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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 September 2014 (25.09.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/152027 Al

(51) International Patent Classification:
C12P 19/34 (2006.01) *C12P 19/26* (2006.01)
CUP 19/30 (2006.01)

(21) International Application Number:
PCT/US2014/026835

(22) International Filing Date:
13 March 2014 (13.03.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/800,049 15 March 2013 (15.03.2013) US

(71) Applicant: **MODERNA THERAPEUTICS, INC.** [US/US]; 200 Technology Square, Cambridge, MA 02139 (US).

(72) Inventors: **BANCEL, Stephane;** Moderna Therapeutics, Inc., 200 Technology Square, Cambridge, MA 02139 (US). **ISSA, William, Joseph;** Moderna Therapeutics, Inc., 200 Technology Square, Cambridge, MA 02139 (US). **AUNINS, John, Grant;** Moderna Therapeutics, Inc., 200 Technology Square, Cambridge, MA 02139 (US). **CHAKRABORTY, Tirtha;** Moderna Therapeutics, Inc., 200 Technology Square, Cambridge, MA 02139 (US).

(74) Agents: **KIM, Jane, N.** et al; Fenwick & West LLP, 801 California Street, Mountain View, CA 94041 (US).

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

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

[Continued on next page]

(54) **Title:** MANUFACTURING METHODS FOR PRODUCTION OF RNA TRANSCRIPTS

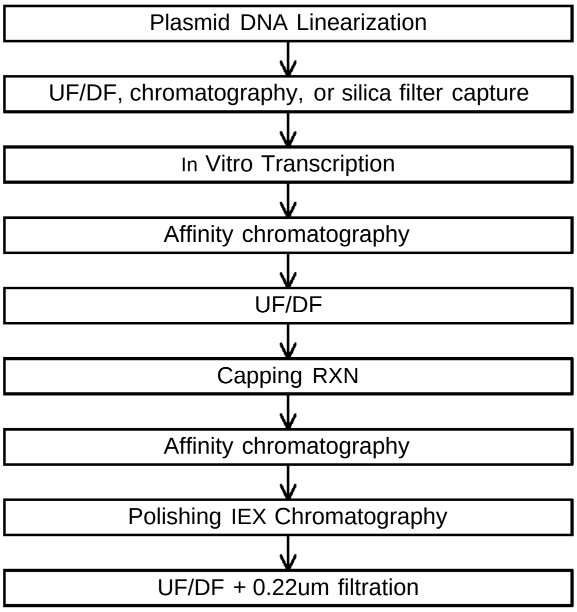


FIG. 1

(57) Abstract: Described are methods for production of RNA transcripts using a non-amplified, linearized DNA tempate in an *in vitro* transcription reaction. Enzymatic 5 ' capping and oligo dT purification can also be included in the methods.

WO 2014/152027 A1

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

TITLE

[0001] Manufacturing Methods for Production of RNA Transcripts.

CROSS REFERENCE TO RELATED APPLICATIONS

[0002] Not applicable.

SEQUENCE LISTING

[0003] Not applicable.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0004] Not applicable.

BACKGROUND OF THE INVENTION

Field of the invention

[0005] The invention relates to methods useful for manufacture of RNA transcripts, e.g., mRNA..

Description of the Related Art

[0006] The manufacturing process used to produce clinical grade chemically modified mRNA therapeutics must be able to generate mRNA of high purity and potency, consistently, reproducibly, and in compliance with current good manufacturing practices (cGMP). The RNA must be as homogeneous as possible; this includes obtaining a uniform cap structure/ 5' terminus, correct sequence, correct poly A tail length and minimizing the formation of product related impurities.

[0007] To generate poly A tail containing mRNAs, several approaches have been taken utilizing in vitro transcription. The first is transcription using a "tailless" DNA template. A 3' Poly A tail is added, post-transcriptionally using a Poly A polymerase. Tails greater than 100 bases are typically generated. The downfall of this approach is the difficulty of controlling tail length and the evolution of tail length distribution in the RNA.

[0008] The use of PCR to generate DNA templates containing a Poly A:T tract is another approach. The poly A:T tracts are introduced via PCR primers, introducing another step for creation of impurities.

SUMMARY OF THE INVENTION

[0009] Disclosed is a method for production of an RNA transcript, e.g., mRNA, using a non-amplified, linearized DNA template in an *in vitro* transcription reaction to generate the RNA transcript. In some embodiments the RNA transcript is capped via enzymatic capping. The method results in production of homogeneous RNA transcripts with high purity and potency. In some embodiments the RNA transcript is purified via chromatographic methods, e.g., use of an oligo dT substrate. In one embodiment the method excludes the use of DNase.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0010] Figure 1 is a flow chart of one embodiment of the method.

[0011] Figure 2 illustrates overhangs of linearized plasmid DNA template with different endonucleases, according to one embodiment.

[0012] Figure 3 illustrates an mRNA cap structure, according to one embodiment.

[0013] Figure 4 illustrates an examplary RNA trasncript diagram, according to one embodiment.

[0014] Figure 5 is a map of plasmid pJ344:91543-TC-GCSF-pA-Vl, according to one embodiment.

[0015] Figure 6 is a DNA sequence of plasmid pJ344:91543-TC-GCSF-pA-Vl, according to one embodiment.

[0016] Figures 7A and 7B are electropherograms of RNA transcripts, according to one embodiment.

[0017] Figure 8 is an electropherogram of capped RNA, according to one embodiment.

[0018] Figure 9 illustrates a bar graph showing results of an ELISA used to quantify expressed protein data, according to one embodiment.

[0019] Figure 10 illustrates a bar graph showing results of an assay for determination of IFN-α induction following transfection of RNAs into PBMC, according to one embodiment.

[0020] Figure 11A is a map of plasmid pJ204:109475-TC-FIX-Hs-3-pA140-Sap, according to one embodiment.

[0021] Figure 1IB is a DNA sequence of plasmid pJ204:109475-TC-FIX-Hs-3-pA140-Sap, according to one embodiment.

DETAILED DESCRIPTION OF THE INVENTION

[0022] Disclosed are methods for production of RNA transcripts, e.g., mRNA, useful for producing clinical grade mRNA of high purity and potency, consistently, reproducibly, and in compliance with current good manufacturing practices (cGMP). The method uses a non-amplified, linearized DNA template in an *in vitro* transcription reaction to generate the RNA transcript. In some embodiments the RNA transcript is purified via oligo dT substrate without the use of DNase. Enzymatic capping is used for 5' capping of the RNA transcript.

[0023] PCR generated DNA templates have disadvantages that can be mitigated using linearized whole plasmid DNA templates A PCR-free process has the following advantages:

[0024] Scalability: Plasmid DNA template can be produced at microgram, milligram and gram scale in a cGMP compliant fashion. Large scale production of PCR generated templates is not commercially viable.

[0025] Higher purity mRNA transcripts than with PCR product templates: There is lower risk that no additional aberrant species are transcribed when using a linearized, non-amplified DNA template. Additionally, PCR amplification creates another intermediate in the manufacturing process where additional impurities may be introduced. When using PCR amplification, the gene of interest is amplified and typically does not contain a poly A tail. The poly A tail is introduced via the reverse primer, not encoded in the plasmid template. Utilizing linearized plasmid DNA template produces RNA transcripts of higher quality and purity

[0026] Increased process efficiency: An additional unit operation (PCR) is removed from the process increasing throughput and reducing production time.

[0027] Whole plasmid allows for simple DNA/RNA separation: Use of whole plasmid DNA template is more amenable to SEC and IEX chromatographic separations compared to use of PCR product due to larger disparity in size and charge density between DNA and RNA. Whole plasmid can also be removed using a poly dT based affinity purification in the flowthrough fraction.

[0028] In addition, there are advantages to the methods use of enzymatic capping. For large scale manufacturing, 5' capping of RNA transcripts is typically performed using a chemical cap analog. This is performed co-transcriptionally where the cap analog : GTP molar ratio in the reaction is ~ 4:1. This typically results in -80% capping efficiency, as well as reduced RNA transcript yields due to consumption of GTP. This high abundance of uncapped species is undesirable when developing therapeutic RNA. Since only capped mRNA is translated into protein, the presence of a high abundance of uncapped species (ie 20%) is problematic as efficacy (protein production/ mg RNA) is reduced by 20%> and 20%> of the final drug substance is an inert impurity, decreasing process productivity.

[0029] The presence of uncapped species is also potentially immunogenic: Presence of a 5' triphosphate motif on uncapped RNAs can be potentially immunostimulatory (see Hornung et. al and Abbas et. al). Use of cap analogs requires additional phosphatase treatment to remove the 5'-triphosphate motif from the RNAs.

[0030] In contrast, the use of enzymatic capping for clinical grade mRNA production is a nearly quantitative capping process that is much more efficient than co-transcriptional capping using cap analogs. This increases potency, process productivity, as well as reduces the potential for immunogenicity.

Definitions

[0031] Terms used in the claims and specification are defined as set forth below unless otherwise specified.

[0032] Polynucleotide. The term "polynucleotide" is interchangeable with nucleic acid, and includes any compound and/or substance that comprise a polymer of nucleotides. RNA transcripts produced by the method of the invention and DNA templates used in the methods of the invention are polynucleotides. Exemplary polynucleotides include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs, including LNA having a β- D-ribo configuration, a-LNA having an a-L-ribo configuration (a diastereomer of LNA), 2'-amino-LNA having a 2'-amino functionalization, and 2'-amino- a-LNA having a 2'-amino functionalization) or hybrids thereof.

[0033] RNA transcript. As used herein, an "RNA transcript" refers to a ribonucleic acid produced by an in vitro transcription reaction using a DNA template and an RNA polymerase. As described in more detail below, an RNA transcript typically includes the coding sequence for a gene of interest and a poly A tail. RNA transcript includes an mRNA. The RNA transcript can include modifications, e.g., modified nucleotides. As used herein, the term RNA transcript includes and is interchangeable with mRNA, modified mRNA "mmRNA" or modified mRNA, and primary construct.

[0034] Gene of interest. As used herein, "gene of interest" refers to a polynucleotide which encodes a polypeptide or protein of interest. Depending on the context, the gene of interest refers to a deoxyribonucleic acid, e.g., a gene of interest in a DNA template which can be transcribed to an RNA transcript, or a ribonucleic acid, e.g., a gene of interest in an RNA transcript which can be translated to produce the encoded polypeptide of interest in vitro, in vivo, in situ or ex vivo. As described in more detail below, a polypeptide of interest includes but is not limited to, biologies, antibodies, vaccines, therapeutic proteins or peptides, etc..

[0035] DNA template. As used herein, a DNA template refers to a polynucleotide template for RNA polymerase. Typically a DNA template includes the sequence for a gene of interest operably linked to a RNA polymerase promoter sequence.

[0036] 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. For example, a gene of interest operably linked to an RNA polymerase promoter allows transcription of the gene of interest.

[0037] Poly A tail. As used herein, "poly A tail" refers to a chain of adenine nucleotides. The term can refer to poly A tail that is to be added to an RNA transcript, or can refer to the poly A tail that already exists at the 3' end of an RNA transcript. As described in more detail below, a poly A tail is typically 5-300 nucleotides in length.

[0038] 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).

[0039] It must be noted that, as used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

Methods of the invention

[0040] Disclosed is a method for production of an RNA transcript, e.g., mRNA, using a non-amplified, linearized DNA template in an *in vitro* transcription reaction to generate the RNA transcript. In some embodiments the RNA transcript is capped via enzymatic capping. The method results in production of homogeneous RNA transcripts with high purity and potency. In some embodiments the RNA transcript is purified via chromatographic methods, e.g., use of an oligo dT substrate. In one embodiment the method excludes the use of DNase.

[0041] In one embodiment, an RNA transcript is synthesized from a non-amplified, linear DNA template coding for the gene of interest via an enzymatic *in vitro* transcription reaction utilizing a T7 phage polymerase and nucleotide triphosphates of the desired chemistry. The RNA transcript is enzymatically capped post transcriptionally at the 5' end using Vaccinia guanylyltransferase, guanosine triphosphate and s-adenosyl-L-methionine to yield cap 0 structure, i.e., an inverted 7-methylguanosine cap is added via a 5' to 5' triphosphate bridge. Alternatively, use of an 2'O-methyltransferase with the Vaccinia guanylyltransferase yields a cap 1 structure where in addition to the cap 0 structure, the 2'OH group is methylated on the penultimate nucleotide. In some embodiments the RNA transcript is chromatographically purified during the manufacturing process using affinity and/or anion exchange methods and diafiltered into the appropriate formulation buffer. This process is PCR free.

DNA templates

[0042] The method used a non amplified, linearized plasmid DNA as the template for *in vitro* transcription. Cells, e.g., bacterial cells, e.g., E. coli, e.g., DH-1 cells are transfected with the plasmid DNA template. The transfected cells are cultured to replicate the plasmid DNA which is then isolated and purified.

[0043] The plasmid DNA template includes a gene of interest coding for, e.g., a polypeptide of interest. A detailed description of genes of interest is found below. In one embodiment, the gene of interest is a GCSF (granular colony stimulating factor) gene. In another embodiment, the gene of interest codes is a Factor IX gene. In another embodiment the gene of interest is a EP (erythropoietin) gene.

[0044] The plasmid DNA template also includes a RNA polymerase promoter, e.g., a T7 promoter located 5' to and operably linked to the gene of interest. Also included is a sequence coding for a poly A tail located 3' to the gene of interest. Additional description of promoters and poly A tails is found below.

[0045] Immediately downstream of the poly A tail coding sequence on the plasmid DNA template is a recognition site for a restriction endonuclease to linearize the plasmid. Linearization of the plasmid mitigates transcriptional readthrough. In one embodiment, the endonuclease produces a 5' overhang on the linearized DNA template. In another embodiment, the endonuclease produces a blunt end on the linearized DNA template. In another embodiment, the endonuclease produces 3' overhang on the linearized DNA template. Examples of endonucleases resulting in 5' overhangs include Xbal, Sapl, Notl, EcoRI, Hindlll, and SPEI. Figure 2 shows examples of such overhangs.Examples of endonucleases resulting in blunt ends include SPEI. Additional endonucleases are well known to one of skill in the art and can be used depending on the application.

[0046] In one embodiment, Xbal is utilized as the restriction endonuclease with a recognition site of 5'TCTAGA3'. The five base overhang left on the DNA template sequence results in additional bases on the 3' end of the RNA transcript post-transcription. In other embodiments, the restriction endonuclease SAP I is used. The resulting overhang on the linearized DNA template sequence does not generate additional bases on the 3' end of the RNA transcript.

[0047] The plasmid into which the gene of interest, promoter, poly A tail sequence, and 5' and 3' UTR, and linearization restriction sites are inserted can be, e.g., one knows to one of skill in the art. Examples include but are not limited to pUC57, pJ204 (from DNA 2.0) and pJ344 (from DNA 2.0).

[0048] Examples of DNA templates to be used in the methods of the invention include, e.g., pJ344:91543 (including a GCSF gene and XbaI site) and pJ204: 109475 (including a Factor IX gene and a SAPI site). Both plasmids are described in detail below.

[0049] In some embodiments, following linearization, the plasmid DNA template is filtered into an appropriate solvent, e.g., water, TE (Tris-EDTA), lOmM Tris HC1 pH 7.5, HEPES/phosphate and the like. Filtration occurs via, e.g., ultrafiltration, diafiltration, or, e.g., tangential flow ultrafiltration/diafiltration.

[0050] The linearized DNA template can be purified before use as a template for *in vitro* transcription. For example, the linearized DNA template can be purified chromatographically, or purified using a silica filter based DNA capture method.

In Vitro Transcription

[0051] The linearized DNA template is used in an in vitro transcription (IVT) system. The system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and an RNA polymerase. The NTPs may be manufactured in house, may be selected from a supplier, or may be synthesized as described herein. The NTPs may be selected from, but are not limited to, those described herein including natural and unnatural (modified) NTPs. Additional description of modified nucleotides is found below.

RNA Polymerases

[0052] Any number of RNA polymerases or variants may be used in the method of the present invention. The polymerase may be selected from, but is not limited to, a phage RNA polymerase, e.g., a T7 RNA polymerase, a T3 RNA polymerase, an SP6 RNa polymerase, and/or mutant polymerases such as, but not limited to, polymerases able to incorporate modified nucleic acids.

[0053] RNA polymerases may be modified by inserting or deleting amino acids of the RNA polymerase sequence. As a non-limiting example, the RNA polymerase may be modified to exhibit an increased ability to incorporate a 2'-modified nucleotide triphosphate compared to an unmodified RNA polymerase (see International Publication WO2008078180 and U.S. Patent 8,101,385; herein incorporated by reference in their entireties).

[0054] Variants may be obtained by evolving an RNA polymerase, optimizing the RNA polymerase amino acid and/or nucleic acid sequence and/or by using other methods known in the art. As a non-limiting example, T7 RNA polymerase variants may be evolved using the continuous directed evolution system set out by Esvelt et al. (Nature (2011) 472(7344) :499-503; herein incorporated by reference in its entirety) where clones of T7 RNA polymerase may encode at least one mutation such as, but not limited to, lysine at position 93 substituted for threonine (K93T), I4M, A7T, E63V, V64D, A65E, D66Y, T76N, C125R, S128R, A136T, N165S, G175R, H176L, Y178H, F182L, L196F, G198V, D208Y, E222K, S228A, Q239R, T243N, G259D, M267I, G280C, H300R, D351A, A354S, E356D, L360P, A383V, Y385C, D388Y, S397R, M401T, N410S, K450R, P451T, G452V, E484A, H523L, H524N, G542V, E565K, K577E, K577M, N601S, S684Y, L699I, K713E, N748D, Q754R, E775K, A827V, D851N or L864F. As another non-limiting example, T7 RNA polymerase variants may encode at least mutation as described in U.S. Pub. Nos. 20100120024 and 20070117112; herein incorporated by reference in their entireties. Variants of RNA polymerase may also include, but are not limited to, substitutional variants, conservative amino acid substitution, insertional variants, deletional variants and/or covalent derivatives.

[0055] In one embodiment, the RNA transcript may be designed to be recognized by the wild type or variant RNA polymerases. In doing so, the RNA transcript may be modified to contain sites or regions of sequence changes from the wild type or parent primary construct.

Transcription reaction conditions

[0056] Typical reaction conditions for *in vitro* transcription can be as follows. One of skill in the art understands that parameters can be changed according to the application, e.g., the polymerase and nucleotides used.

[0057] The *in vitro* transcription reaction includes the following: an RNA polymerase, e.g., a T7 RNA polymerase at a final concentration of, e.g., 1000-12000 U/mL, e.g., 7000 U/mL; the DNA template at a final concentration of, e.g., 10-70 nM, e.g., 40 nM; nucleotides (NTPs) at a final concentration of e.g., 0.5-10 nM, e.g., 7.5 nM each; magnesisum at a final concentration of, e.g., 12-60 mM, e.g., magnesium acetate at 40 mM; a buffer such as, e.g., HEPES or Tris at a pH of, e.g., 7-8.5, e.g. 40 mM Tris HCl, pH 8

[0058] In some embodiments 5 mM dithiothreitol (DTT) and/or 1 mM spermidine is included. In some embodiments, an RNase inhibitor is included in the *in vitro* transcription reaction to ensure no RNase induced degradation during the transcription reaction.. For example, murine RNase inhibitor can be utilized at a final concentration of 1000 U/mL, In some embodiments a pyrophosphatase is included in the *in vitro* transcription reaction to cleave the inorganic pyrophosphate generated following each nucleotide incorporation into two units of inorganic phosphate. This ensures that magnesium, which is essential for transcription, remains in solution and does not precipitate as magnesium pyrophosphate. For example, an E. Coli inorganic pyrophosphatase can utilized at a final concentration of 1 U/mL.

[0059] The *in vitro* transcription reaction is allowed to proceed, for example, under constant mixing at 37°C for 4 hours. Typical yields can be, e.g., 5 mg of RNA per mL of transcription reaction.

[0060] A typical *in vitro* transcription reaction might have the following properties:

Lot #											
Template:											
Process:											
Duration/Temp:											
Projected Total RNA Yield	**30**	**mg**									
			300	RXNs	** Enter			20	uL/rxn		
	Stock Concentration						effective concentration				
water			2170.1	uL							
DNA template	1450	ng/uL	379.9	uL or	550.8	ug	40	nM	enter***	91.8	ug/mL DNA
T7 Buffer	10	X	600.0	uL			1	X			
ATP (100mM)	100	mM	450.0	uL			7.5	mM	1.84	ug/rxn DNA template	
GTP (100mM)	100	mM	450.0	uL			7.5	mM			
CTP (100mM)	100	mM	450.0	uL			7.5	mM			
UTP (100mM)	100	mM	450.0	uL			7.5	mM			
RNase inhibitor	40,000	U/mL	150.0	uL			1000	U/mL			
T7 Polymerase	50000	U/mL	840.0	uL			7000	u/mL			
pyrophosphatase	100	U/mL	60.0	uL			1	U/mL			
Total Volume			**6000**	**uL**							
Conc of DNA template:			**1450**	**ng/uL**	**** Enter**			Do not use DNase if using whole plasmid as a DNA template			
RNA yield projection			5	mg RNA/mL IVT		30	mg		15.9	uM RNA in Solution	
Length of RNA			922	bases	** Enter						
RNA MW			313,480	Da							
DNA MW			2,295,067	Da							
Length of DNA Template			3,781	BP	For Plasmid: Enter length of whole Plasmid						
					For PCR : Enter length of PCR product calculated below						

Nucelotides

[0061] The *in vitro* transcription reaction includes nucleotides (NTPs). The nucleotides can be unmodified NTPs, e.g., A, C, C, and U ribonucleotides, or modified nucleotides, or a combination. A more detailed description of modified nucleotides is found below. Examples of nucleotide combinations useful for *in vitro* transcription are found in the table below.

Chemistry	A	G	C	U
Gen 0	ATP	GTP	CTP	UTP
Gen 1	ATP	GTP	5mCTP	ψTP
Gen 2	ATP	GTP	5mCTP	N-1-meψTP
Gen 3	ATP	GTP	5mCTP	2-thio UTP
Gen 4	ATP	GTP	CTP	ψTP
Gen 5	ATP	GTP	CTP	N-1-meψTP

Example combinations of nucleotides for *in vitro* transcription.

Key: ATP = Adenosine triphosphate; GTP = Guanosine triphosphate; CTP= Cytidine Triphosphate; UTP= Uridine Triphosphate; 5mCTP = 5-methylcytidine triphosphate; 2-thioUTP = 2-thiouridine triphosphate; ψTP= pseudouridine triphosphate; N-l-me ψTP = 1-methylpseudouridine triphosphate.

