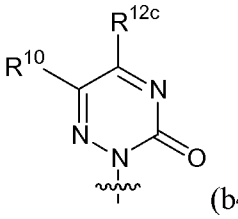
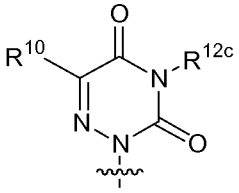
each $Y^4$ is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, or optionally substituted amino;

each $Y^5$ is, independently, O, S, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene; and

n is an integer from 10 to 10,000.

53. The isolated polynucleotide of claim 35, wherein the n number of linked nucleosides comprise Formula (IIn-1) or (II-n2).

54. The isolated polynucleotide of any one of claims 35-53, wherein each B independently has a formula selected from Formula (b1)-(b5):


(b1),

(b2),

(b3),

(b4), or

(b5), or a pharmaceutically acceptable salt or stereoisomer thereof,

wherein

 is a single or double bond;

each of $T^{1'}$, $T^{1''}$, $T^{2'}$, and $T^{2''}$ is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of $T^{1'}$ and $T^{1''}$ or the combination of $T^{2'}$ and $T^{2''}$ join together (e.g., as in $T^2$) to form O (oxo), S (thio), or Se (seleno);

each of $V^1$ and $V^2$ is, independently, O, S, $N(R^{Vb})_{nv}$, or $C(R^{Vb})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vb}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted acylaminoalkyl, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, or optionally substituted alkoxycarbonylalkoxy;

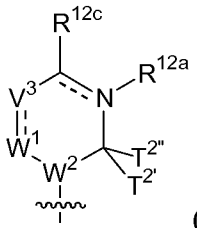
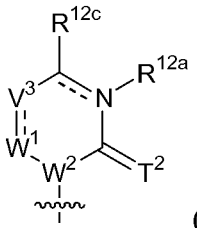
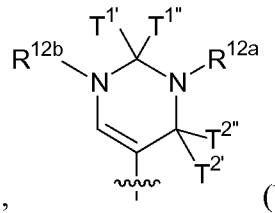
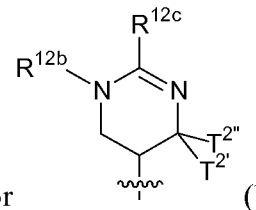
$R^{10}$ is H, halo, optionally substituted amino acid, hydroxy, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aminoalkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl;

$R^{11}$ is H or optionally substituted alkyl;

$R^{12a}$ is H, optionally substituted alkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl, optionally substituted carboxyalkyl, optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, or optionally substituted carbamoylalkyl; and

$R^{12c}$ is H, halo, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted amino, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl.

55. The isolated polynucleotide of any one of claims 35-54, wherein B comprises Formula (b6)-(b9):


(b6),

(b7),

(b8), or

(b9),

or a pharmaceutically acceptable salt or stereoisomer thereof,

wherein

 is a single or double bond;

each of $T^{1'}$, $T^{1''}$, $T^{2'}$, and $T^{2''}$ is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of $T^{1'}$ and $T^{1''}$ join together or the combination of $T^{2'}$ and $T^{2''}$ join together to form O (oxo), S (thio), or Se (seleno), or each $T^1$ and $T^2$ is, independently, O (oxo), S (thio), or Se (seleno);

each of $W^1$ and $W^2$ is, independently, $N(R^{Wa})_{nw}$ or $C(R^{Wa})_{nw}$, wherein nw is an integer from 0 to 2 and each $R^{Wa}$ is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy;

each $V^3$ is, independently, O, S, $N(R^{Va})_{nv}$, or $C(R^{Va})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Va}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted alkoxy, optionally substituted alkenyloxy, or optionally substituted alkynyloxy, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted acylaminoalkyl, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylacyl, optionally substituted

alkoxycarbonylalkoxy, optionally substituted carboxyalkyl, optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, or optionally substituted carbamoylalkyl, and wherein $R^{Va}$ and $R^{12c}$ taken together with the carbon atoms to which they are attached can form optionally substituted cycloalkyl, optionally substituted aryl, or optionally substituted heterocyclyl;