Capping of RNA transcripts

[0062] In some embodiments the RNA transcript is enzymatically capped at the 5/ end after *in vitro* transcription. Capping can be performed either before or after further purification of the RNA transcript, e.g., oligo dT purification. If not performing purification prior to capping, an ultrafiltration/diafiltration step must be performed (e.g., buffer exchange).

[0063] Non-limiting examples of 5' cap structures are those which, among other things, have enhanced binding of cap binding proteins, increased half life, reduced susceptibility to 5' endonucleases and/or reduced 5'decapping, as compared to synthetic 5'cap structures known in the art (or to a wild-type, natural or physiological 5'cap structure). For example, recombinant Vaccinia Virus Capping Enzyme and recombinant 2'-0-methyltransferase enzyme can create a canonical 5'-5'-triphosphate linkage between the 5'-terminal nucleotide of an mRNA and a guanine cap nucleotide wherein the cap guanine contains an N7 methylation and the 5'-terminal nucleotide of the mRNA contains a 2'-0-methyl. Such a structure is termed the Capl structure. This cap results in a higher translational-competency and cellular stability and a reduced activation of cellular pro-inflammatory cytokines, as compared, e.g., to other 5'cap analog structures known in the art. Cap structures include, but are not limited to, 7mG(5')ppp(5')N,pN2p (cap 0), 7mG(5')ppp(5')NlmpNp (cap 1), and 7mG(5')-ppp(5')NlmpN2mp (cap 2). An example of a cap structure is illustrated in Figure 3.

[0064] For example, the RNA transcript can be enzymatically capped at the 5' end using Vaccinia guanylyltransferase, guanosine triphosphate and s-adenosyl-L-methionine to yield cap 0 structure. An inverted 7-methylguanosine cap is added via a 5' to 5' triphosphate bridge. Alternatively, use of a 2'0-methyltransferase with Vaccinia guanylyltransferase yields the cap 1 structure where in addition to the cap 0 structure, the 2'O H group is methylated on the penultimate nucleotide. S-adenosyl-L-methionine (SAM) is a cofactor utilized as a methyl transfer reagent.

[0065] In one embodiment enzymatic 5' capping is performed as follows. S-adenosylmethione chloride * 2HC1 is dissolved at 20mM in 5mM HC1 10/90 v/v% ethanol/ water as a prepared stock. RNase inhibitor are utilized as a safeguard to ensure no RNase degradation is observed during the reaction. The final IX buffer conditions consist of the following: 50mM Tris HC1 pH 8, 5mM KC1, ImM MgC12, and ImM dithiothreitol. The reaction is run under constant mixing at 37°C for 2 hours. Enzymatic capping is of considerably higher efficiency than performing co-transcription through the use of dinucleotide cap analogs.

[0066] In another embodiment, the RNA transcript is co-transcriptionall capped. The 5' terminal caps may include endogenous caps or cap analogs. A 5' terminal cap may comprise a guanine analog. Useful guanine analogs include, but are not limited to, inosine, Nl-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

RNA transcripts

[0067] An RNA transcript is the polynucleotide product of the *in vitro* transcription reaction. Typically the RNA transcript includes a gene of interest and a poly-A tail. In some embodiments, the RNA transcript includes a 5'UTR and a 3'UTR. In one embodiment the RNA transcript includes a 5' cap, typically added post trasncriptionally.

[0068] Figure 4 illustrates an exemplary RNA transcript. The RNA transcript 100 here contains a first region of linked nucleotides 102 that is flanked by a first flanking region 104 and a second flaking region 106. As used herein, the "first region" may be referred to as a "coding region" or "region encoding" or simply the "first region." This first region may include, but is not limited to, the gene of interest. The polypeptide of interest may comprise at its 5' terminus one or more signal sequences which are encoded by a signal sequence region 103. The flanking region 104 may comprise a region of linked nucleotides comprising one or more complete or incomplete 5' UTRs sequences. The flanking region 104 may also comprise a 5' terminal cap 108. The second flanking region 106 may comprise a region of linked nucleotides comprising one or more complete or incomplete 3' UTRs. The flanking region 106 may also comprise a 3' tailing sequence 110, e.g., a poly A tail sequence.

[0069] Bridging the 5' terminus of the first region 102 and the first flanking region 104 is a first operational region 105. Traditionally this operational region comprises a Start codon. The operational region may alternatively comprise any translation initiation sequence or signal including a Start codon.

[0070] Bridging the 3' terminus of the first region 102 and the second flanking region 106 is a second operational region 107. Traditionally this operational region comprises a Stop codon. The operational region may alternatively comprise any translation initiation sequence or signal including a Stop codon. According to the present invention, multiple serial stop codons may also be used.

[0071] Generally, the shortest length of the first region of the RNA transcript, e.g., the gene of interest can be the length of a nucleic acid sequence that is sufficient to encode for a dipeptide, a tripeptide, a tetrapeptide, a pentapeptide, a hexapeptide, a heptapeptide, an octapeptide, a nonapeptide, or a decapeptide. In another embodiment, the length may be sufficient to encode a peptide of 2-30 amino acids, e.g. 5-30, 10-30, 2-25, 5-25, 10-25, or 10-20 amino acids. The length may be sufficient to encode for a peptide of at least 11, 12, 13, 14, 15, 17, 20, 25 or 30 amino acids, or a peptide that is no longer than 40 amino acids, e.g. no longer than 35, 30, 25, 20, 17, 15, 14, 13, 12, 11 or 10 amino acids. Examples of dipeptides that the polynucleotide sequences can encode or include, but are not limited to, carnosine and anserine.

[0072] Generally, the length of the first region encoding the polypeptide of interest of the present invention is greater than about 30 nucleotides in length (e.g., at least or greater than about 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, and 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000 or up to and including 100,000 nucleotides).

[0073] In some embodiments, the RNA transcript includes from about 30 to about 100,000 nucleotides (e.g., from 30 to 50, from 30 to 100, from 30 to 250, from 30 to 500, from 30 to 1,000, from 30 to 1,500, from 30 to 3,000, from 30 to 5,000, from 30 to 7,000, from 30 to 10,000, from 30 to 25,000, from 30 to 50,000, from 30 to 70,000, from 100 to 250, from 100 to 500, from 100 to 1,000, from 100 to 1,500, from 100 to 3,000, from 100 to 5,000, from 100 to 7,000, from 100 to 10,000, from 100 to 25,000, from 100 to 50,000, from 100 to 70,000, from 100 to 100,000, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to

3,000, from 500 to 5,000, from 500 to 7,000, from 500 to 10,000, from 500 to 25,000, from 500 to 50,000, from 500 to 70,000, from 500 to 100,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 3,000, from 1,000 to 5,000, from 1,000 to 7,000, from 1,000 to 10,000, from 1,000 to 25,000, from 1,000 to 50,000, from 1,000 to 70,000, from 1,000 to 100,000, from 1,500 to 3,000, from 1,500 to 5,000, from 1,500 to 7,000, from 1,500 to 10,000, from 1,500 to 25,000, from 1,500 to 50,000, from 1,500 to 70,000, from 1,500 to 100,000, from 2,000 to 3,000, from 2,000 to 5,000, from 2,000 to 7,000, from 2,000 to 10,000, from 2,000 to 25,000, from 2,000 to 50,000, from 2,000 to 70,000, and from 2,000 to 100,000).

[0074] In some embodiments, the first and second flanking regions may range independently from 15-1,000 nucleotides in length (e.g., greater than 30, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, and 900 nucleotides or at least 30, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, and 1,000 nucleotides).

[0075] The tailing sequence may range from 1 to 500 nucleotides in length (e.g., at least 60, 70, 80, 90, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, or 500 nucleotides). Where the tailing region is a polyA tail, the length may be determined in units of or as a function of polyA Binding Protein binding. In this embodiment, the polyA tail is long enough to bind at least 4 monomers of PolyA Binding Protein. PolyA Binding Protein monomers bind to stretches of approximately 38 nucleotides. As such, it has been observed that polyA tails of about 80 nucleotides and 160 nucleotides are functional.

[0076] According to the present invention, the capping region may comprise a single cap or a series of nucleotides forming the cap. In this embodiment the capping region may be from 1 to 10, e.g. 2-9, 3-8, 4-7, 1-5, 5-10, or at least 2, or 10 or fewer nucleotides in length. In some embodiments, the cap is absent.

[0077] According to the present invention, the first and second operational regions may range from 3 to 40, e.g., 5-30, 10-20, 15, or at least 4, or 30 or fewer nucleotides in length and may comprise, in addition to a Start and/or Stop codon, one or more signal and/or restriction sequences.

[0078] The RNA transcripts can comprise one or more structural and/or chemical modifications or alterations which impart useful properties to the polynucleotide including, in some embodiments, the lack of a substantial induction of the innate immune response of a cell into which the polynucleotide is introduced. As used herein, a "structural" feature or modification is one in which two or more linked nucleotides are inserted, deleted, duplicated, inverted or randomized in an RNA transcript without significant chemical modification to the nucleotides themselves. Because chemical bonds will necessarily be broken and reformed to effect a structural modification, structural modifications are of a chemical nature and hence are chemical modifications. However, structural modifications will result in a different sequence of nucleotides. For example, the polynucleotide "ATCG" may be chemically modified to "AT-5meC-G". The same polynucleotide may be structurally modified from "ATCG" to "ATCCCG". Here, the dinucleotide "CC" has been inserted, resulting in a structural modification to the polynucleotide.

Removal of DNA template from RNA transcript

[0079] The linearized plasmid DNA template is removed from the *in vitro* transcription, e.g., the DNA template is separated from the RNA transcript. In one embodiment, the DNA template is removed chromatographically using an poly A capture, e.g., oligo dT, based affinity purification step . The RNA transcript binds affinity substrate while the DNA template flow through and is removed. .

[0080] It is typical to utilize DNase I to enzymatically digest DNA template immediately following in vitro transcription. In the methods of the invention, DNase is not utilized. Whole plasmid removal is preferred to enzymatic digestion due to the fact that the risk of degraded DNA fragments hybridizing to the transcribed mRNA is mitigated.

[0081] In one embodiment the poly A capture based affinity purification is oligo dT purification. For example, a polythymidine ligand is immobilized to a derivatized chromatography resin. The mechanism of purification involves hybridization of the poly A tail of the RNA transcript to the oligonucleotide ligand. The DNA template will not bind. In addition, RNA transcripts that do not contain Poly A stretches (short aborts and other truncates formed during in vitro transcription) will not bind to the resin and will not form a duplex with the affinity ligand. Poly Adenylated RNA can then be eluted from the resin utilizing a low ionic strength buffer or a competitive binding oligonucleotide solution. A one pot purification method can yield highly purified poly A containing RNA with recoveries >80% actively removes endotoxin, DNA template, and enzymes utilized in the production of RNA using a simple capture and elute methodology with no subsequent fraction of captured poly A containing RNA. This purification increases mRNA product purity and in turn significantly increases target protein expression.

Additional purification steps

[0082] The method for production of an RNA transcript can include additional purification steps after the *in vitro* transcription, e.g., an ion exchange chromatography step.

Characterization and analysis of RNA transcript

[0083] The RNA transcript produced by the methods of the invention can be analyzed and characterized using any number of methods. Analysis can be performed before or after capping. Alternatively, analysis can be performed before or after poly A capture based affinity purification. In another embodiment, analysis can be performed before or after additional purification steps, e.g., anion exchange chromatography and the like.

[0084] For example, RNA template quality can be determined using Bioanalyzer chip based electrophoresis system. In other embodiments, RNA template purity is analysed using analytical reverse phase HPLC respectively. Capping efficiency can be analyzed using, e.g., total nuclease digestion followed by MS/MS quantitation of the dinucleotide cap species vs. uncapped GTP species. In vitro efficacy can be analyzed by, e.g., transfecting RNA transcript into a human cell line,e .g., HeLA, PBMC, BJ Fibroblasts, Hek 293). Protein expression of the polypeptide of interest can be quantified using methods such as ELISA or flow cytometry. Immunogenicity can be analyzed by, e.g., transfecting RNA transcripts into cell lines that indicate innate immune stimulation, e.g., PBMCs. Cytokine induction can be analyzed using, e.g., methods such as ELISA to quantify a cytokine, e.g., Interferon-a.

[0085] The method of producing RNA transcript of the invention can produce RNA transcript that is at least 30% full length transcript, or at least 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, or at least 95%> full length transcript. Purity can be determined as described herein, e.g., via reverse phase HPLC or Bioanalyzer chip based electrophoresis and measure by, e.g., peak area of full length RNA transcript relative to total peak.

Genes of interest

[0086] The DNA template and resulting RNA transcript of the present invention include a gene of interest. The gene of interest encodes a polypeptide of interest selected from, e.g., biologies, antibodies, vaccines, therapeutic proteins or peptides, cell penetrating peptides, secreted proteins, plasma membrane proteins, cytoplasmic or cytoskeletal proteins, intracellular membrane bound proteins, nuclear proteins, proteins associated with human disease, targeting moieties or those proteins encoded by the human genome for which no therapeutic indication has been identified but which nonetheless have utility in areas of research and discovery. The sequence for a particular gene of interest is readily identified by one of skill in the art using public and private databases, e.g., GenBank.

[0087] In one embodiment the gene of interest may encode variant polypeptides which have a certain identity with a reference polypeptide sequence. As used herein, a "reference polypeptide sequence" refers to a starting polypeptide sequence. Reference sequences may be wild type sequences or any sequence to which reference is made in the design of another sequence.

[0088] 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).

[0089] In some embodiments, the polypeptide variant may have the same or a similar activity as the reference polypeptide. Alternatively, the variant may have an altered activity (e.g., increased or decreased) relative to a reference polypeptide. Generally, variants of a particular polynucleotide or polypeptide of the invention will have at least about 40%, 45%>, 50%>, *55%,* 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% but less than 100% 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. Such tools for alignment include those of the BLAST suite (Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.) Other tools are described herein, specifically in the definition of "Identity."

[0090] Default parameters in the BLAST algorithm include, for example, an expect threshold of 10, Word size of 28, Match/Mismatch Scores 1, -2, Gap costs Linear. Any filter can be applied as well as a selection for species specific repeats, e.g., Homo sapiens.

Biologies

[0091] The gene of interest can encode one or more biologies. As used herein, a "biologic" is a polypeptide-based molecule produced by the methods provided herein and which may be used to treat, cure, mitigate, prevent, or diagnose a serious or life-threatening disease or medical condition. Biologies, according to the present invention include, but are not limited to, allergenic extracts (e.g. for allergy shots and tests), blood components, gene therapy products, human tissue or cellular products used in transplantation, vaccines, monoclonal antibodies, cytokines, growth factors, enzymes, thrombolytics, and immunomodulators, among others.

[0092] According to the present invention, one or more biologies currently being marketed or in development may be encoded by the polynucleotides, primary constructs or mrnRNA of the present invention. While not wishing to be bound by theory, it is believed that incorporation of the encoding polynucleotides of a known biologic into the primary constructs or mmRNA of the invention will result in improved therapeutic efficacy due at least in part to the specificity, purity and/or selectivity of the construct designs.

Antibodies

[0093] The gene of interest can encode one or more antibodies or fragments thereof. The term "antibody" includes monoclonal antibodies (including full length antibodies which have an immunoglobulin Fc region), antibody compositions with polyepitopic specificity, multispecific antibodies (e.g., bispecific antibodies, diabodies, and single-chain molecules), as well as antibody fragments. The term "immunoglobulin" (Ig) is used interchangeably with "antibody" herein. As used herein, the term "monoclonal antibody" refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations and/or post-translation modifications (e.g., isomerizations, amidations) that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site.

[0094] The monoclonal antibodies herein specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is(are) identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity. Chimeric antibodies of interest herein include, but are not limited to, "primatized" antibodies comprising variable domain antigen-binding sequences derived from a non-human primate (e.g., Old World Monkey, Ape etc.) and human constant region sequences.

[0095] An "antibody fragment" comprises a portion of an intact antibody, preferably the antigen binding and/or the variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')2 and Fv fragments; diabodies; linear antibodies; nanobodies; single-chain antibody molecules and multispecific antibodies formed from antibody fragments.

[0096] Any of the five classes of immunoglobulins, IgA, IgD, IgE, IgG and IgM, may be encoded by the mmRNA of the invention, including the heavy chains designated alpha, delta, epsilon, gamma and mu, respectively. Also included are polynucleotide sequences encoding the subclasses, gamma and mu. Hence any of the subclasses of antibodies may be encoded in part or in whole and include the following subclasses: IgGl, IgG2, IgG3, IgG4, IgAl and IgA2.

[0097] In one embodiment, the gene of interest can encode monoclonal antibodies and/or variants thereof. Variants of antibodies may also include, but are not limited to, substitutional variants, conservative amino acid substitution, insertional variants, deletional variants and/or covalent derivatives. In one embodiment, the primary construct and/or mmRNA disclosed herein may encode an immunoglobulin Fc region. In another embodiment, the primary constructs and/or mmRNA may encode a variant immunoglobulin Fc region. As a non-limiting example, the primary constructs and/or mmRNA may encode an antibody having a variant immunoglobulin Fc region as described in U.S. Pat. No. 8,217,147 herein incorporated by reference in its entirety.

Vaccines

[0098] The gene of interest can encode one or more vaccines. As used herein, a "vaccine" is a biological preparation that improves immunity to a particular disease or infectious agent. According to the present invention, one or more vaccines currently being marketed or in development may be encoded by the polynucleotides, primary constructs or mmRNA of the present invention. While not wishing to be bound by theory, it is believed that incorporation into the primary constructs or mmRNA of the invention will result in improved therapeutic efficacy due at least in part to the specificity, purity and selectivity of the construct designs.

[0099] Vaccines encoded in the polynucleotides, primary constructs or mmRNA of the invention may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, cardiovascular, CNS, dermatology, endocrinology, oncology, immunology, respiratory, and anti-infective.

Therapeutic proteins or peptides

[00100] The gene of interest can encode one or more validated or "in testing" therapeutic proteins or peptides. "Therapeutic protein" refers to a protein that, when administered to a cell has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect.

[00101] Therapeutic proteins and peptides may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, blood, cardiovascular, CNS, poisoning (including antivenoms), dermatology, endocrinology, genetic, genitourinary, gastrointestinal, musculoskeletal, oncology, and immunology, respiratory, sensory and anti-infective.

Cell-Penetrating Polypeptides

[00102] The gene of interest can encode one or more cell-penetrating polypeptides. As used herein, "cell-penetrating polypeptide" or CPP refers to a polypeptide which may facilitate the cellular uptake of molecules. A cell-penetrating polypeptide of the present invention may contain one or more detectable labels. The polypeptides may be partially labeled or completely labeled throughout. The gene of interest can encode the detectable label completely, partially or not at all. The cell-penetrating peptide may also include a signal sequence. As used herein, a "signal sequence" refers to a sequence of amino acid residues bound at the amino terminus of a nascent protein during protein translation. The signal sequence may be used to signal the secretion of the cell-penetrating polypeptide.

[00103] In one embodiment, the gene of interest can encode a fusion protein. The fusion protein may be created by operably linking a charged protein to a therapeutic protein. As used herein, "operably linked" refers to the therapeutic protein and the charged protein being connected in such a way to permit the expression of the complex when introduced into the cell. As used herein, "charged protein" refers to a protein that carries a positive, negative or overall neutral electrical charge. Preferably, the therapeutic protein may be covalently linked to the charged protein in the formation of the fusion protein. The ratio of surface charge to total or surface amino acids may be approximately 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9.

[00104] The cell-penetrating polypeptide may form a complex after being translated. The complex may comprise a charged protein linked, e.g. covalently linked, to the cell-penetrating polypeptide.

[00105] In one embodiment, the cell-penetrating polypeptide may comprise a first domain and a second domain. The first domain may comprise a supercharged polypeptide. The second domain may comprise a protein-binding partner. As used herein, "protein-binding partner" includes, but is not limited to, antibodies and functional fragments thereof, scaffold proteins, or peptides. The cell-penetrating polypeptide may further comprise an intracellular binding partner for the protein-binding partner. The cell-penetrating polypeptide may be capable of being secreted from a cell where the polynucleotide, primary construct or mmRNA may be introduced. The cell-penetrating polypeptide may also be capable of penetrating the first cell.

[00106] In a further embodiment, the cell-penetrating polypeptide is capable of penetrating a second cell. The second cell may be from the same area as the first cell, or it may be from a different area. The area may include, but is not limited to, tissues and organs. The second cell may also be proximal or distal to the first cell.

[00107] In one embodiment, the cell-penetrating polypeptide which may comprise a protein-binding partner. The protein binding partner may include, but is not limited to, an antibody, a supercharged antibody or a functional fragment. The polynucleotides, primary constructs or mmRNA may be introduced into the cell where a cell-penetrating polypeptide comprising the protein-binding partner is introduced.

Additional proteins

[00108] The gene of interest can encode a secreted protein. The secreted proteins may be selected from, e.g., those in US Patent Publication, 20100255574, the contents of which are incorporated herein by reference in their entirety.

[00109] Alternatively, the gene of interest can encode a protein of the plasma membrane; a cytoplasmic or cytoskeletal proteins; a intracellular membrane bound proteins; or a nuclear protein; a targeting moiety (e.g., 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); antimicrobial peptides (AMP) or antiviral peptides (AVP).

Flanking Regions: Untranslated Regions (UTRs)

[00110] The DNA template and RNA transcript can include UTRs. Untranslated regions (UTRs) of a gene are transcribed but not translated. The 5'UTR starts at the transcription start site and continues to the start codon but does not include the start codon; whereas, the 3'UTR starts immediately following the stop codon and continues until the transcriptional termination signal. There is growing body of evidence about the regulatory roles played by the UTRs in terms of stability of the nucleic acid molecule and translation. The regulatory features of a UTR can be incorporated into the polynucleotides, primary constructs and/or mmRNA of the present invention to enhance the stability of the molecule. The specific features can also be incorporated to ensure controlled down-regulation of the transcript in case they are misdirected to undesired organs sites.

5 ' UTR and Translation Initiation

[00111] Natural 5'UTRs bear features which play roles in translation initiation. They harbor signatures like Kozak sequences which are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus CCR(A/G)CCAUGG, where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G'. 5'UTR also have been known to form secondary structures which are involved in elongation factor binding.

[00112] By engineering the features typically found in abundantly expressed genes of specific target organs, one can enhance the stability and protein production of the polynucleotides, primary constructs or mmRNA of the invention. For example, introduction of 5' UTR of liver-expressed mRNA, such as albumin, serum amyloid A, Apolipoprotein A/B/E, transferrin, alpha fetoprotein, erythropoietin, or Factor VIII, could be used to enhance expression of a nucleic acid molecule, such as a mmRNA, in hepatic cell lines or liver. Likewise, use of 5' UTR from other tissue-specific mRNA to improve expression in that tissue is possible for muscle (MyoD, Myosin, Myoglobin, Myogenin, Herculin), for endothelial cells (Tie-1, CD36), for myeloid cells (C/EBP, AML1, G-CSF, GM-CSF, CDl lb, MSR, Fr-1, i-NOS), for leukocytes (CD45, CD18), for adipose tissue (CD36, GLUT4, ACRP30, adiponectin) and for lung epithelial cells (SP-A/B/C/D).

[00113] Other non-UTR sequences may be incorporated into the 5' (or 3' UTR) UTRs. For example, introns or portions of introns sequences may be incorporated into the flanking regions of the polynucleotides, primary constructs or mmRNA of the invention. Incorporation of intronic sequences may increase protein production as well as mRNA levels.

3 ' UTR and the A U Rich Elements

[00114] 3' UTRs are known to have stretches of Adenosines and Uridines embedded in them. These AU rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU rich elements (AREs) can be separated into three classes (Chen et al, 1995): Class I AREs contain several dispersed copies of an AUUUA motif within U-rich regions. C-Myc and MyoD contain class I AREs. Class II AREs possess two or more overlapping UUAUUUA(U/A)(U/A) nonamers. Molecules containing this type of AREs include GM-CSF and TNF-a. Class III ARES are less well defined. These U rich regions do not contain an AUUUA motif. c-Jun and Myogenin are two well-studied examples of this class. Most proteins binding to the AREs are known to destabilize the messenger, whereas members of the ELAV family, most notably HuR, have been documented to increase the stability of mRNA. HuR binds to AREs of all the three classes. Engineering the HuR specific binding sites into the 3' UTR of nucleic acid molecules will lead to HuR binding and thus, stabilization of the message in vivo.

[00115] Introduction, removal or modification of 3' UTR AU rich elements (AREs) can be used to modulate the stability of polynucleotides, primary constructs or mrnRNA of the invention. When engineering specific polynucleotides, primary constructs or mrnRNA, one or more copies of an ARE can be introduced to make polynucleotides, primary constructs or mrnRNA of the invention less stable and thereby curtail translation and decrease production of the resultant protein. Likewise, AREs can be identified and removed or mutated to increase the intracellular stability and thus increase translation and production of the resultant protein. Transfection experiments can be conducted in relevant cell lines, using polynucleotides, primary constructs or mrnRNA of the invention and protein production can be assayed at various time points post-transfection. For example, cells can be transfected with different ARE-engineering molecules and by using an ELISA kit to the relevant protein and assaying protein produced at 6 hour, 12 hour, 24 hour, 48 hour, and 7 days post-transfection.

Incorporating microRNA Binding Sites

[00116] microRNAs (or miRNA) are 19-25 nucleotide long noncoding RNAs that bind to the 3'UTR of nucleic acid molecules and down-regulate gene expression either by reducing nucleic acid molecule stability or by inhibiting translation. The polynucleotides, primary constructs or mrnRNA of the invention may comprise one or more microRNA target sequences, microRNA seqences, or microRNA seeds. Such sequences may correspond to any known microRNA such as those taught in US Publication US2005/0261218 and US Publication US2005/0059005, the contents of which are incorporated herein by reference in their entirety.

[00117] A microRNA sequence comprises a "seed" region, i.e., a sequence in the region of positions 2-8 of the mature microRNA, which sequence has perfect Watson-Crick complementarity to the miRNA target sequence. A microRNA seed may comprise positions 2-8 or 2-7 of the mature microRNA. In some embodiments, a microRNA seed may comprise 7 nucleotides (e.g., nucleotides 2-8 of the mature microRNA), wherein the seed-complementary site in the corresponding miRNA target is flanked by an adenine (A) opposed to microRNA position 1. In some embodiments, a microRNA seed may comprise 6 nucleotides (e.g., nucleotides 2-7 of the mature microRNA), wherein the seed-complementary site in the corresponding miRNA target is flanked byan adenine (A) opposed to microRNA position 1. See for example, Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP; Mol Cell. 2007 Jul 6;27(1):91-105; each of which is herein incorporated by reference in their entirety. The bases of the microRNA seed have complete complementarity with the target sequence. By engineering microRNA target sequences into the 3'UTR of RNA transcripts of the invention one can target the molecule for degradation or reduced translation, provided the microRNA in question is available. This process will reduce the hazard of off target effects upon nucleic acid molecule delivery. Identification of microRNA, microRNA target regions, and their expression patterns and role in biology have been reported (Bonauer et al, Curr Drug Targets 2010 11:943-949; Anand and Cheresh Curr Opin Hematol 2011 18:171-176; Contreras and Rao Leukemia 2012 26:404-413 (2011 Dec 20. doi: 10.1038/leu.2011.356); Bartel Cell 2009 136:215-233; Landgraf et al, Cell, 2007 129:1401-1414; each of which is herein incorporated by reference in its entirety).

[00118] For example, if the nucleic acid molecule is an mRNA and is not intended to be delivered to the liver but ends up there, then miR-122, a microRNA abundant in liver, can inhibit the expression of the gene of interest if one or multiple target sites of miR-122 are engineered into the 3' UTR of the RNA transcripts. Introduction of one or multiple binding sites for different microRNA can be engineered to further decrease the longevity, stability, and protein translation of a RNA transcripts.

[00119] As used herein, the term "microRNA site" refers to a microRNA target site or a microRNA recognition site, or any nucleotide sequence to which a microRNA binds or associates. It should be understood that "binding" may follow traditional Watson-Crick hybridization rules or may reflect any stable association of the microRNA with the target sequence at or adjacent to the microRNA site.

[00120] Conversely, for the purposes of the RNA transcripts of the present invention, microRNA binding sites can be engineered out of (i.e. removed from) sequences in which they naturally occur in order to increase protein expression in specific tissues. For example, miR- 122 binding sites may be removed to improve protein expression in the liver. Regulation of expression in multiple tissues can be accomplished through introduction or removal or one or several microRNA binding sites.

[00121] Examples of tissues where microRNA are known to regulate mRNA, and thereby protein expression, include, but are not limited to, liver (miR-122), muscle (miR-133, miR-206, miR-208), endothelial cells (miR- 17-92, miR- 126), myeloid cells (miR-142-3p, miR-142-5p, miR-16, miR-21, miR-223, miR-24, miR-27), adipose tissue (let-7, miR-30c), heart (miR- Id, miR- 149), kidney (miR- 192, miR- 194, miR-204), and lung epithelial cells (let-7, miR-133, miR- 126). MicroRNA can also regulate complex biological processes such as angiogenesis (miR- 132) (Anand and Cheresh Curr Opin Hematol 201 1 18:171-176; herein incorporated by reference in its entirety). In the RNA transcripts of the present invention, binding sites for microRNAs that are involved in such processes may be removed or introduced, in order to tailor the expression of the RNA transcripts expression to biologically relevant cell types or to the context of relevant biological processes. A listing of MicroRNA, miR sequences and miR binding sites is listed in Table 9 of U.S. Provisional Application No. 61/753,661 filed January 17, 2013, in Table 9 of U.S. Provisional Application No. 61/754,159 filed January 18, 2013, and in Table 7 of U.S. Provisional Application No. 61/758,921 filed January 31, 2013, each of which are herein incorporated by reference in their entireties. g

[00122] Lastly, through an understanding of the expression patterns of microRNA in different cell types, RNA transcripts can be engineered for more targeted expression in specific cell types or only under specific biological conditions. Through introduction of tissue-specific microRNA binding sites, RNA transcripts could be designed that would be optimal for protein expression in a tissue or in the context of a biological condition.

[00123] Transfection experiments can be conducted in relevant cell lines, using engineered RNA transcripts and protein production can be assayed at various time points post-transfection. For example, cells can be transfected with different microRNA binding site-engineering RNA transcripts and by using an ELISA kit to the relevant protein and assaying protein produced at 6 hour, 12 hour, 24 hour, 48 hour, 72 hour and 7 days post-transfection. In vivo experiments can also be conducted using microRNA-binding site-engineered molecules to examine changes in tissue-specific expression of formulated RNA transcripts.

Viral Sequences

[00124] Additional viral sequences such as, but not limited to, the translation enhancer sequence of the barley yellow dwarf virus (BYDV-PAV), the Jaagsiekte sheep retrovirus (JSRV) and/or the Enzootic nasal tumor virus (See e.g., International Pub. No. WO2012129648; herein incorporated by reference in its entirety) can be engineered and inserted in the 3' UTR of the RNA transcripts of the invention and can stimulate the translation of the construct in vitro and in vivo. Transfection experiments can be conducted in relevant cell lines at and protein production can be assayed by ELISA at 12hr, 24hr, 48hr, 72 hr and day 7 post-transfection.

IRES Sequences

[00125] Further, provided are RNA transcripts which may contain an internal ribosome entry site (IRES). First identified as a feature Picorna virus RNA, IRES plays an important role in initiating protein synthesis in absence of the 5' cap structure. An IRES may act as the sole ribosome binding site, or may serve as one of multiple ribosome binding sites of an mRNA. RNA transcripts containing more than one functional ribosome binding site may encode several peptides or polypeptides that are translated independently by the ribosomes ("multicistronic nucleic acid molecules"). When RNA transcripts are provided with an IRES, further optionally provided is a second translatable region. Examples of IRES sequences that can be used according to the invention include without limitation, those from picornaviruses (e.g. FMDV), pest viruses (CFFV), polio viruses (PV), encephalomyocarditis viruses (ECMV), foot-and-mouth disease viruses (FMDV), hepatitis C viruses (HCV), classical swine fever viruses (CSFV), murine leukemia virus (MLV), simian immune deficiency viruses (SIV) or cricket paralysis viruses (CrPV).

Poly-A tails

[00126] During RNA processing, a long chain of adenine nucleotides (poly-A tail) may be added to a polynucleotide such as an mRNA molecules in order to increase stability. Immediately after transcription, the 3' end of the transcript may be cleaved to free a 3' hydroxyl. Then poly-A polymerase adds a chain of adenine nucleotides to the RNA. The process, called polyadenylation, adds a poly-A tail that can be between, for example, approximately 100 and 250 residues long.

[00127] It has been discovered that unique poly-A tail lengths provide certain advantages to the RNA transcripts of the present invention.

[00128] In one embodiment the poly A tail is 5-300 nucleotides in length. In another embodiment, the poly A tail is 60-160 nucleotides in length.

[00129] In another embodiment, the poly-A tail is greater than 35 nucleotides in length (e.g., at least or greater than about 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, and 3,000 nucleotides). In some embodiments, the polynucleotide, primary construct, or mmRNA includes from about 30 to about 3,000 nucleotides (e.g., from 30 to 50, from 30 to 100, from 30 to 250, from 30 to 500, from 30 to 750, from 30 to 1,000, from 30 to 1,500, from 30 to 2,000, from 30 to 2,500, from 50 to 100, from 50 to 250, from 50 to 500, from 50 to 750, from 50 to 1,000, from 50 to 1,500, from 50 to 2,000, from 50 to 2,500, from 50 to 3,000, from 100 to 500, from 100 to 750, from 100 to 1,000, from 100 to 1,500, from 100 to 2,000, from 100 to 2,500, from 100 to 3,000, from 500 to 750, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to 2,500, from 500 to 3,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 2,500, from 1,000 to 3,000, from 1,500 to 2,000, from 1,500 to 2,500, from 1,500 to 3,000, from 2,000 to 3,000, from 2,000 to 2,500, and from 2,500 to 3,000).

[00130] In one embodiment, the poly-A tail is designed relative to the length of the overall RNA transcripts. This design may be based on the length of the coding region, the length of a particular feature or region (such as the first or flanking regions), or based on the length of the ultimate product expressed from the RNA transcripts.

[00131] In this context the poly-A tail may be 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100% greater in length than the RNA transcripts or feature thereof. The poly-A tail may also be designed as a fraction of RNA transcripts to which it belongs. In this context, the poly-A tail may be 10, 20, 30, 40, 50, 60, 70, 80, or 90% or more of the total length of the construct or the total length of the construct minus the poly-A tail. Further, engineered binding sites and conjugation of RNA transcripts for Poly-A binding protein may enhance expression.

[00132] Additionally, multiple distinct RNA transcripts may be linked together to the PABP (Poly-A binding protein) through the 3'-end using modified nucleotides at the 3'-terminus of the poly-A tail. Transfection experiments can be conducted in relevant cell lines at and protein production can be assayed by ELISA at 12hr, 24hr, 48hr, 72 hr and day 7 post-transfection.

[00133] In one embodiment, the polynucleotide primary constructs of the present invention are designed to include a polyA-G Quartet. The G-quartet is a cyclic hydrogen bonded array of four guanine nucleotides that can be formed by G-rich sequences in both DNA and RNA. In this embodiment, the G-quartet is incorporated at the end of the poly-A tail. The resultant mmRNA construct is assayed for stability, protein production and other parameters including half-life at various time points. It has been discovered that the polyA-G quartet results in protein production equivalent to at least 75% of that seen using a poly-A tail of 120 nucleotides alone.

Modified nucleotides

[00134] Herein, in a polynucleotide (such as a primary construct or an 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)

[00135] The modifications may be various distinct modifications. In some embodiments, 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, primary construct, or mmRNA introduced to a cell may exhibit reduced degradation in the cell, as compared to an unmodified polynucleotide, primary construct, or mmRNA.

[00136] The polynucleotides, primary constructs, and mmRNA 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). One or more atoms of a pyrimidine nucleobase 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), threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs) or hybrids thereof). Additional modifications are described herein.

[00137] As described herein, the polynucleotides, primary constructs, and mmRNA of the invention do not substantially induce an innate immune response of a cell into which the mRNA is introduced. Featues 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.

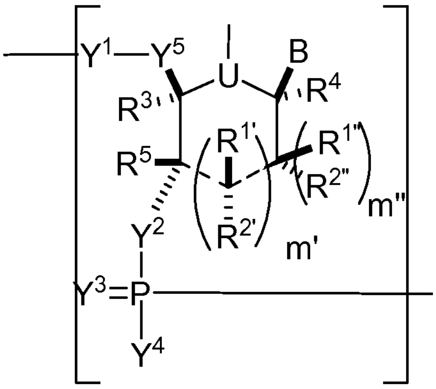
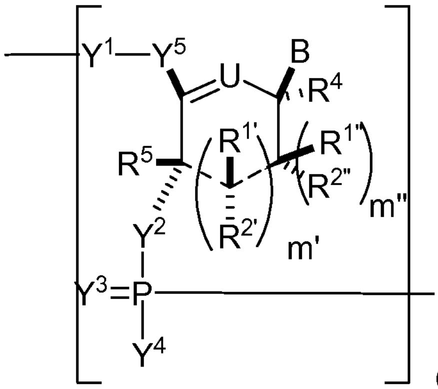
[00138] In certain embodiments, it may desirable to intracellularly degrade a modified nucleic acid molecule introduced into the cell. For example, degradation of a modified nucleic acid molecule may be preferable if precise timing of protein production is desired. Thus, in some embodiments, the invention 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).

[00139] The polynucleotides, primary constructs, and mmRNA 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, primary constructs, or mmRNA may include one or more messenger RNAs (mRNAs) and one or more modified nucleoside or nucleotides (e.g., mmRNA molecules). Details for these polynucleotides, primary constructs, and mmRNA follow.

Polynucleotides and Primary Constructs

[00140] The polynucleotides, primary constructs, and mmRNA 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.

[00141] In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (la) or Formula (Ia-1):


(Ia)

(Ia-1)

or a pharmaceutically acceptable salt or stereoisomer thereof,

[00142] wherein

[00143] U is O, S, N(RU)nu, or C(RU)nu, wherein nu is an integer from 0 to 2 and each RU is, independently, H, halo, or optionally substituted alkyl;

[00144] — is a single bond or absent;

[00145] each of R1', R2', Rl", R2", Rl, R2, R3, R4, and R5 is, independently, if present, 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 R3 with one or more of R1', Rl", R2', R2", or R5 (e.g., the combination of R1' and R3, the combination of Rl" and R3, the combination of R2' and R3, the combination of R2" and R3, or the combination of R5 and R3) 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 R5 with one or more of R1', Rl", R2', or R2" (e.g., the combination of R1' and R5, the combination of Rl" and R5, the combination of R2' and R5, or the combination of R2" and R5) 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 R4 and one or more of R1', Rl", R2', R2", R3, or R5 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);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);

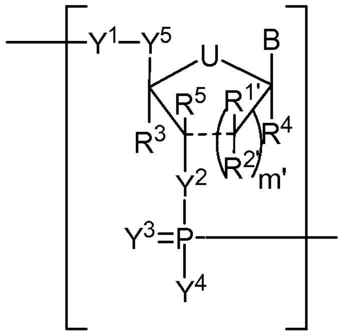
[00146] each of Yl, Y2, and Y3, is, independently, O, S, Se, -NRN1-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein RN1 is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, or absent;

[00147] each Y4 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;

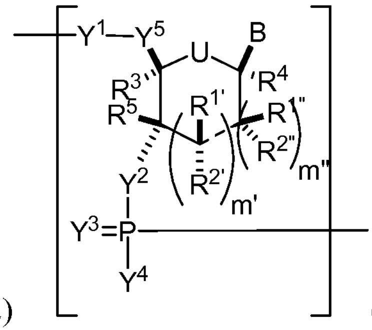
[00148] each Y5 is, independently, O, S, Se, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

[00149] n is an integer from 1 to 100,000; and

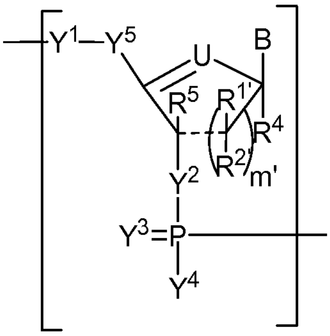
[00150] B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof), wherein the combination of B and Rl', the combination of B and R2', the combination of B and Rl", or the combination of B and R2" can, taken together with the carbons to which they are attached, optionally form a bicyclic group (e.g., a bicyclic heterocyclyl) or wherein the combination of B, Rl", and R3 or the combination of B, R2", and R3 can optionally form a tricyclic or tetracyclic group (e.g., a tricyclic or tetracyclic heterocyclyl, such as in Formula (IIo)-(IIp) herein).In some embodiments, the polynucleotide, primary construct, or mmRNA includes a modified ribose. In some embodiments, the polynucleotide, primary construct, or mmRNA (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.



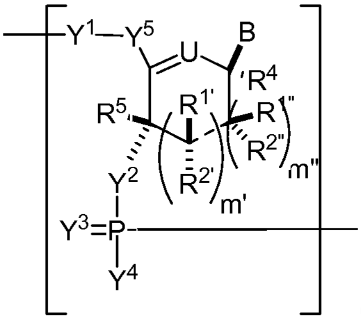
(Ia-2)



(Ia-3)

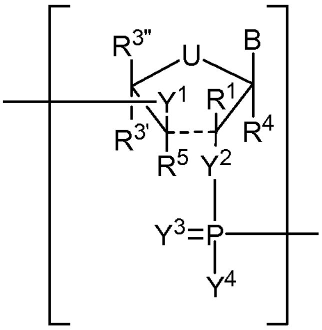


(Ia-4),

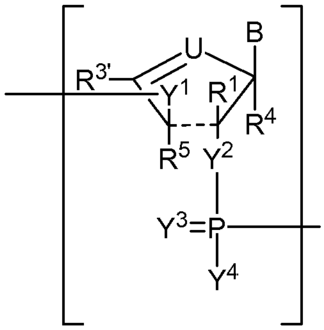


(Ia-5).

[00151] In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (lb) or Formula (Ib-1):



(Ib),



(Ib-1)

[00152] or a pharmaceutically acceptable salt or stereoisomer thereof,

[00153] wherein

[00154] U is O, S, N(RU)nu, or C(RU)nu, wherein nu is an integer from **0** to 2 and each RU is, independently, H, halo, or optionally substituted alkyl;

[00155] — is a single bond or absent;

[00156] each of Rl, R3', R3", and R4 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 Rl and R3' or the combination of Rl and R3" can be taken together to form optionally substituted alkylene or optionally substituted heteroalkylene (e.g., to produce a locked nucleic acid);

[00157] each R5 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;

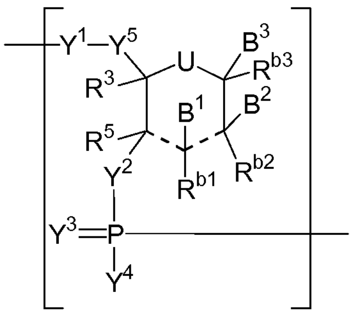
[00158] each of Yl, Y2, and Y3 is, independently, O, S, Se, -NRN1-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein RN1 is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl;

[00159] each Y4 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;

[00160] n is an integer from 1 to 100,000; and

[00161] B is a nucleobase.

[00162] In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (Ic):



(Ic),

[00163] or a pharmaceutically acceptable salt or stereoisomer thereof,

[00164] wherein

[00165] U is O, S, N(RU)nu, or C(RU)nu, wherein nu is an integer from 0 to 2 and each RU is, independently, H, halo, or optionally substituted alkyl;

[00166] — is a single bond or absent;

[00167] each of Bl, B2, and B3 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 Bl, B2, and B3 is a nucleobase;

[00168] each of Rbl, Rb2, Rb3, R3, and R5 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;

[00169] each of Yl, Y2, and Y3, is, independently, O, S, Se, -NRN1-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein RN1 is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl;

[00170] each Y4 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;

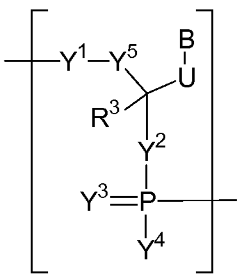
[00171] each Y5 is, independently, O, S, Se, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

[00172] n is an integer from 1 to 100,000; and

[00173] wherein the ring including U can include one or more double bonds.

[00174] In particular embodiments, the ring including U does not have a double bond between U-CB3Rb3 or between CB3Rb3-CB2Rb2.

[00175] In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (Id):



(Id),

[00176] or a pharmaceutically acceptable salt or stereoisomer thereof,

[00177] wherein

[00178] U is O, S, N(RU)nu, or C(RU)nu, wherein nu is an integer from 0 to 2 and each RU is, independently, H, halo, or optionally substituted alkyl;

[00179] each R3 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;

[00180] each of Yl, Y2, and Y3, is, independently, O, S, Se, -NRN1-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein RN1 is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl;

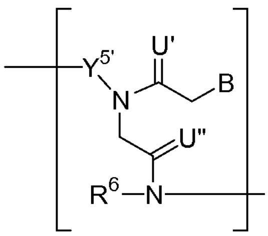
[00181] each Y4 is, independently, H, hydroxy, thiol, boranyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino;

[00182] each Y5 is, independently, O, S, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

[00183] n is an integer from 1 to 100,000; and

[00184] B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof).