$R^{12a}$ is H, optionally substituted alkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted carboxyalkyl, optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, optionally substituted carbamoylalkyl, or absent;

$R^{12b}$ is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkaryl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted amino acid, optionally substituted alkoxycarbonylacyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkyl, optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, or optionally substituted carbamoylalkyl,

wherein the combination of $R^{12b}$ and $T^{1'}$ or the combination of $R^{12b}$ and $R^{12c}$ can join together to form optionally substituted heterocyclyl; and

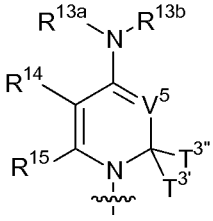
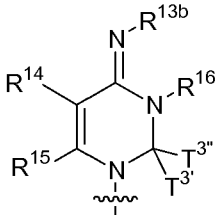
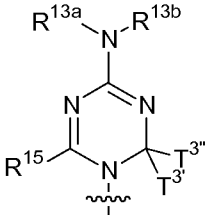
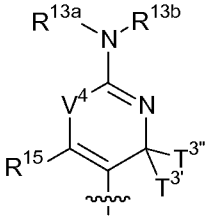
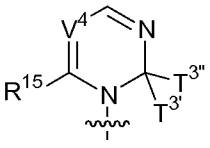
$R^{12c}$ is H, halo, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted amino, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl.

56. The isolated polynucleotide of claim 55, wherein $R^{12a}$, $R^{12b}$, $R^{12c}$, or $R^{Va}$ is substituted with -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10, each of s2 and s3, independently, is an integer from 0 to 10, and R' is H or $C_{1-20}$ alkyl); or -$NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10, each of s2 and

s3, independently, is an integer from 0 to 10, and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl.

57. The isolated polynucleotide of claim 55, wherein B comprises Formula (b28)-(b31).

58. The isolated polynucleotide of any one of claims 35-57, wherein B comprises Formula (b10)-(b14):


(b10),

(b11),

(b12),

(b13), or

(b14), or a pharmaceutically acceptable salt or stereoisomer thereof,

wherein

each of $T^{3'}$ and $T^{3''}$ is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of $T^{3'}$ and $T^{3''}$ join together to form O (oxo), S (thio), or Se (seleno);

each $V^4$ is, independently, O, S, $N(R^{Vc})_{nv}$, or $C(R^{Vc})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vc}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, or optionally substituted alkynyloxy, wherein the combination of $R^{13b}$ and $R^{Vc}$ can be taken together to form optionally substituted heterocyclyl;

each $V^5$ is, independently, $N(R^{Vd})_{nv}$, or $C(R^{Vd})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vd}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, or optionally substituted alkynyloxy;

each of $R^{13a}$ and $R^{13b}$ is, independently, H, optionally substituted acyl, optionally substituted acyloxyalkyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the

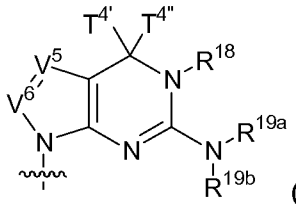
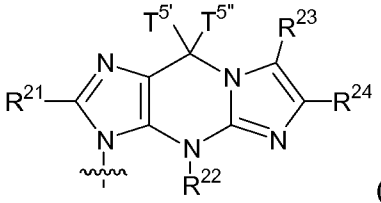
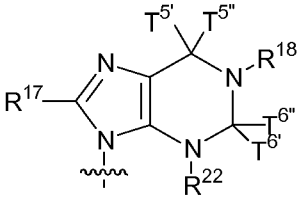
combination of $R^{13b}$ and $R^{14}$ can be taken together to form optionally substituted heterocyclyl;

each $R^{14}$ is, independently, H, halo, hydroxy, thiol, optionally substituted acyl, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted acyloxyalkyl, optionally substituted amino, azido, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl; and

each of $R^{15}$ and $R^{16}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl.