[00185] In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, first flanking region, or second flanking region) includes n number of linked nucleosides having Formula (Ie):



(Ie),

[00186] or a pharmaceutically acceptable salt or stereoisomer thereof,

[00187] wherein

[00188] each of U' and U" is, independently, O, S, N(RU)nu, or C(RU)nu, wherein nu is an integer from 0 to 2 and each RU is, independently, H, halo, or optionally substituted alkyl;

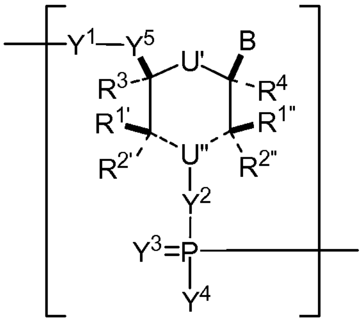
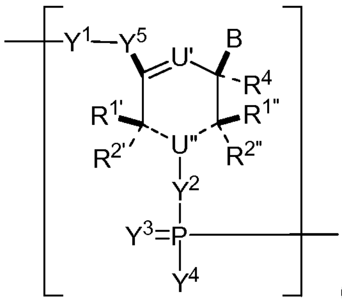
[00189] each R6 is, independently, H, halo, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, optionally substituted hydroxyalkoxy, optionally substituted amino, azido, optionally substituted aryl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, or optionally substituted aminoalkynyl;

[00190] each Y5' is, independently, O, S, optionally substituted alkylene (e.g., methylene or ethylene), or optionally substituted heteroalkylene;

[00191] n is an integer from 1 to 100,000; and

[00192] B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof).

[00193] In some embodiments, the polynucleotide, primary construct, or mmRNA (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),

[00194] or a pharmaceutically acceptable salt or stereoisomer thereof,

[00195] wherein

[00196] each of U' and U" is, independently, O, S, N, N(RU)nu, or C(RU)nu, wherein nu is an integer from 0 to 2 and each RU is, independently, H, halo, or optionally substituted alkyl (e.g., U' is O and U" is N);

[00197] ── is a single bond or absent;

[00198] each of Rl', R2′, Rl", R2", R3, and R4 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 Rl' and R3, the combination of Rl" and R3, the combination of R2' and R3, or the combination of R2" and R3 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);

[00199] each of Yl, Y2, and Y3, is, independently, O, S, Se, -NRN1-, optionally substituted alkylene, or optionally substituted heteroalkylene, wherein RN1 is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aryl, or absent;

[00200] each Y4 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;

[00201] each Y5 is, independently, O, S, Se, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene;

[00202] n is an integer from 1 to 100,000; and

[00203] B is a nucleobase (e.g., a purine, a pyrimidine, or derivatives thereof).

[00204] In some embodiments of the polynucleotides, primary constructs, or mmRNA (e.g., Formulas (la), (Ia-l)-(Ia-3), (Ib)-(If), and (Ila)-(IIp)), the ring including U has one or two double bonds.

[00205] In some embodiments of the polynucleotides, primary constructs, or mmRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), each of Rl, Rl', and Rl", if present, is H. In further embodiments, each of R2, R2', and R2", 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 -(CH2)s2(OCH2CH2)sl(CH2)s30R\ wherein si 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 CI-20 alkyl). In some embodiments, s2 is 0, si is 1 or 2, s3 is 0 or 1, and R' is CI-6 alkyl.

[00206] In some embodiments of the polynucleotides, primary constructs, or mmRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), each of R2, R2′, and R2", if present, is H. In further embodiments, each of Rl, Rl', and Rl", 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 -(CH2)s2(OCH2CH2)sl(CH2)s30R\ wherein si 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 CI-20 alkyl). In some embodiments, s2 is 0, si is 1 or 2, s3 is 0 or 1, and R' is CI-6 alkyl.

[00207] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), each of R3, R4, and R5 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, R3 is H, R4 is H, R5 is H, or R3, R4, and R5 are all H. In particular embodiments, R3 is CI-6 alkyl, R4 is CI-6 alkyl, R5 is CI-6 alkyl, or R3, R4, and R5 are all CI-6 alkyl. In particular embodiments, R3 and R4 are both H, and R5 is CI-6 alkyl.

[00208] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), R3 and R5 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 R3 and R5 join together to form heteroalkylene (e.g., -(CH2)blO(CH2)b20(CH2)b3-, wherein each of bl, b2, and b3 are, independently, an integer from 0 to 3).

[00209] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), R3 and one or more of Rl', Rl", R2', R2", or R5 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, R3 and one or more of Rl', Rl", R2', R2", or R5 join together to form heteroalkylene (e.g., -(CH2)blO(CH2)b20(CH2)b3-, wherein each of bl, b2, and b3 are, independently, an integer from 0 to 3).

[00210] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), R5 and one or more of Rl', Rl", R2', or R2" 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, R5 and one or more of R1', Rl", R2', or R2" join together to form heteroalkylene (e.g., -(CH2)blO(CH2)b20(CH2)b3-, wherein each of bl, b2, and b3 are, independently, an integer from 0 to 3).

[00211] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), each Y2 is, independently, O, S, or -NRN1-, wherein RNl is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl. In particular embodiments, Y2 is NRN1-, wherein RNl is H or optionally substituted alkyl (e.g., Cl-6 alkyl, such as methyl, ethyl, isopropyl, or n-propyl).

[00212] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), each Y3 is, independently, O or S.

[00213] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), R1 is H; each R2 is, independently, H, halo (e.g., fluoro), hydroxy, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy (e.g., -(CH2)s2(OCH2CH2)sl(CH2)s30R', wherein si 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 CI-20 alkyl, such as wherein s2 is 0, si is 1 or 2, s3 is 0 or 1, and R' is Cl-6 alkyl); each Y2 is, independently, O or -NRN1-, wherein RNl is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein RNl is H or optionally substituted alkyl (e.g., Cl-6 alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each Y3 is, independently, O or S (e.g., S). In further embodiments, R3 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 Y1 is , independently, O or -NRN1-, wherein RNl is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein RNl is H or optionally substituted alkyl (e.g., Cl-6 alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each Y4 is, independently, H, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino.

[00214] In some embodiments of the polynucleotides, primary constructs, or mmRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), each R1 is, independently, H, halo (e.g., fluoro), hydroxy, optionally substituted alkoxy (e.g., methoxy or ethoxy), or optionally substituted alkoxyalkoxy (e.g., -(CH2)s2(OCH2CH2)sl(CH2)s30R\ wherein si 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 CI-20 alkyl, such as wherein s2 is 0, si is 1 or 2, s3 is 0 or 1, and R' is Cl-6 alkyl); R2 is H; each Y2 is, independently, O or -NRN1-, wherein RNl is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein RNl is H or optionally substituted alkyl (e.g., Cl-6 alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each Y3 is, independently, O or S (e.g., S). In further embodiments, R3 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 Y1 is , independently, O or -NRN1-, wherein RNl is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl (e.g., wherein RNl is H or optionally substituted alkyl (e.g., Cl-6 alkyl, such as methyl, ethyl, isopropyl, or n-propyl)); and each Y4 is, independently, H, hydroxy, thiol, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted thioalkoxy, optionally substituted alkoxyalkoxy, or optionally substituted amino.

[00215] In some embodiments of the polynucleotides, primary constructs, or mmRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), the ring including U is in the β-D (e.g., β-D-ribo) configuration.

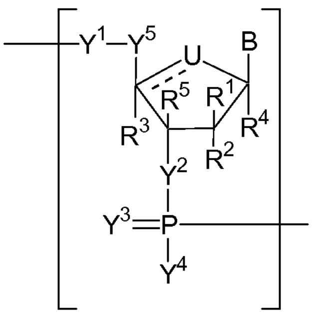
[00216] In some embodiments of the polynucleotides, primary constructs, or mmRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), the ring including U is in the a-L (e.g., a-L-ribo) configuration.

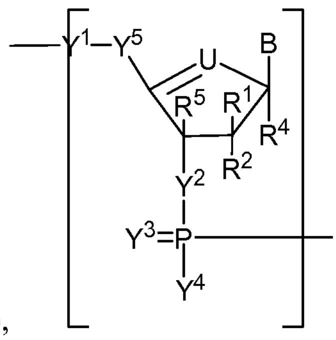
[00217] In some embodiments of the polynucleotides, primary constructs, or mmRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-

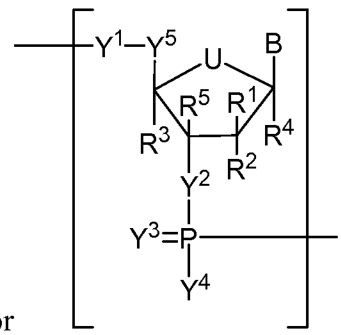
2), (IVa)-(IVl), and (IXa)-(IXr)), one or more B is not pseudouridine (ψ) or 5-methyl-cytidine (m5C). In some embodiments, about 10% to about *100%* of n number of B nucleobases is not ψ or m5C (e.g., from 10% to 20%, from 10% to 35%, from 10% to 50%, from 10% to 60%, from 10% to 75%, from 10% to 90%, 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 m5C). In some embodiments, B is not ψ or m5C.

[00218] In some embodiments of the polynucleotides, primary constructs, or mrnRNA (e.g., Formulas (Ia)-(Ia-5), (Ib)-(If-l), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(IIc-2), (IIn-1), (Iln-2), (IVa)-(IVl), and (IXa)-(IXr)), when B is an unmodified nucleobase selected from cytosine, guanine, uracil and adenine, then at least one of Yl, Y2, or Y3 is not O.

[00219] In some embodiments, the polynucleotide, primary construct, or mmRNA includes a modified ribose. In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (Ila)-(IIc):



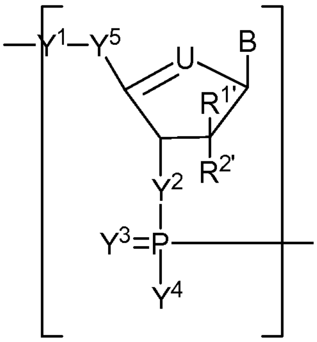
(IIa),


(IIb), or


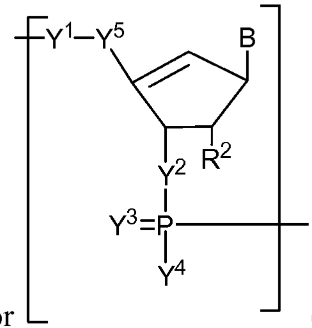
(IIc),

[00220] or a pharmaceutically acceptable salt or stereoisomer thereof. In particular embodiments, U is O or C(RU)nu, wherein nu is an integer from 0 to 2 and each RU is, independently, H, halo, or optionally substituted alkyl (e.g., U is -CH2- or -CH-). In other embodiments, each of Rl, R2, R3, R4, and R5 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 R1 and R2 is, independently, H, halo, hydroxy, optionally substituted alkyl, or optionally substituted alkoxy; each R3 and R4 is, independently, H or optionally substituted alkyl; and R5 is H or hydroxy), and - - - is a single bond or double bond.

[00221] In particular embodiments, the polynucleotidesor mmRNA includes n number of linked nucleosides having Formula (lib- 1)-(IIb-2):



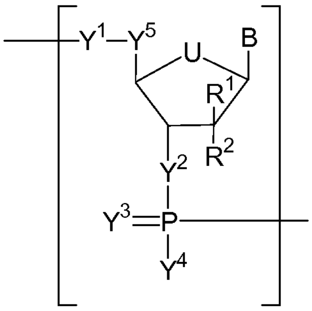
(IIb-1) or



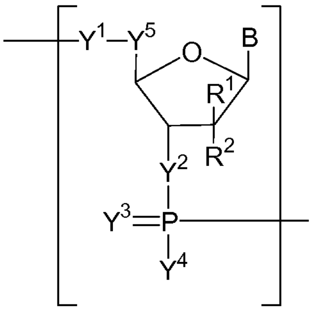
(IIb-2)

[00222] or a pharmaceutically acceptable salt or stereoisomer thereof. In some embodiments, U is O or C(RU)nu, wherein nu is an integer from 0 to 2 and each RU is, independently, H, halo, or optionally substituted alkyl (e.g., U is -CH2- or -CH-). In other embodiments, each of R1 and R2 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 R1 and R2 is, independently, H, halo, hydroxy, optionally substituted alkyl, or optionally substituted alkoxy, e.g., H, halo, hydroxy, alkyl, or alkoxy). In particular embodiments, R2 is hydroxy or optionally substituted alkoxy (e.g., methoxy, ethoxy, or any described herein).

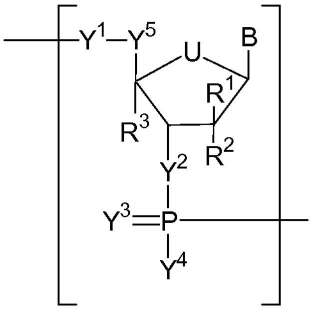
[00223] In particular embodiments, the polynucleotide, primary construct, or mmRNA includes n number of linked nucleosides having Formula (IIc-l)-(IIc-4):



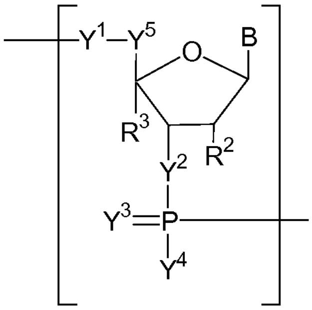
(IIc-1),



(IIc-2),



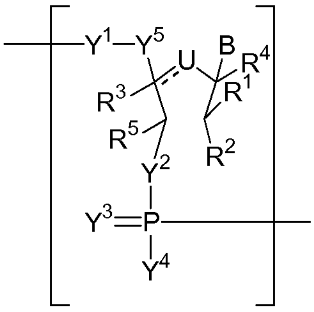
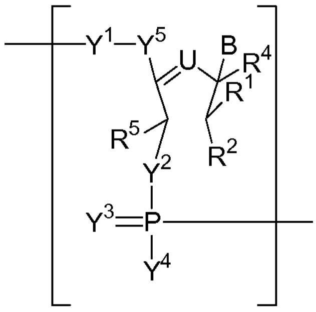
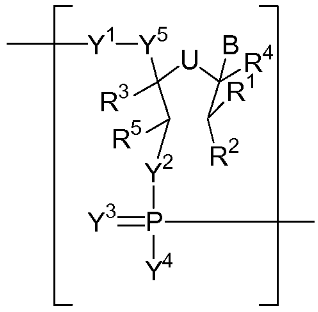
(IIc-3), or



(IIc-4),

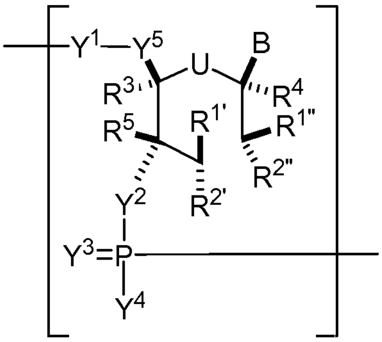
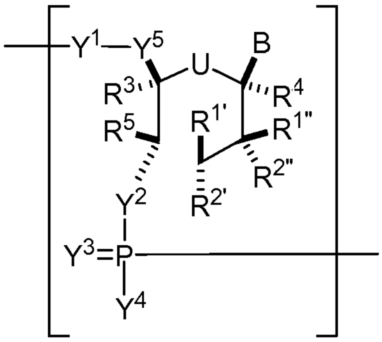
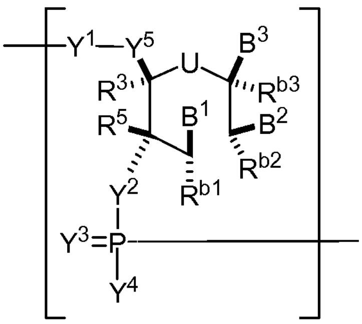
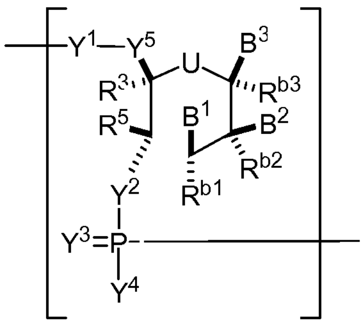
[00224] or a pharmaceutically acceptable salt or stereoisomer thereof. In some embodiments, U is O or C(RU)nu, wherein nu is an integer from **0** to 2 and each RU is, independently, H, halo, or optionally substituted alkyl (e.g., U is -CH2- or -CH-). In some embodiments, each of Rl, R2, and R3 is, independently, 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 R1 and R2 is, independently, H, halo, hydroxy, optionally substituted alkyl, or optionally substituted alkoxy, e.g., H, halo, hydroxy, alkyl, or alkoxy; and each R3 is, independently, H or optionally substituted alkyl)). In particular embodiments, R2 is optionally substituted alkoxy (e.g., methoxy or ethoxy, or any described herein). In particular embodiments, R1 is optionally substituted alkyl, and R2 is hydroxy. In other embodiments, R1 is hydroxy, and R2 is optionally substituted alkyl. In further embodiments, R3 is optionally substituted alkyl.

[00225] In some embodiments, the polynucleotide, primary construct, or mmRNA includes an acyclic modified ribose. In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (Ild)-(IIf):


(IId),

(IIe), or

(IIf),

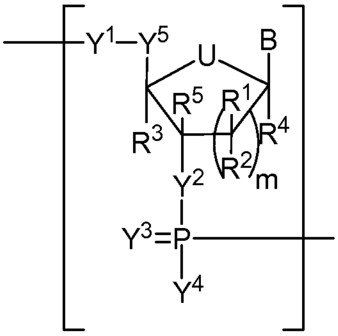
[00226] or a pharmaceutically acceptable salt or stereoisomer thereof.

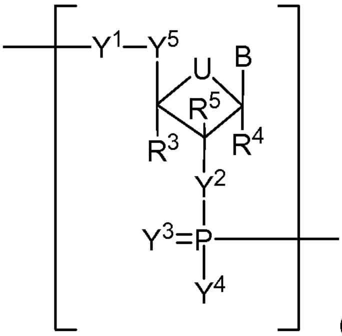
[00227] In some embodiments, the polynucleotide, primary construct, or mmRNA includes an acyclic modified hexitol. In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides Formula (Ilg)-(IIj):


(IIg),

(IIh),

(IIi), or

(IIj),

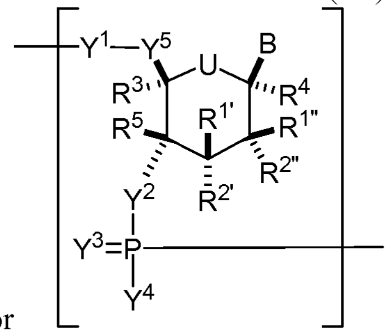
[00228] or a pharmaceutically acceptable salt or stereoisomer thereof.

[00229] In some embodiments, the polynucleotide, primary construct, or mmRNA includes a sugar moiety having a contracted or an expanded ribose ring. In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (Ilk)-(IIm):


(IIk),

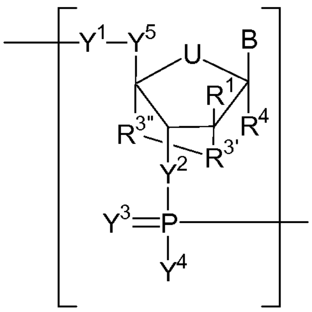

(IIl),

or


(IIm),

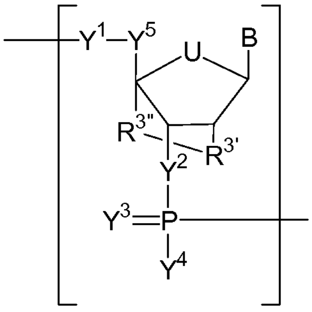
[00230] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each of Rl', Rl", R2', and R2" is, independently, H, halo, hydroxy, optionally substituted alkyl, optionally substituted alkoxy, optionally substituted alkenyloxy, optionally substituted alkynyloxy, optionally substituted aminoalkoxy, optionally substituted alkoxyalkoxy, or absent; and wherein the combination of R2' and R3 or the combination of R2" and R3 can be taken together to form optionally substituted alkylene or optionally substituted heteroalkylene.

[00231] In some embodiments, the polynucleotide, primary construct, or mmRNA includes a locked modified ribose. In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (Iln):

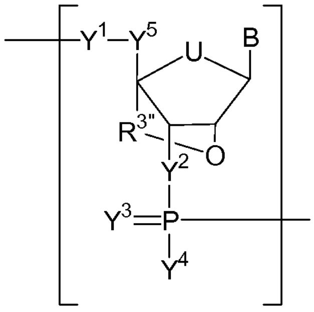

(IIn),

[00232] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein R3' is O, S, or -NRN1-, wherein RNl is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl and R3" is optionally substituted alkylene (e.g., -CH2-, -CH2CH2-, or -CH2CH2CH2-) or optionally substituted heteroalkylene (e.g., -CH2NH-, -CH2CH2NH-, -CH20CH2-, or -CH2CH20CH2-)(e.g., R3' is O and R3" is optionally substituted alkylene (e.g., -CH2-, -CH2CH2-, or -CH2CH2CH2-)).

[00233] In some embodiments, the polynucleotide, primary construct, or mmRNA includes n number of linked nucleosides having Formula (IIn-l)-(II-n2):



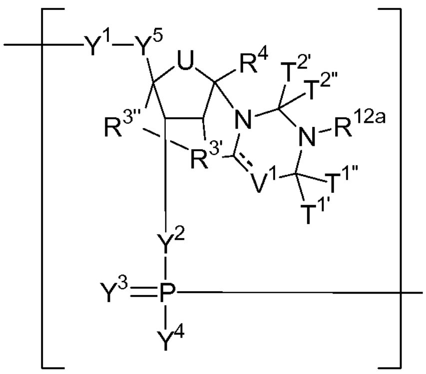
(IIn-1) or



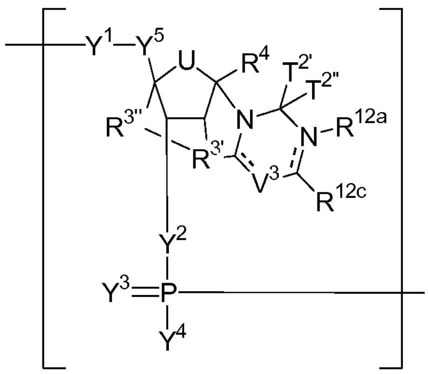
(IIn-2),

[00234] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein R3' is O, S, or -NRN1-, wherein RNl is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted aryl and R3" is optionally substituted alkylene (e.g., -CH2-, -CH2CH2-, or -CH2CH2CH2-) or optionally substituted heteroalkylene (e.g., -CH2NH-, -CH2CH2NH-, -CH20CH2-, or -CH2CH20CH2-) (e.g., R3' is O and R3" is optionally substituted alkylene (e.g., -CH2-, -CH2CH2-, or -CH2CH2CH2-)).

[00235] In some embodiments, the polynucleotide, primary construct, or mmRNA includes a locked modified ribose that forms a tetracyclic heterocyclyl. In some embodiments, the polynucleotide, primary construct, or mmRNA (e.g., the first region, the first flanking region, or the second flanking region) includes n number of linked nucleosides having Formula (Ho):



(IIo) or

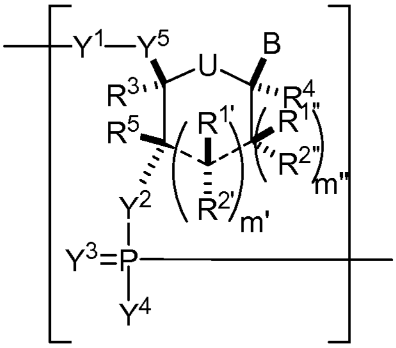


(IIp),

[00236] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein R12a, R12c, Tl', T1",T2', T2", VI, and V3 are as described herein.

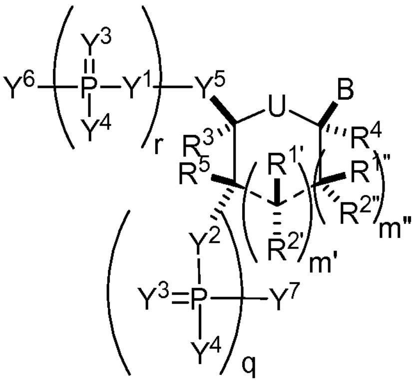
[00237] Any of the formulas for the polynucleotides, primary constructs, or mmRNA can include one or more nucleobases described herein (e.g., Formulas (bl)-(b43)).

[00238] In one embodiment, the present invention provides methods of preparing a polynucleotide, primary construct, or mmRNA, wherein the polynucleotide comprises n number of nucleosides having Formula (la), as defined herein:



(Ia),

[00239] the method comprising reacting a compound of Formula (Ilia), as defined herein:

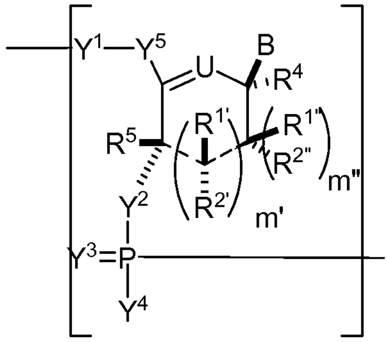


(IIIa),

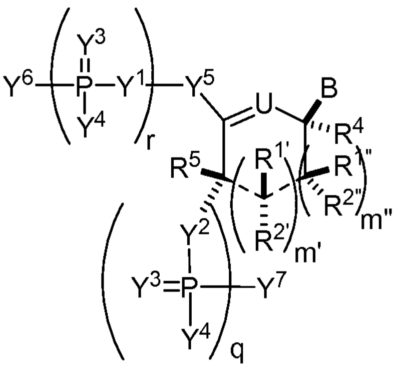
[00240] with an RNA polymerase, and a cDNA template.

[00241] In a further embodiment, the present invention provides methods of amplifying a polynucleotide, primary construct, or mmRNA comprising at least one nucleotide (e.g., mmRNA molecule), the method comprising: reacting a compound of Formula (Ilia), as defined herein, with a primer, a cDNA template, and an RNA polymerase.

[00242] In one embodiment, the present invention provides methods of preparing a polynucleotide, primary construct, or mmRNA comprising at least one nucleotide (e.g., mmRNA molecule), wherein the polynucleotide comprises n number of nucleosides having Formula (la), as defined herein:


(Ia-1),

[00243] the method comprising reacting a compound of Formula (IIIa-1), as defined herein:

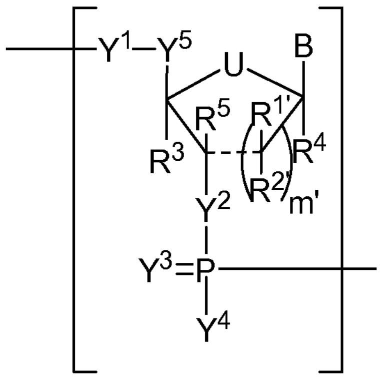

(IIIa-1),

[00244] with an RNA polymerase, and a cDNA template.

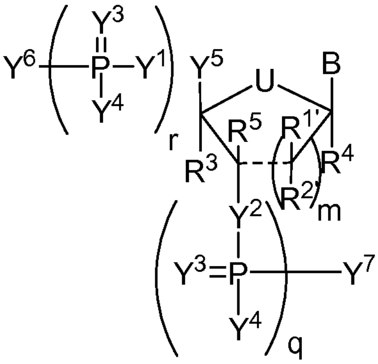
[00245] In a further embodiment, the present invention provides methods of amplifying a polynucleotide, primary construct, or mmRNA comprising at least one nucleotide (e.g., mmRNA molecule), the method comprising:

[00246] reacting a compound of Formula (IIIa-1), as defined herein, with a primer, a cDNA template, and an RNA polymerase.

[00247] In one embodiment, the present invention provides methods of preparing a modified mRNA comprising at least one nucleotide (e.g., mmRNA molecule), wherein the polynucleotide comprises n number of nucleosides having Formula (Ia-2), as defined herein:


(Ia-2),

[00248] the method comprising reacting a compound of Formula (IIIa-2), as defined herein:



(IIIa-2),

[00249] with an RNA polymerase, and a cDNA template.

[00250] In a further embodiment, the present invention provides methods of amplifying a modified mRNA comprising at least one nucleotide (e.g., mmRNA molecule), the method comprising: reacting a compound of Formula (IIIa-2), as defined herein, with a primer, a cDNA template, and an RNA polymerase.

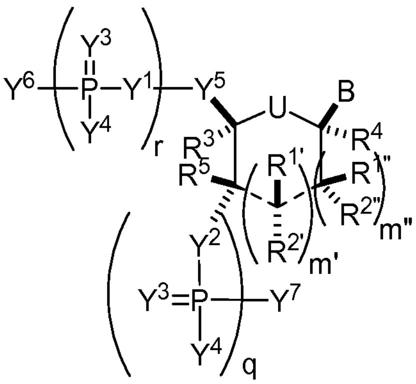
[00251] In some embodiments, the reaction may be repeated from 1 to about 7,000 times. In any of the embodiments herein, B may be a nucleobase of Formula (bl)-(b43).

[00252] The polynucleotides, primary constructs, and mmRNA can optionally include 5' and/or 3' flanking regions, which are described herein.

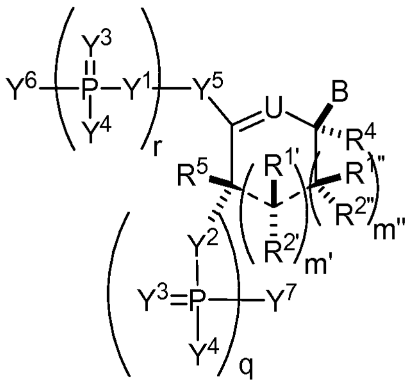
Modified RNA (mmRNA) Molecules

[00253] The present invention also includes building blocks, e.g., modified ribonucleosides, modified ribonucleotides, of modified RNA (mmRNA) molecules. For example, these building blocks can be useful for preparing the polynucleotides, primary constructs, or mmRNA of the invention.

[00254] In some embodiments, the building block molecule has Formula (Ilia) or (Ilia- 1):



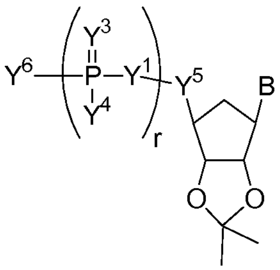
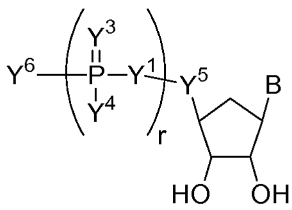
(Ilia),



(IIIa-1)

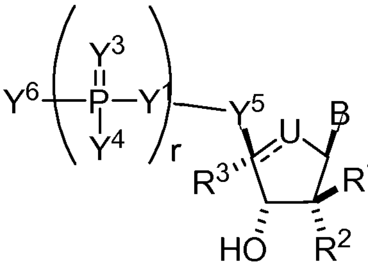
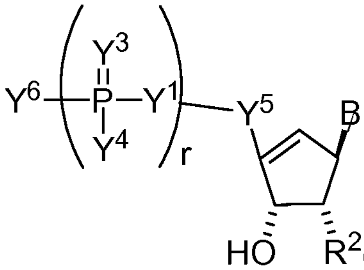
[00255] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein the substituents are as described herein (e.g., for Formula (la) and (Ia-1)), and wherein when B is an unmodified nucleobase selected from cytosine, guanine, uracil and adenine, then at least one of Yl, Y2, or Y3 is not O.

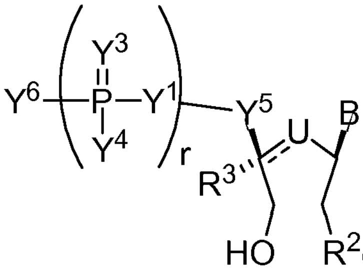
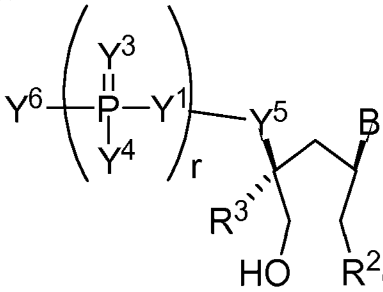
[00256] In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, has Formula (IVa)-(IVb):

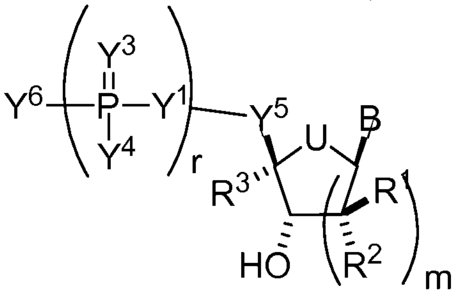
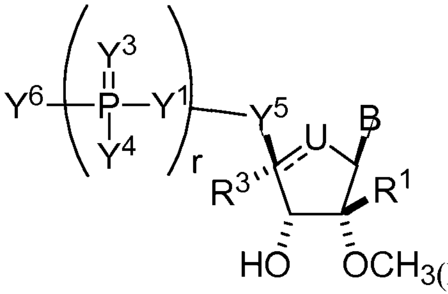

(IVa) or

(IVb),

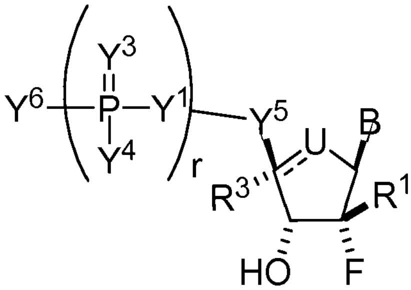
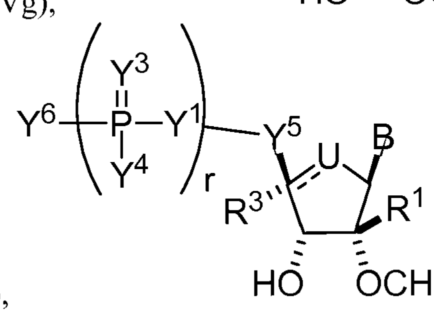
[00257] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (bl)-(b43)). In particular embodiments, Formula (IVa) or (IVb) is combined with a modified uracil (e.g., any one of formulas (bl)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (bl), (b8), (b28), (b29), or (b30)). In particular embodiments, Formula (IVa) or (IVb) is combined with a modified cytosine (e.g., any one of formulas (M0)-(bl4), (b24), (b25), and (b32)-(b36), such as formula (blO) or (b32)). In particular embodiments, Formula (IVa) or (IVb) is combined with a modified guanine (e.g., any one of formulas (bl5)-(bl7) and (b37)-(b40)). In particular embodiments, Formula (IVa) or (IVb) is combined with a modified adenine (e.g., any one of formulas (M8)-(b20) and (b41)-(b43)).

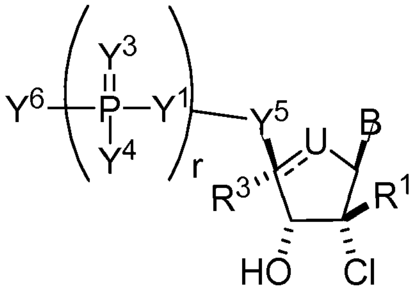
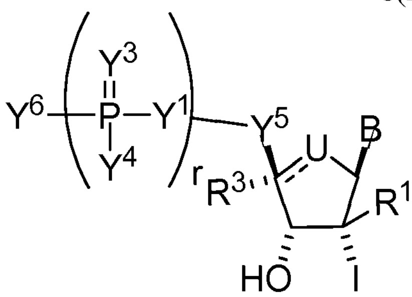
[00258] In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, has Formula (IVc)-(IVk):

 (IVc),  (IVd),

 (IVe),  (IVf),

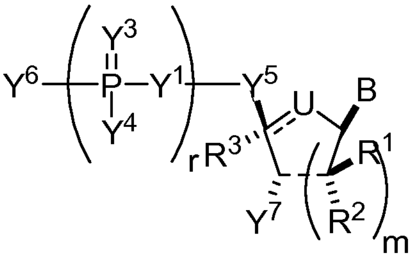
 (IVg),  (IVh),

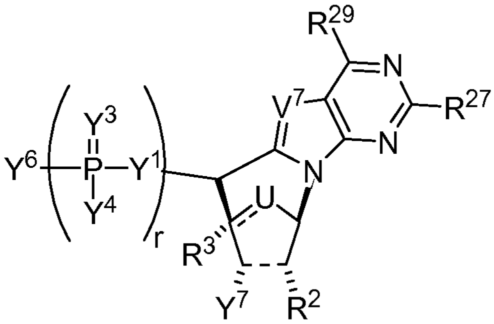
 (IVi),  (IVj),

 (IVk), or  (IVl),

[00259] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (bl)-(b43)). In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified uracil (e.g., any one of formulas (bl)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (bl), (b8), (b28), (b29), or (b30)). In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified cytosine (e.g., any one of formulas (M0)-(bl4), (b24), (b25), and (b32)-(b36), such as formula (blO) or (b32)). In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified guanine (e.g., any one of formulas (bl5)-(bl7) and (b37)-(b40)). In particular embodiments, one of Formulas (IVc)-(IVk) is combined with a modified adenine (e.g., any one of formulas (M8)-(b20) and (b41)-(b43)).

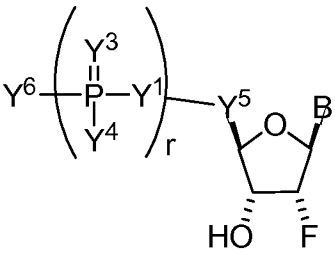
[00260] In other embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, has Formula (Va) or (Vb):

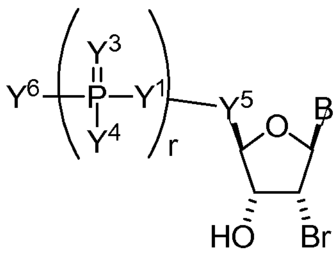

(Va) or

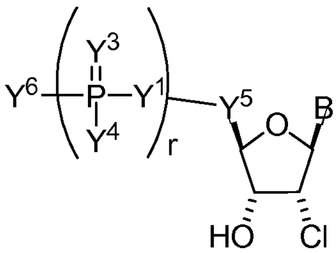

(Vb),

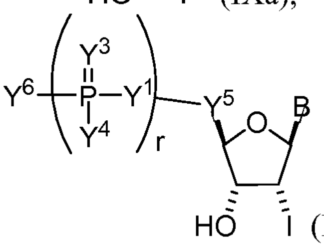
[00261] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (bl)-(b43)).

[00262] In other embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, has Formula (IXa)-(IXd):


(IXa),

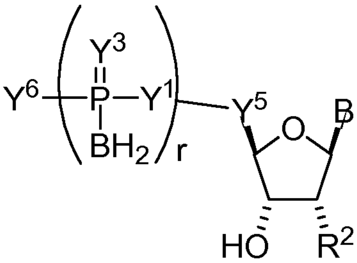

(IXb),

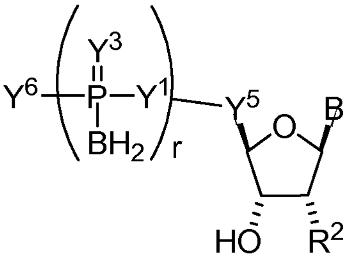

(IXc), or

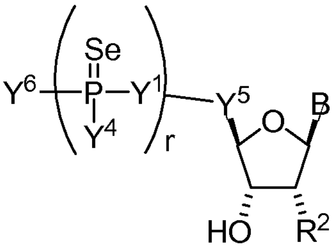

(IXd),

[00263] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (bl)-(b43)). In particular embodiments, one of Formulas (IXa)-(IXd) is combined with a modified uracil (e.g., any one of formulas (bl)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (bl), (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 (M0)-(bl4), (b24), (b25), and (b32)-(b36), such as formula (blO) or (b32)). In particular embodiments, one of Formulas (IXa)-(IXd) is combined with a modified guanine (e.g., any one of formulas (bl5)-(bl7) and (b37)-(b40)). In particular embodiments, one of Formulas (IXa)-(IXd) is combined with a modified adenine (e.g., any one of formulas (M8)-(b20) and (b41)-(b43)).

[00264] In other embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, has Formula (IXe)-(IXg):

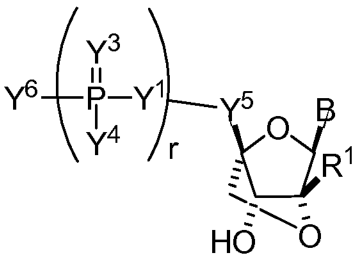
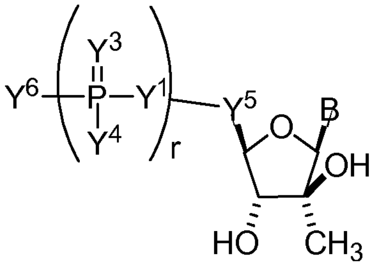

(IXe),

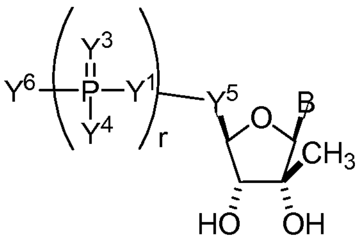
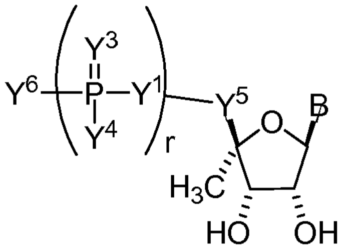

(IXf), or


(IXg),

[00265] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (bl)-(b43)). In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified uracil (e.g., any one of formulas (bl)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (bl), (b8), (b28), (b29), or (b30)). In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified cytosine (e.g., any one of formulas (M0)-(bl4), (b24), (b25), and (b32)-(b36), such as formula (blO) or (b32)). In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified guanine (e.g., any one of formulas (bl5)-(bl7) and (b37)-(b40)). In particular embodiments, one of Formulas (IXe)-(IXg) is combined with a modified adenine (e.g., any one of formulas (M8)-(b20) and (b41)-(b43)).

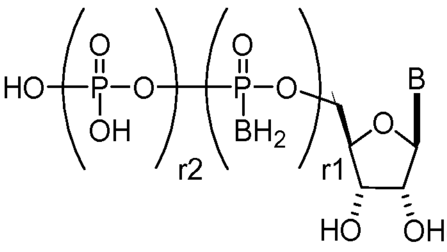
[00266] In other embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, has Formula (IXh)-(IXk):

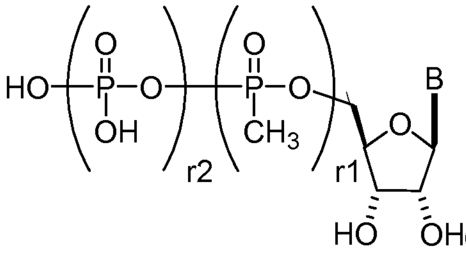
 (IXh),  (IXi),

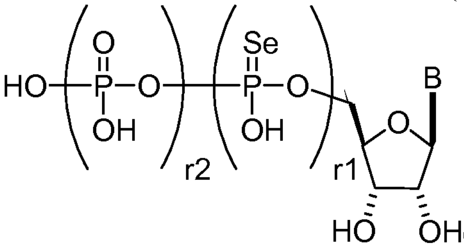
 (IXj), or (IXk),

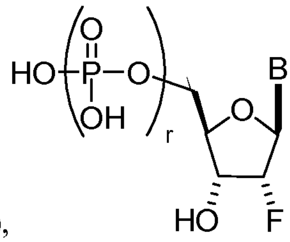
[00267] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein B is as described herein (e.g., any one of (bl)-(b43)). In particular embodiments, one of Formulas (IXh)-(IXk) is combined with a modified uracil (e.g., any one of formulas (bl)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (bl), (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 (M0)-(bl4), (b24), (b25), and (b32)-(b36), such as formula (blO) or (b32)). In particular embodiments, one of Formulas (IXh)-(IXk) is combined with a modified guanine (e.g., any one of formulas (bl5)-(bl7) and (b37)-(b40)). In particular embodiments, one of Formulas (IXh)-(IXk) is combined with a modified adenine (e.g., any one of formulas (M8)-(b20) and (b41)-(b43)).