59. The isolated polynucleotide of claim 5858, wherein B comprises Formula (b32)-(b36).

60. The isolated polynucleotide of any one of claims 35-59, wherein B comprises Formula (b15)-(b17):


(b15),

(b16), or


(b17) or a pharmaceutically acceptable salt or stereoisomer thereof,

wherein

each of $T^{4'}$, $T^{4''}$, $T^{5'}$, $T^{5''}$, $T^{6'}$, and $T^{6''}$ is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy, and wherein the combination of $T^{4'}$ and $T^{4''}$ or the combination of $T^{5'}$ and $T^{5''}$ or the combination of $T^{6'}$ and $T^{6''}$ join together form O (oxo), S (thio), or Se (seleno);

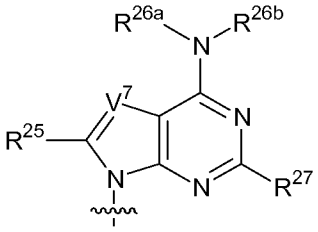
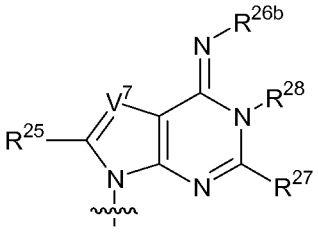
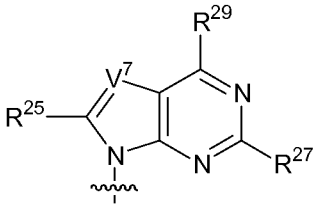
each of $V^5$ and $V^6$ is, independently, O, S, $N(R^{Vd})_{nv}$, or $C(R^{Vd})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Vd}$ is, independently, H, halo, thiol, optionally substituted amino acid, cyano, amidine, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted alkyl, optionally substituted

alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, or optionally substituted amino; and

each of $R^{17}$, $R^{18}$, $R^{19a}$, $R^{19b}$, $R^{21}$, $R^{22}$, $R^{23}$, and $R^{24}$ is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, optionally substituted amino, or optionally substituted amino acid.

61. The isolated polynucleotide of claim 60, wherein B comprises Formula (b37)-(b40).

62. The isolated polynucleotide of any one of claims 35-59, wherein B comprises Formula (b18)-(b20):


(b18),

(b19), or


(b20), or a pharmaceutically acceptable salt or stereoisomer thereof

wherein

each $V^7$ is, independently, O, S, $N(R^{Ve})_{nv}$, or $C(R^{Ve})_{nv}$, wherein nv is an integer from 0 to 2 and each $R^{Ve}$ is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, or optionally substituted alkynyloxy;

each $R^{25}$ is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, or optionally substituted amino;

each of $R^{26a}$ and $R^{26b}$ is, independently, H, optionally substituted acyl, optionally substituted amino acid, optionally substituted carbamoylalkyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkoxy, or polyethylene glycol group, or an amino-polyethylene glycol group;

each $R^{27}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted thioalkoxy, or optionally substituted amino;

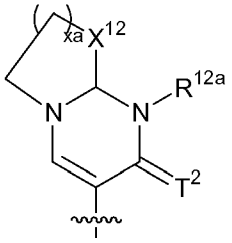
each $R^{28}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl; and

each $R^{29}$ is, independently, H, optionally substituted acyl, optionally substituted amino acid, optionally substituted carbamoylalkyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted alkoxy, or optionally substituted amino.

63. The isolated polynucleotide of claim 62, wherein B comprises Formula (b41)-(b43).

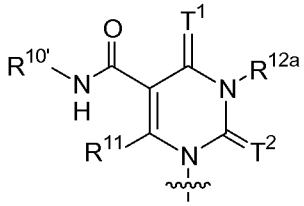
64. The isolated polynucleotide of claim 62, wherein $R^{26a}$, $R^{26b}$, or $R^{29}$ is substituted with -$(CH_2)_{s2}(OCH_2CH_2)_{s1}(CH_2)_{s3}OR'$, wherein s1 is an integer from 1 to 10, each of s2 and s3, independently, is an integer from 0 to 10, and R' is H or $C_{1-20}$ alkyl); or -$NR^{N1}(CH_2)_{s2}(CH_2CH_2O)_{s1}(CH_2)_{s3}NR^{N1}$, wherein s1 is an integer from 1 to 10, each of s2 and s3, independently, is an integer from 0 to 10, and each $R^{N1}$ is, independently, hydrogen or optionally substituted $C_{1-6}$ alkyl.