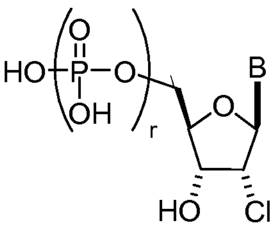
[00268] In other embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, has Formula (IXl)-(IXr):

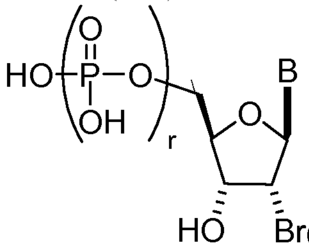
 (IXl),

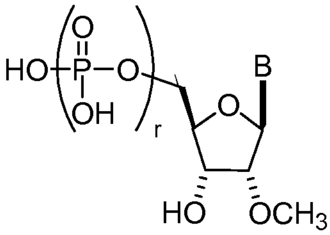
 (IXm),

 (IXn),

 (IXo),

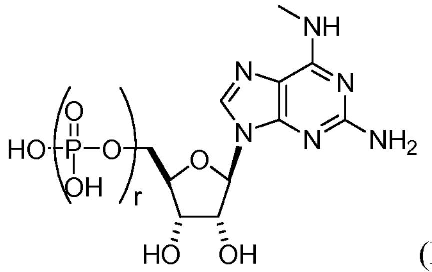
 (IXp),

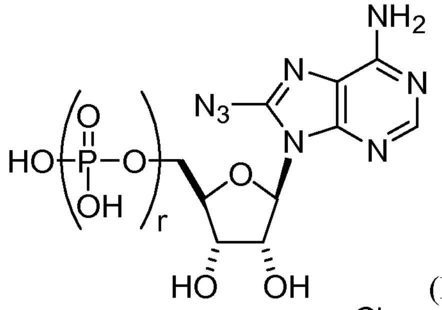
 (IXq), or

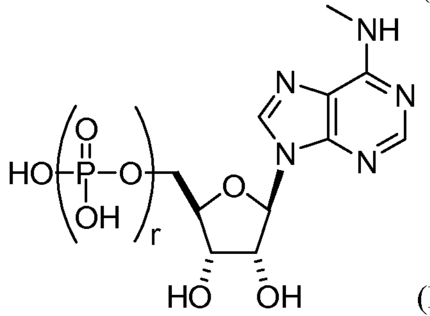
 (IXr)

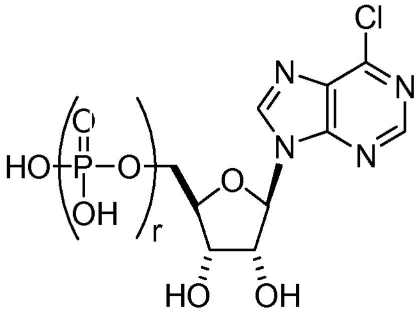
[00269] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r1 and r2 is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5) and B is as described herein (e.g., any one of (bl)-(b43)). In particular embodiments, one of Formulas (IXl)-(IXr) is combined with a modified uracil (e.g., any one of formulas (bl)-(b9), (b21)-(b23), and (b28)-(b31), such as formula (bl), (b8), (b28), (b29), or (b30)). In particular embodiments, one of Formulas (IXl)-(IXr) is combined with a modified cytosine (e.g., any one of formulas (M0)-(bl4), (b24), (b25), and (b32)-(b36), such as formula (blO) or (b32)). In particular embodiments, one of Formulas (IXl)-(IXr) is combined with a modified guanine (e.g., any one of formulas (bl5)-(bl7) and (b37)-(b40)). In particular embodiments, one of Formulas (IXl)-(IXr) is combined with a modified adenine (e.g., any one of formulas (b 18)-(b20) and (b41)-(b43)).

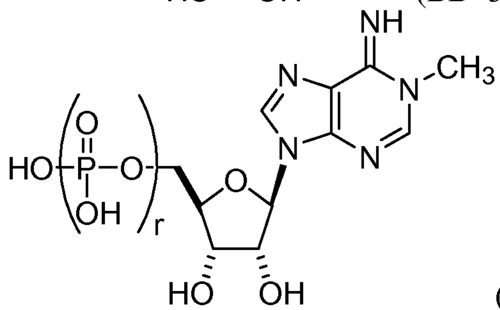
[00270] In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mrnRNA, can be selected from the group consisting of:

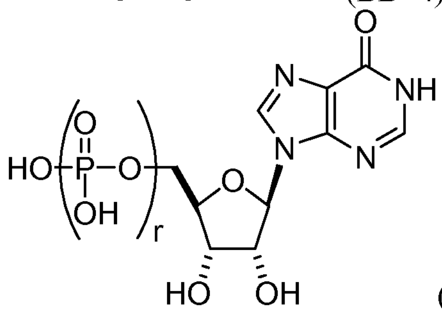

(BB- 1),

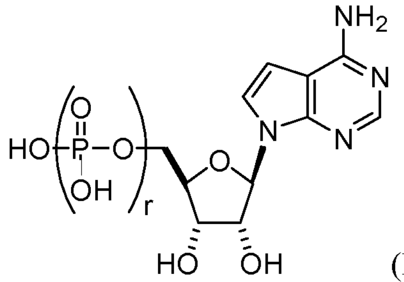

(BB- 2),

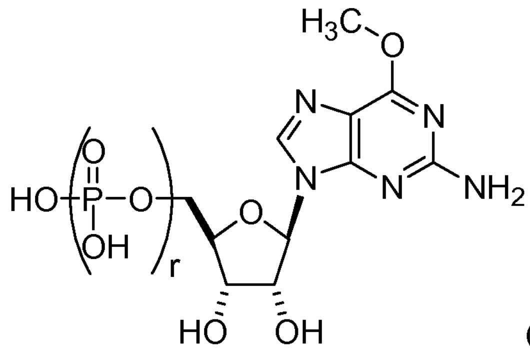

(BB- 3),

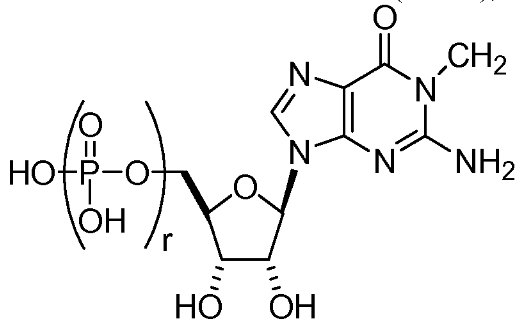

(BB- 4),

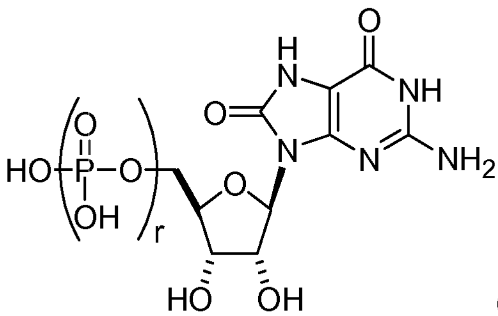

(BB- 5),

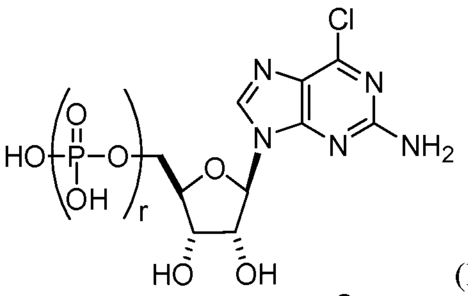

(BB- 6),


(BB- 7),

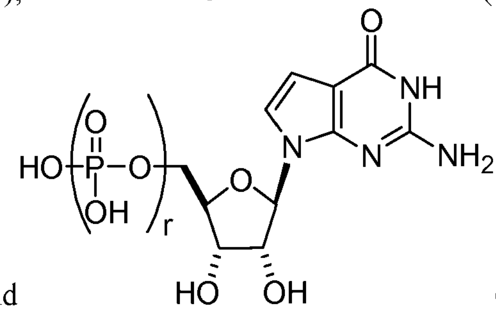

(BB- 8),


(BB- 9),


(BB-



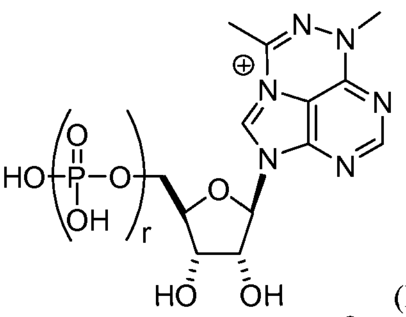
10), (BB- 11),



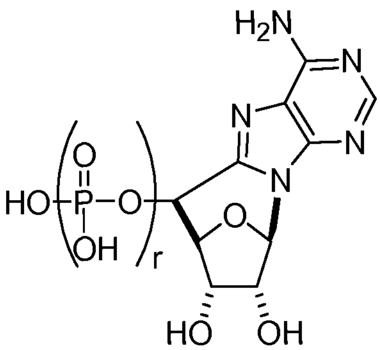
and (BB- 12),

[00271] 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).

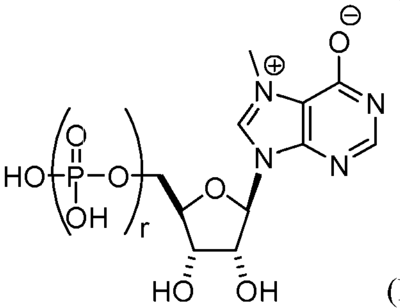
[00272] In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, can be selected from the group consisting of:



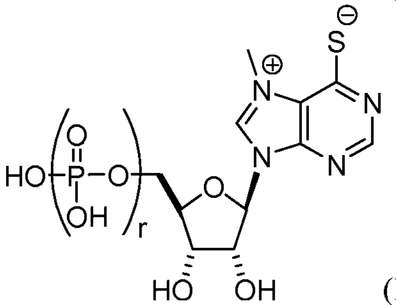
(BB- 13),



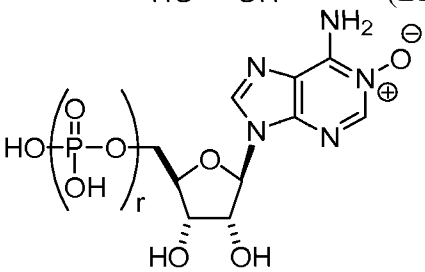
(BB- 14),



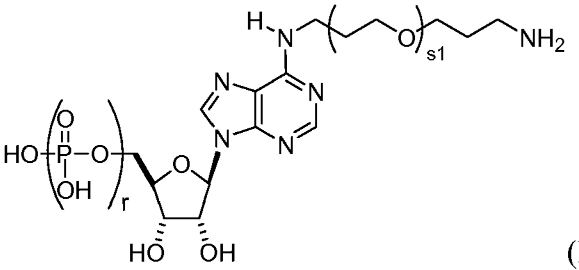
(BB- 15),

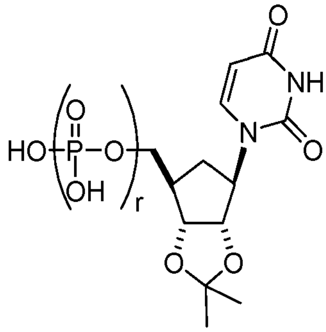


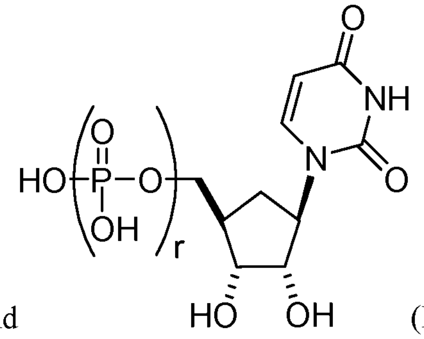
(BB- 16),



(BB- 17),

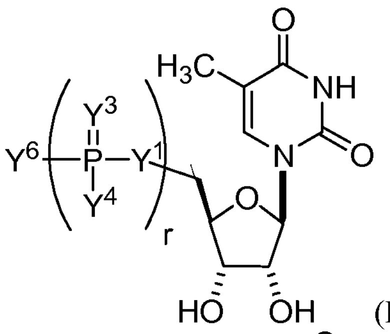

(BB- 18),

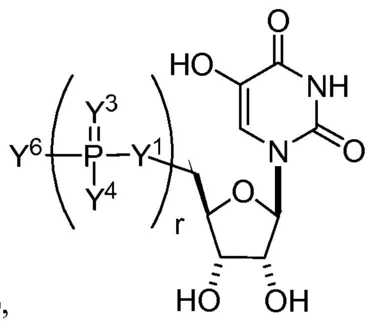

(BB- 19),

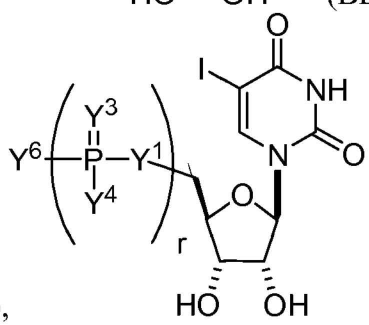
and

(BB- 20),

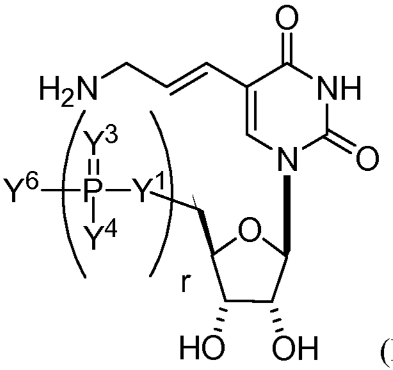
[00273] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5) and si is as described herein.

[00274] In some embodiments, the building block molecule, which may be incorporated into a nucleic acid (e.g., RNA, mRNA, polynucleotide, primary construct, or mmRNA), is a modified uridine (e.g., selected from the group consisting of:

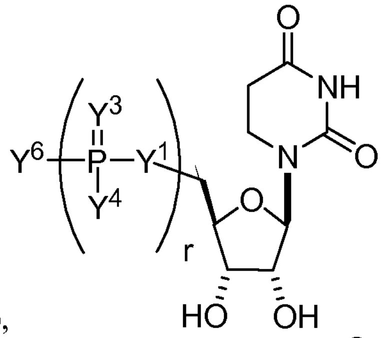

(BB- 21),


(BB- 22),


(BB- 23),

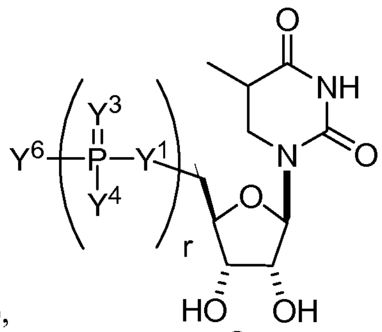


(BB- 24),

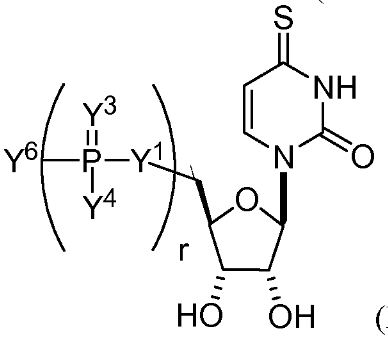


(BB-

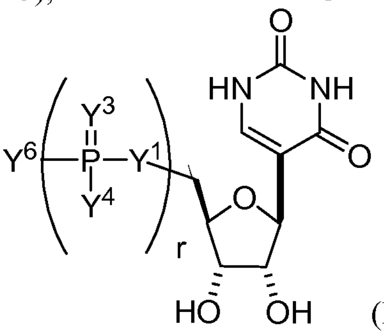
25),



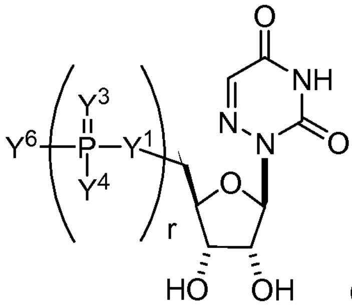
(BB- 26),



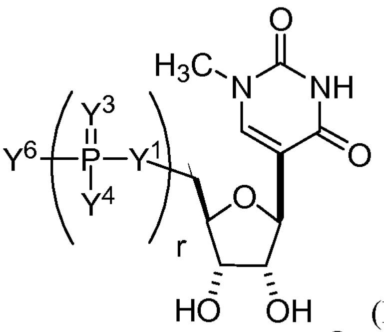
(BB- 27),



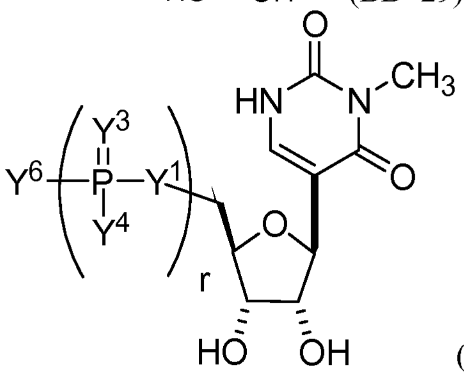
(BB- 28),



(BB- 29),

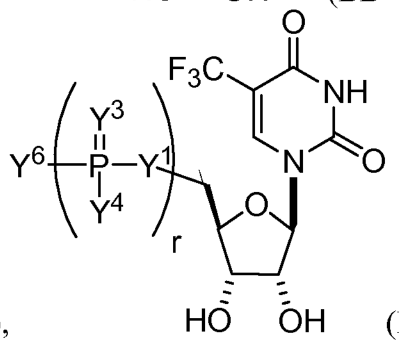


(BB- 30),

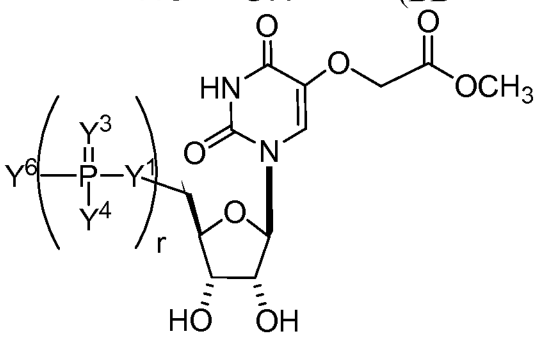


(BB-

31),

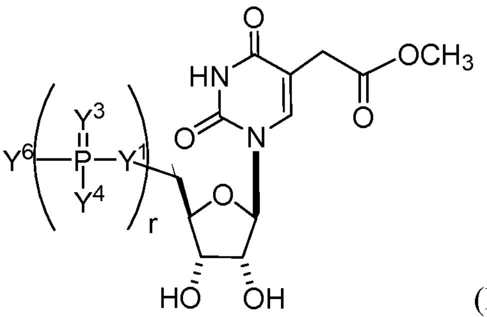


(BB- 32),

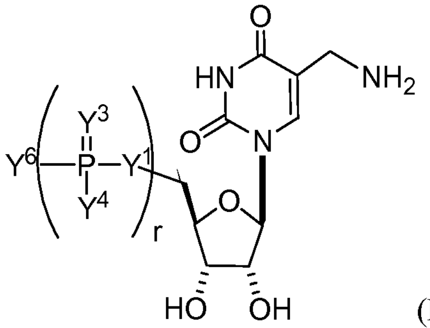


(BB-

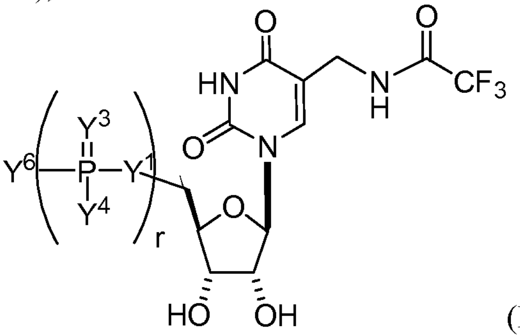
33),



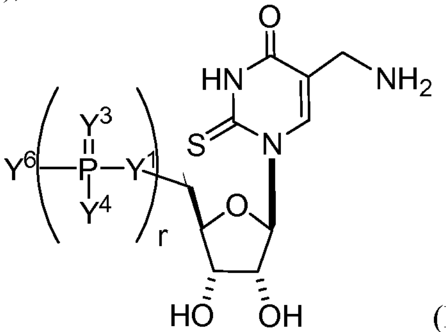
(BB- 34),



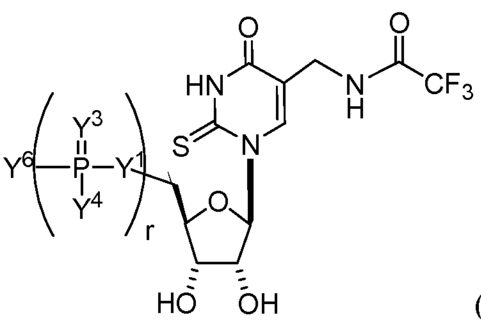
(BB- 35),



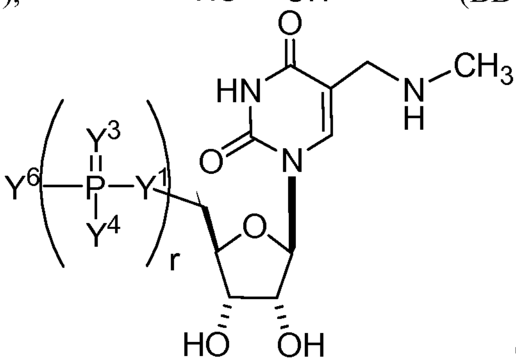
(BB- 36),



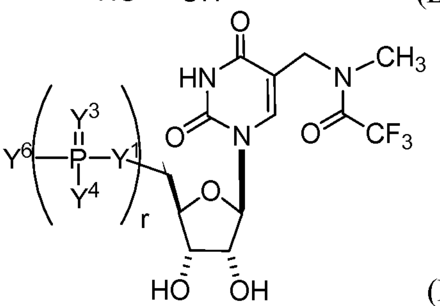
(BB- 37),



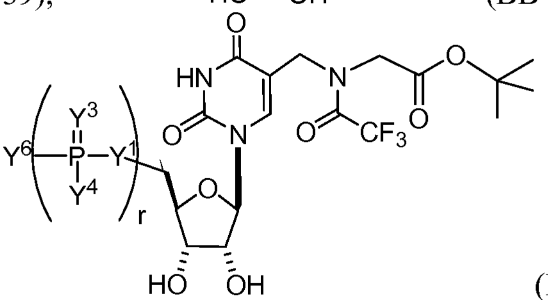
(BB- 38),



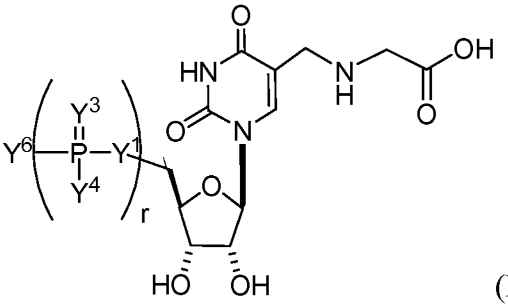
(BB-
39),



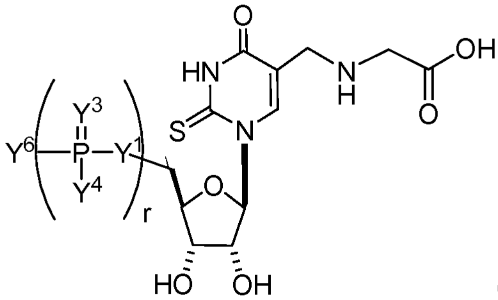
(BB- 40),



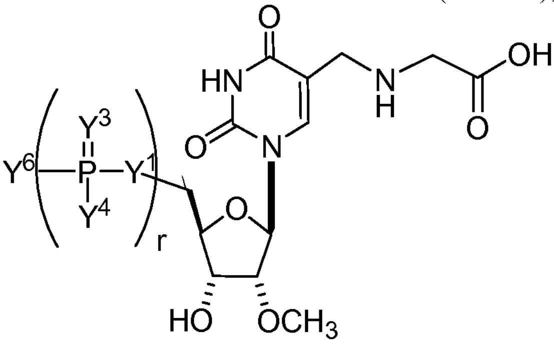
(BB- 41),



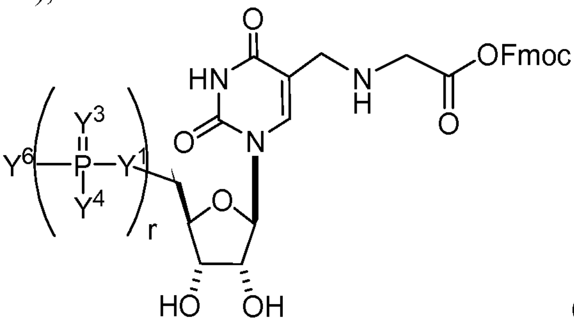
(BB- 42),



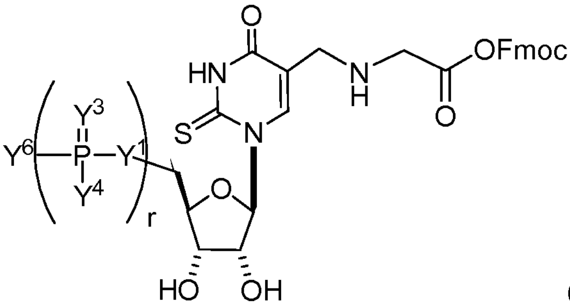
(BB- 43),



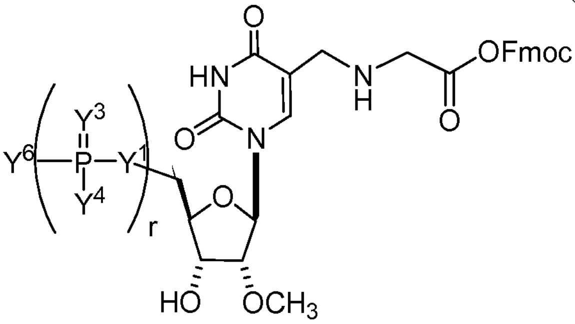
(BB- 44),



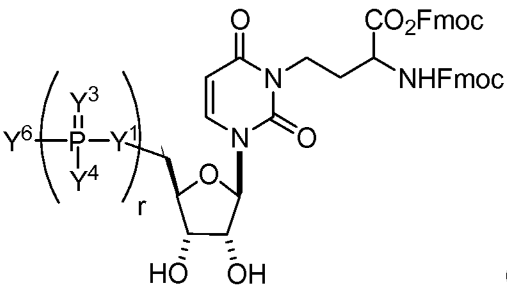
(BB- 45),



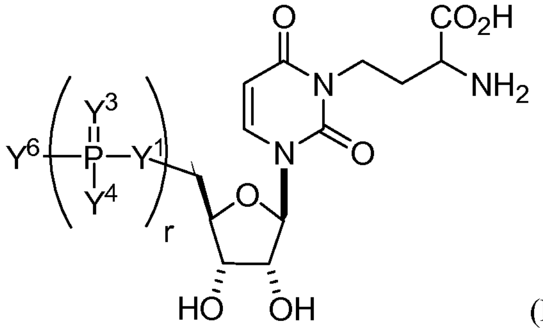
(BB-46),



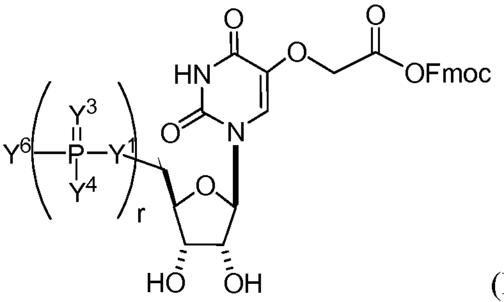
(BB- 47),



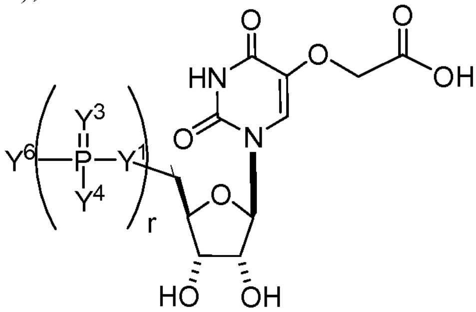
(BB- 48),



(BB- 49),

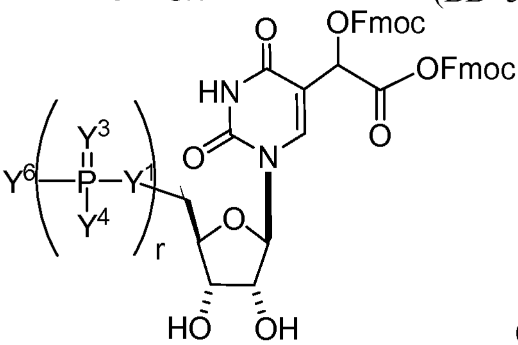


(BB- 50),

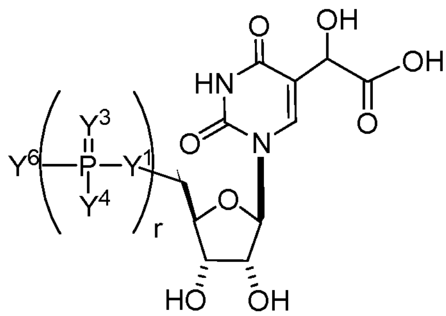


(BB-

51),

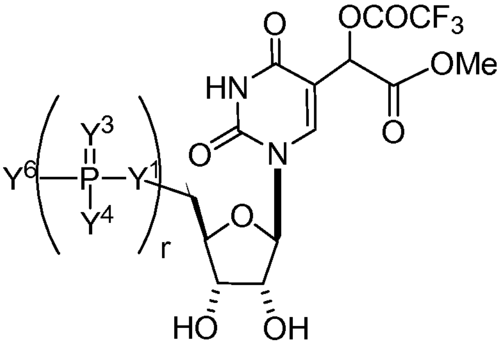


(BB- 52),

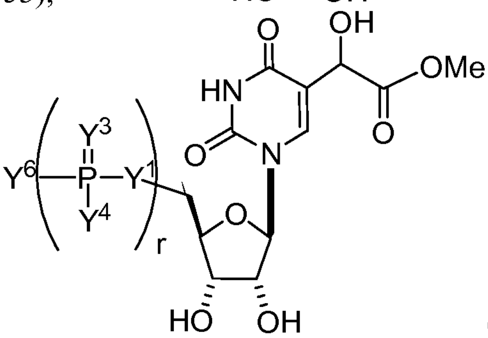


(BB-

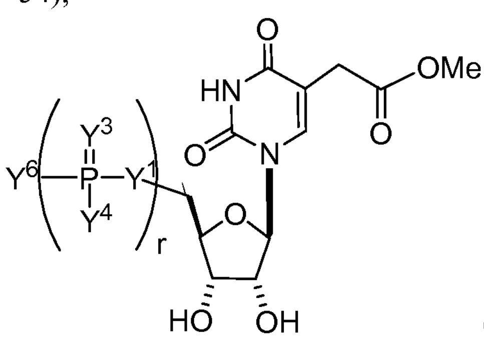
53),



(BB- 54),

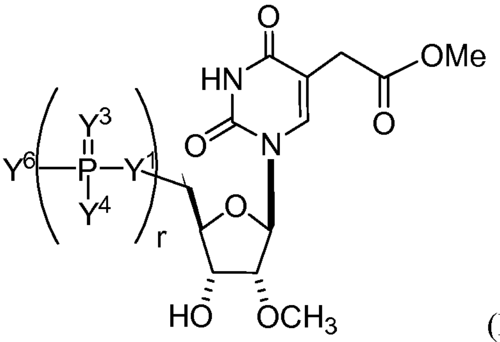


(BB- 55),

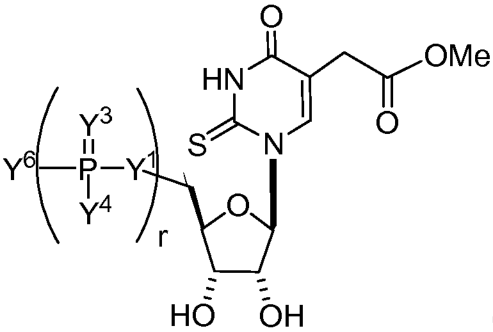


(BB-

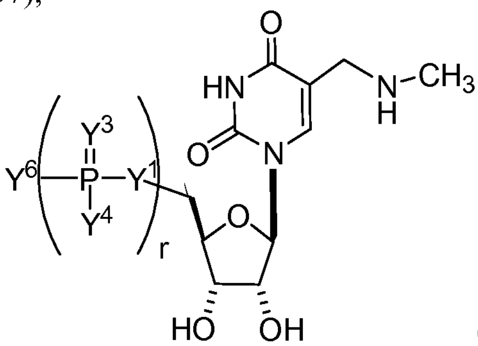
56),



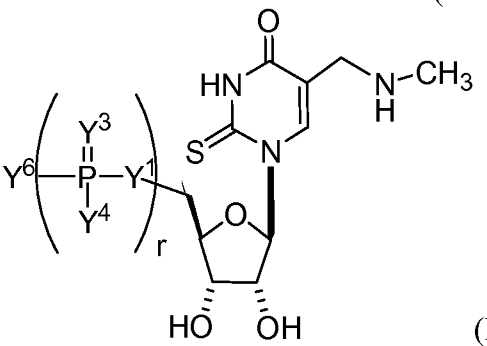
(BB-57),



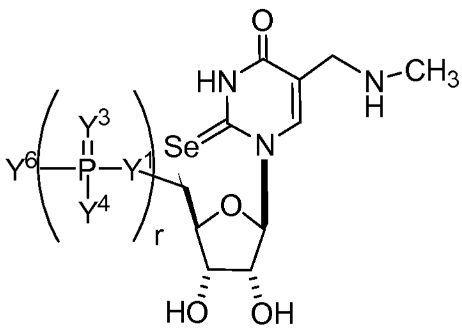
(BB- 58),



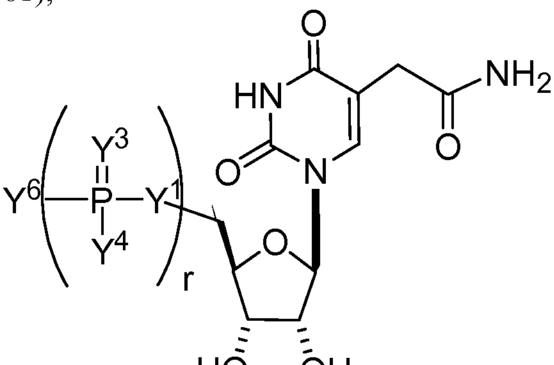
(BB- 59),



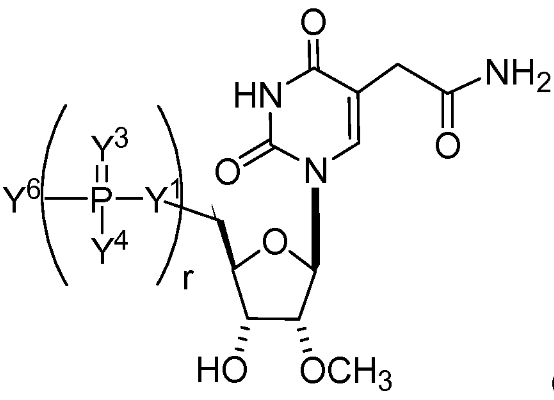
(BB- 60),



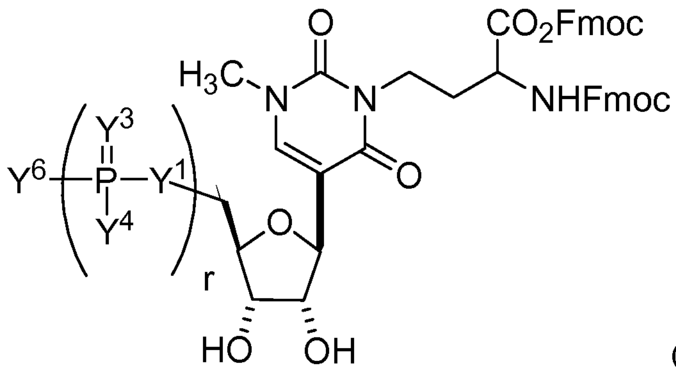
(BB- 61),



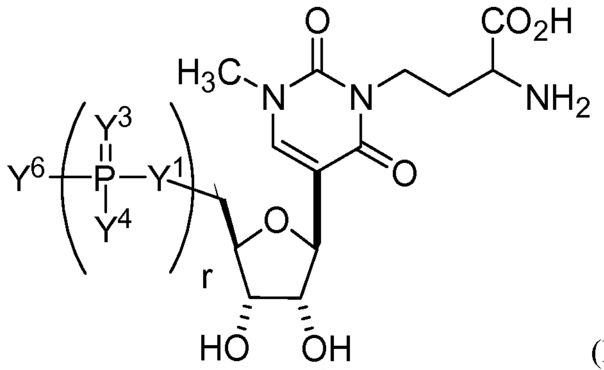
(BB- 62),



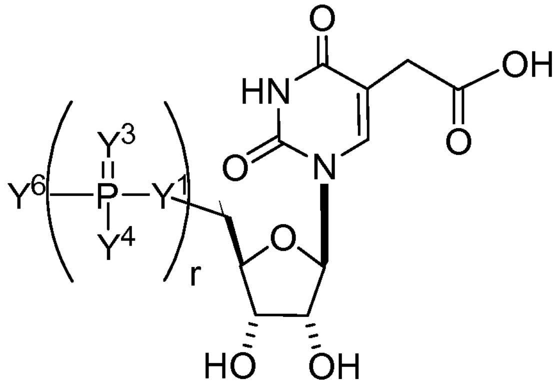
(BB- 63),



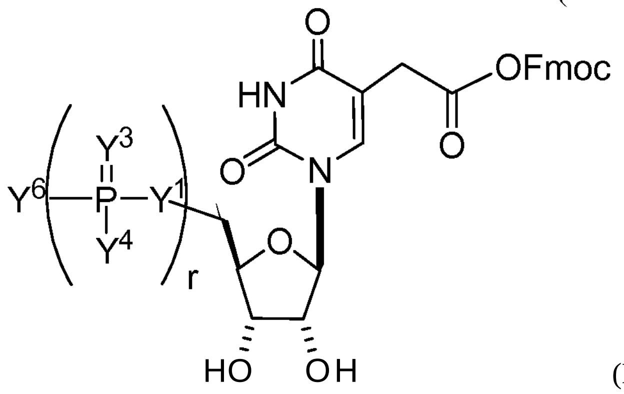
(BB- 64),



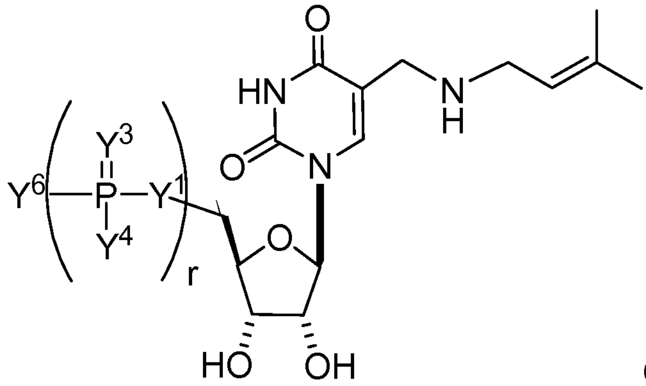
(BB- 65),



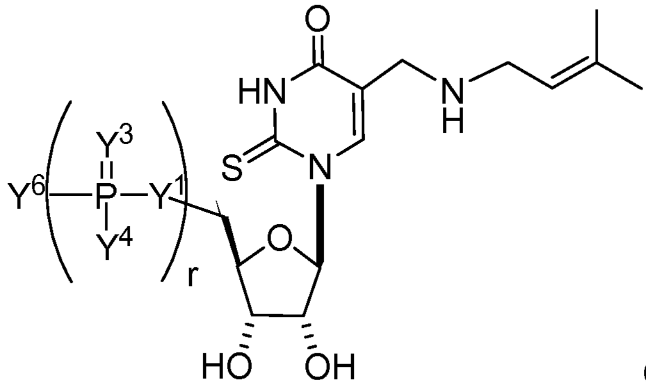