65. The isolated polynucleotide of any one of claims 35-64, wherein B comprises Formula (b21):



(b21), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $X^{12}$ is, independently, O, S, optionally substituted alkylene, or optionally substituted heteroalkylene; xa is an integer from 0 to 3; $R^{12a}$ is H, optionally substituted alkyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent; and $T^2$ is O, S, or Se.

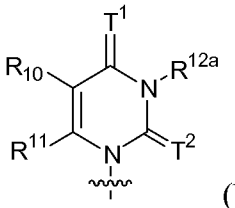
66. The isolated polynucleotide of any one of claims 35-65, wherein B comprises Formula (b22):



(b22), or a pharmaceutically acceptable salt or stereoisomer thereof, wherein $R^{10'}$ is, independently, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted

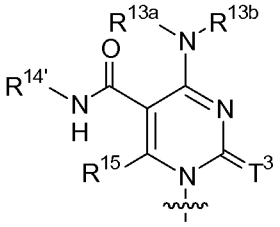
aminoalkynyl, optionally substituted alkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl; $R^{11}$ is H or optionally substituted alkyl; $R^{12a}$ is H, optionally substituted alkyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent; and each of $T^1$ and $T^2$ is, independently, O, S, or Se.

67. The isolated polynucleotide of any one of claims 35-66, wherein B comprises Formula (b23):

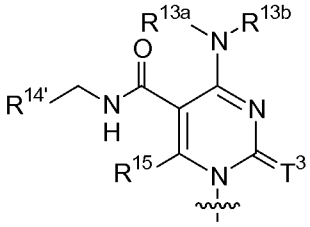


(b23), wherein $R^{10}$ is optionally substituted heterocyclyl or optionally substituted aryl; $R^{11}$ is H or optionally substituted alkyl; $R^{12a}$ is H, optionally substituted alkyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, or absent; and each of $T^1$ and $T^2$ is, independently, O, S, or Se.

68. The isolated polynucleotide of any one of claims 35-67, wherein B comprises Formula (b24) or (b25):



(b24) or



(b25), wherein

$T^3$ is O, S, or Se;

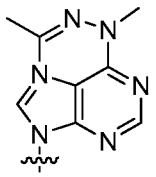
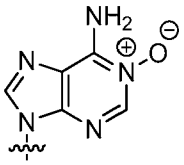
each of $R^{13a}$ and $R^{13b}$ is, independently, H, optionally substituted acyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the combination of $R^{13b}$ and $R^{14}$ can be taken together to form optionally substituted heterocyclyl;

$R^{14'}$ is, independently, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, optionally substituted heterocyclyl, optionally substituted alkheterocyclyl, optionally substituted alkaryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynynyl, optionally substituted alkoxy, optionally substituted alkoxycarbonylalkyl, optionally substituted

alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl; and

each $R^{15}$ is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl.

69. The isolated polynucleotide of any one of claims 35-68, wherein B comprises Formula (b26) or (b27):


(b26) or 
(b27).

70. The polynucleotide of any one of claims 35-69, wherein B comprises Formula (b28)-b43).

71. The isolated polynucleotide of claim 35, wherein said isolated polynucleotide is prepared from one or more building blocks selected from BB- 1 to BB- 274, or a pharmaceutically acceptable salt or stereoisomer thereof.

72. The isolated polynucleotide of claim 35, wherein said isolated polynucleotide is prepared from one or more building blocks selected from compound 1-50, or a pharmaceutically acceptable salt or stereoisomer thereof.