(BB- 66),



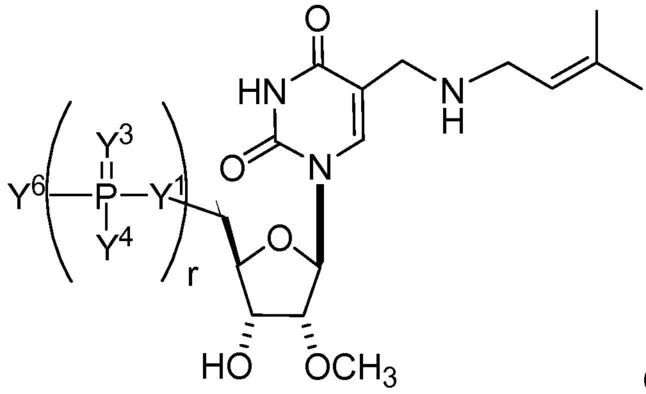
(BB- 67),



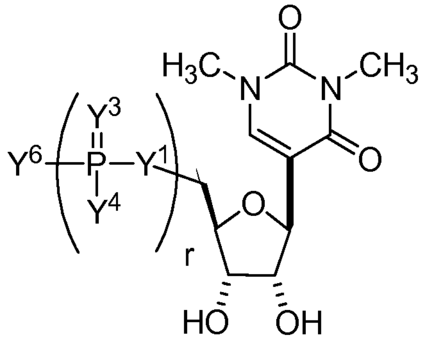
(BB- 68),

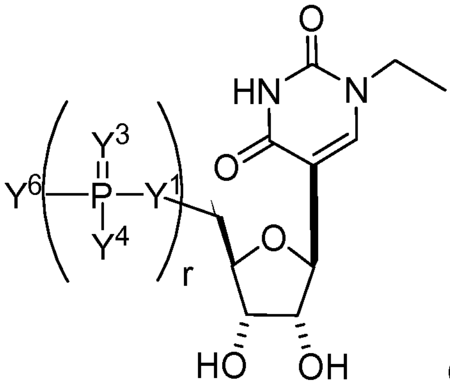


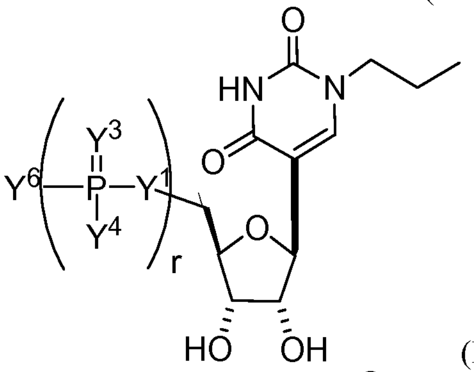
(BB- 69),

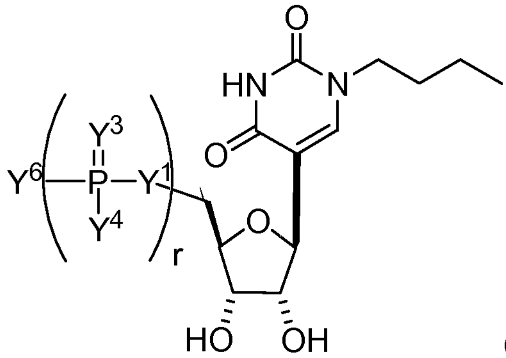


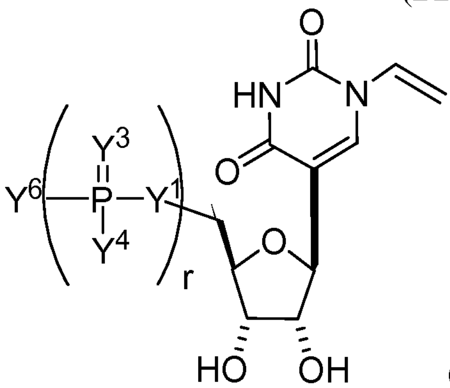
(BB- 70),

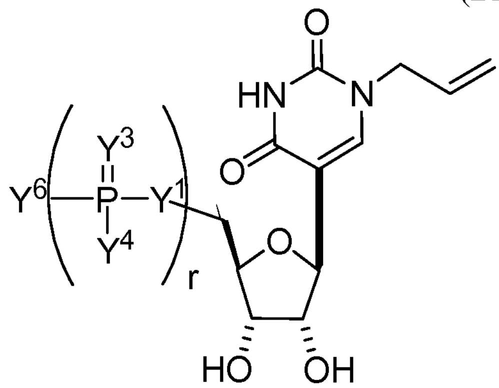

(BB- 71),

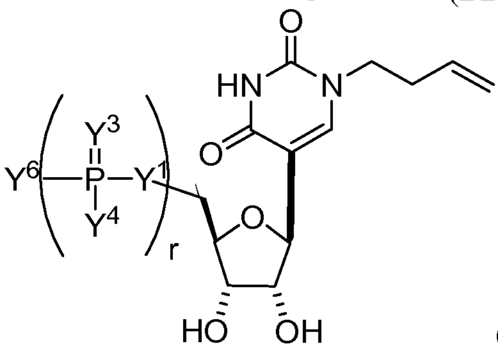

(BB- 72),

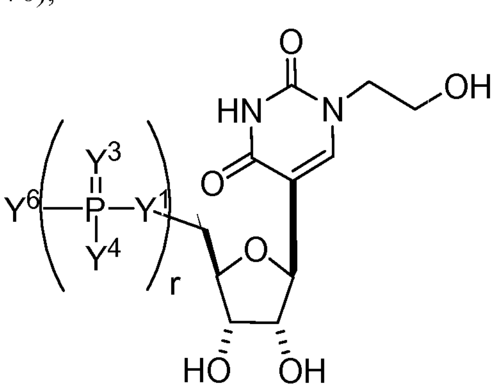

(BB- 73),

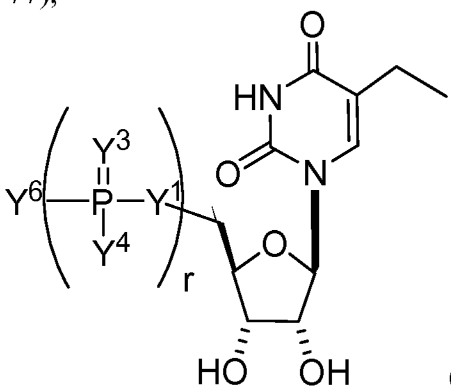

(BB- 74),

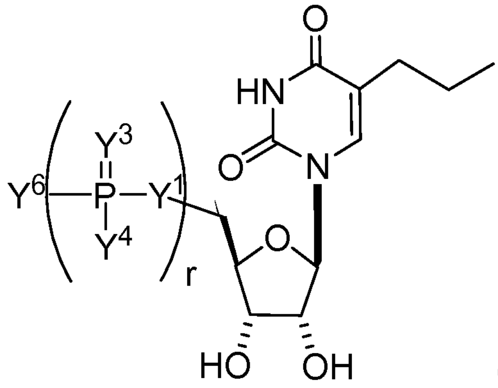

(BB- 75),


(BB- 76),

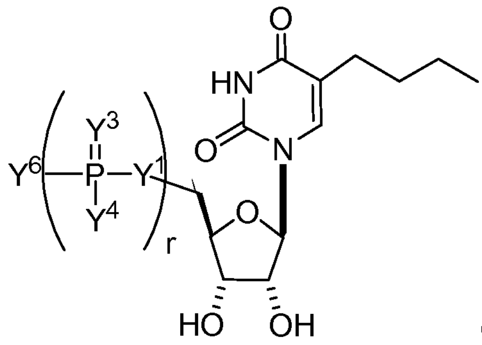

(BB- 77),


(BB- 78),

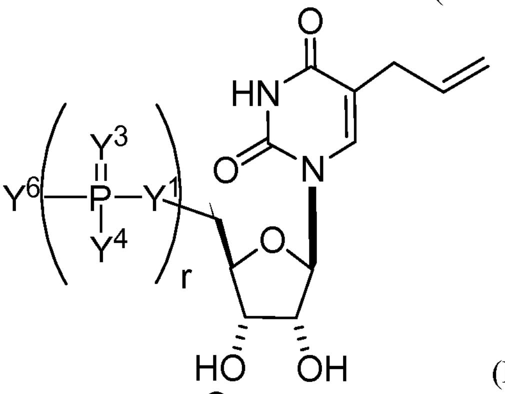

(BB- 79),



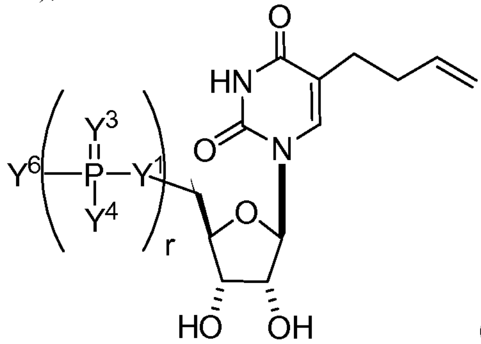
(BB- 80),



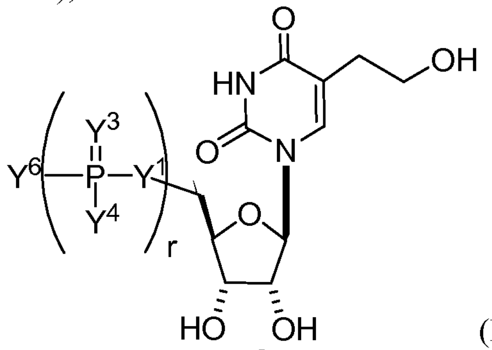
(BB- 81),



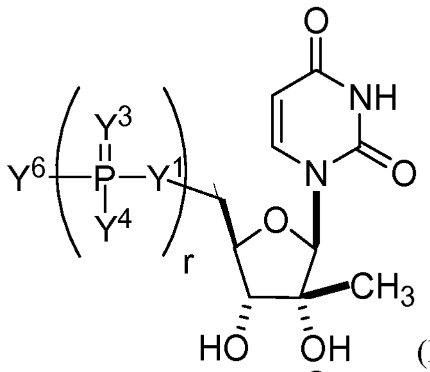
(BB- 82),



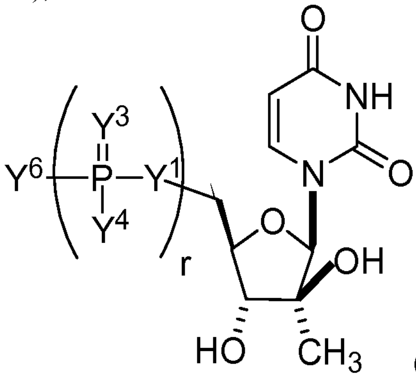
(BB- 83),



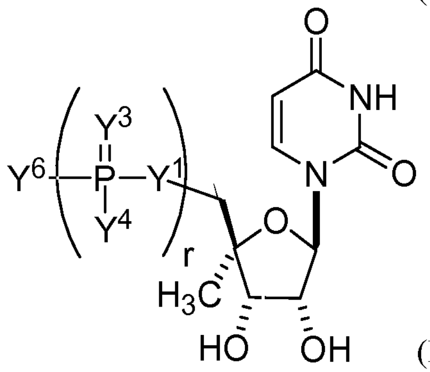
(BB- 84),



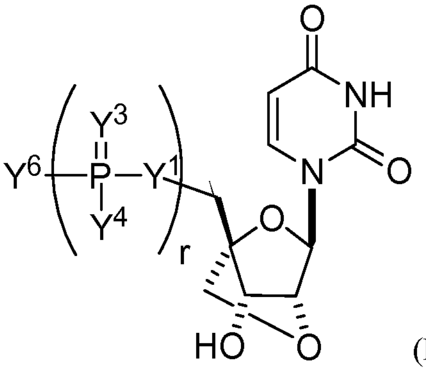
(BB- 85),



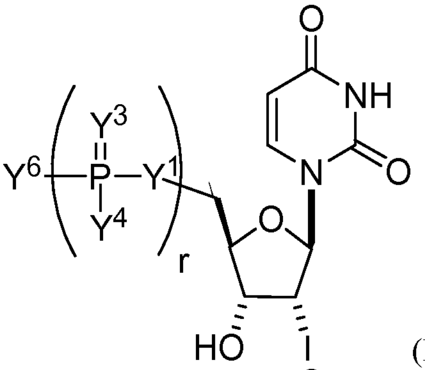
(BB- 86),

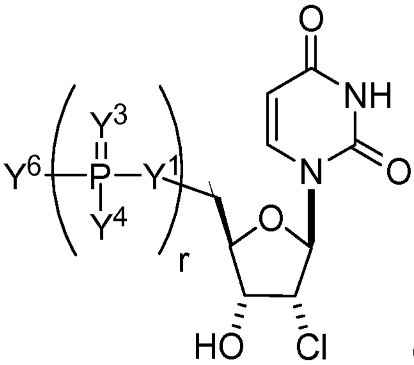


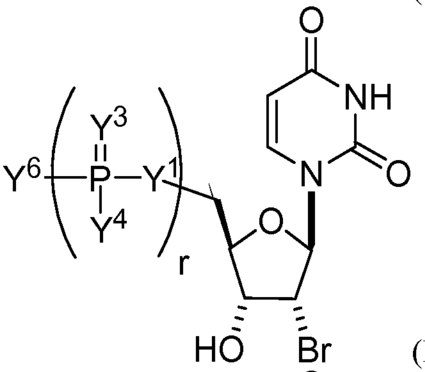
(BB- 87),

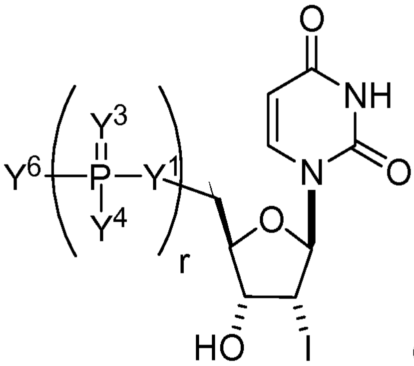


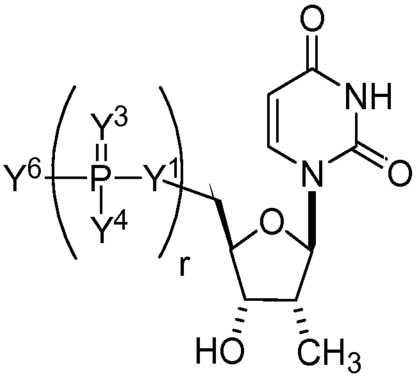
(BB- 88),

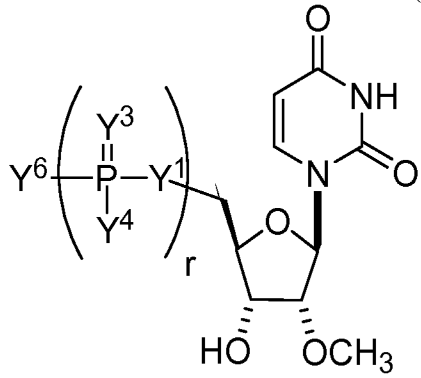

(BB- 89),

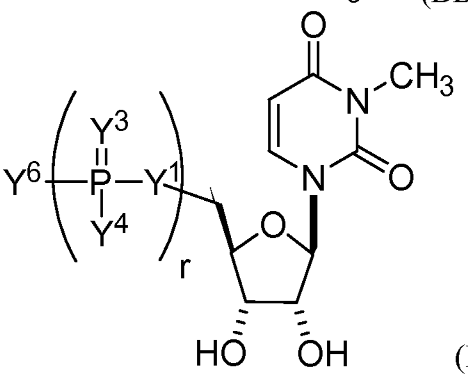

(BB- 90),

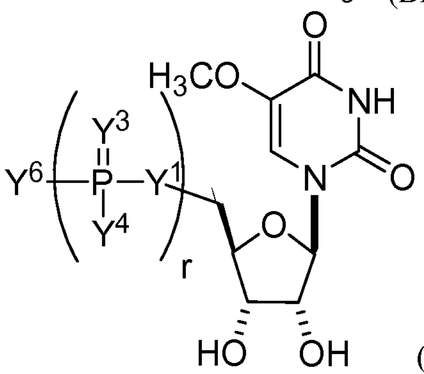

(BB- 91),

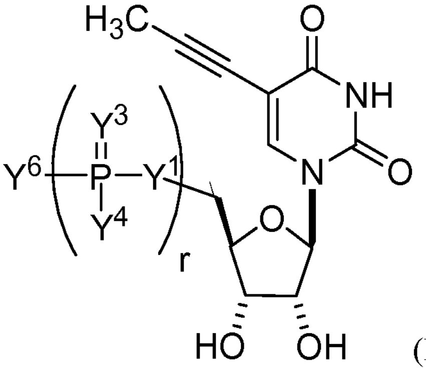

(BB- 92),

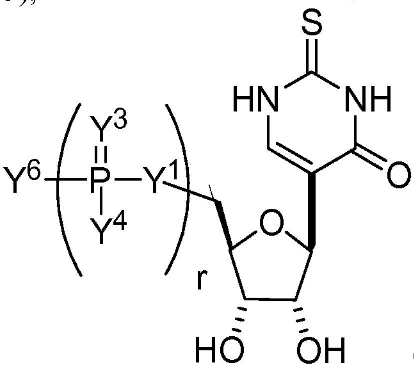

(BB- 93),

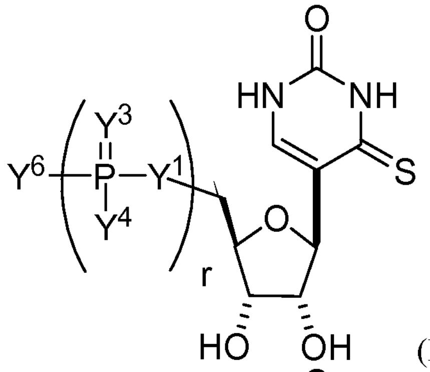

(BB- 94),

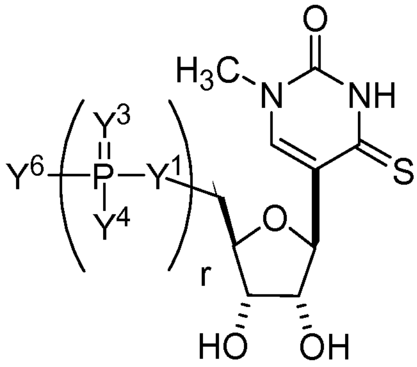

(BB- 95),

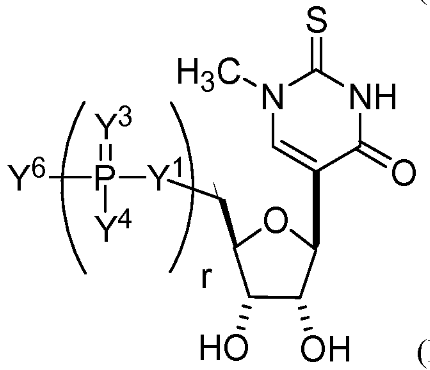

(BB- 96),

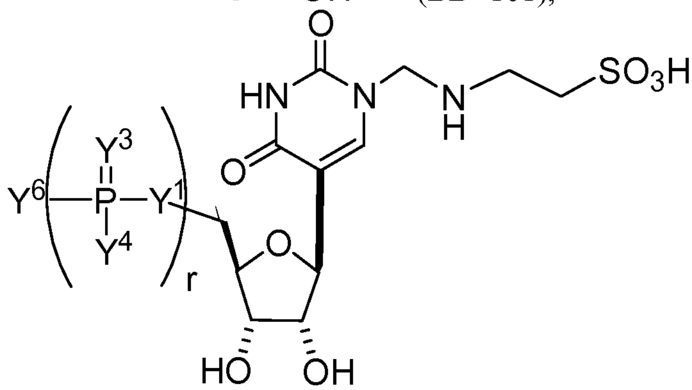

(BB- 97),

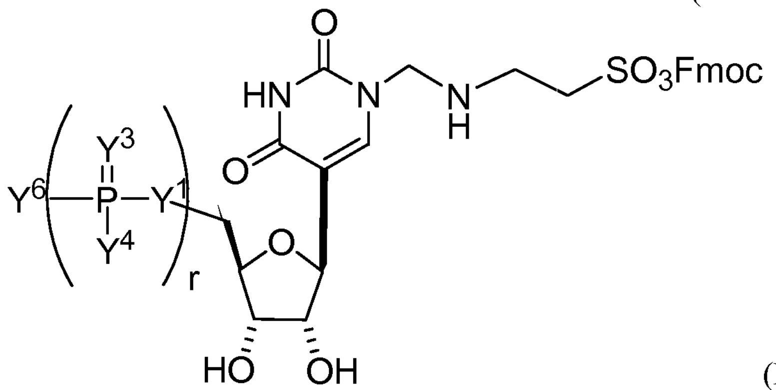

(BB- 98),

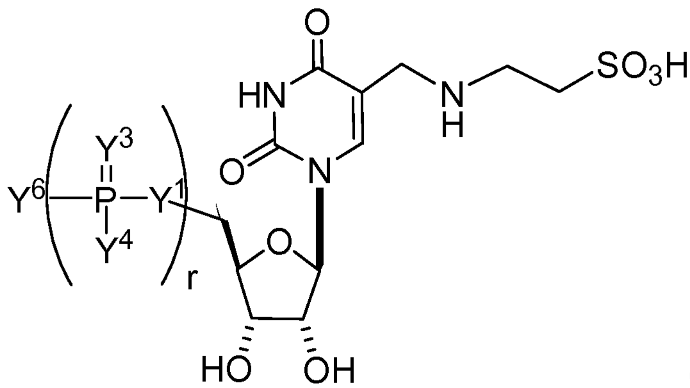

(BB- 99),

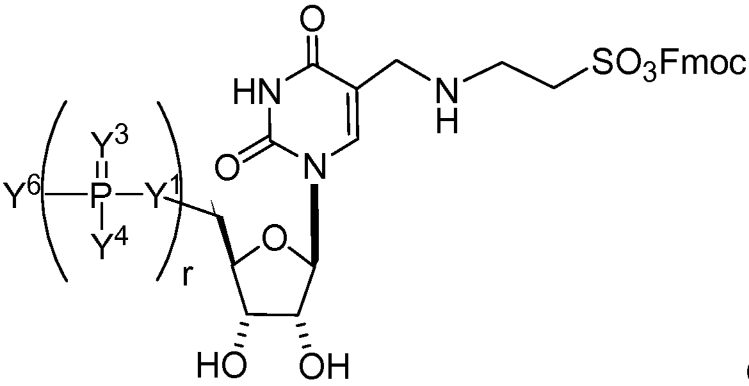

(BB- 100),

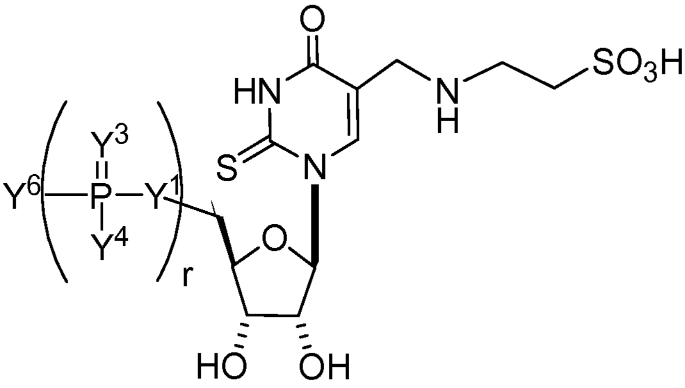

(BB- 101),

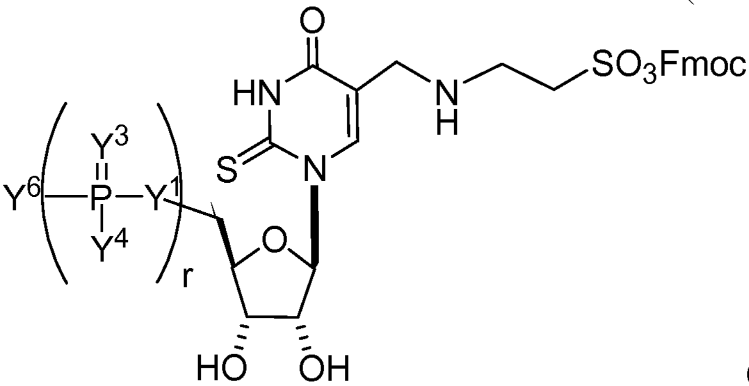

(BB- 102),

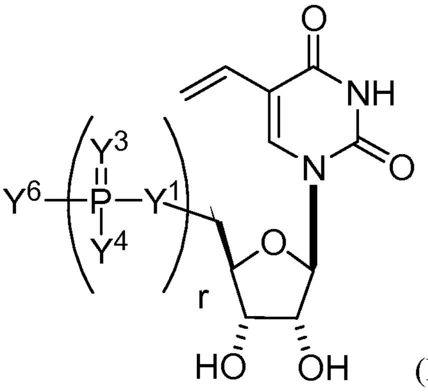

(BB- 103),

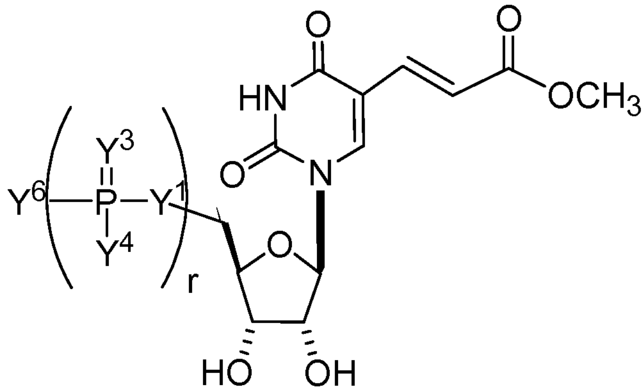

(BB- 104),


(BB- 105),

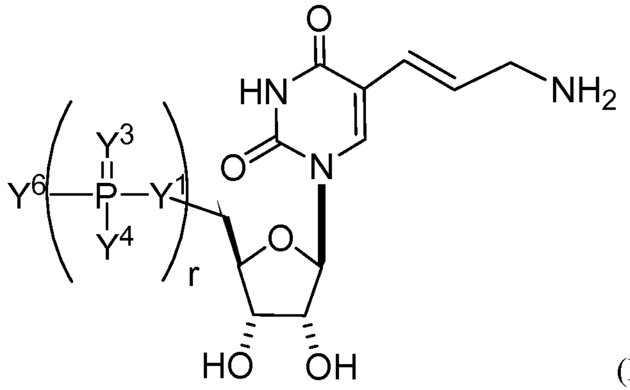

(BB- 106),


(BB- 107),

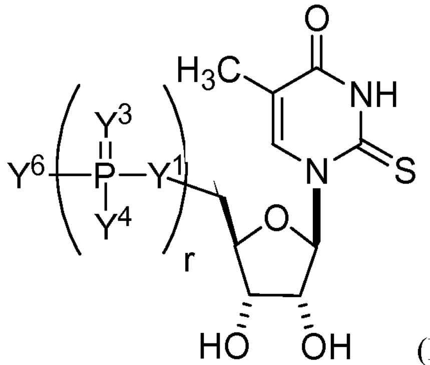

(BB- 108),



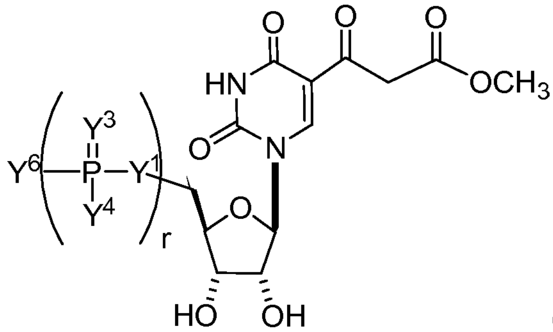
(BB- 109),



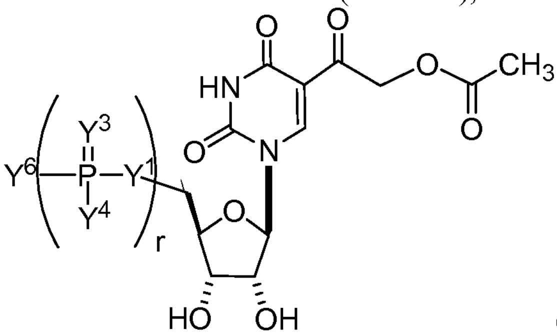
(BB- 110),



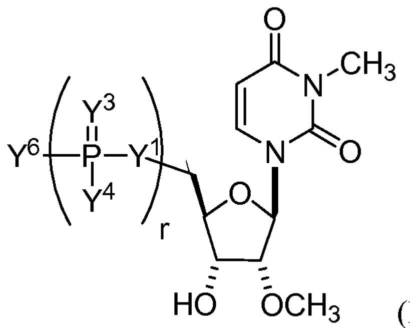
(BB- 111),



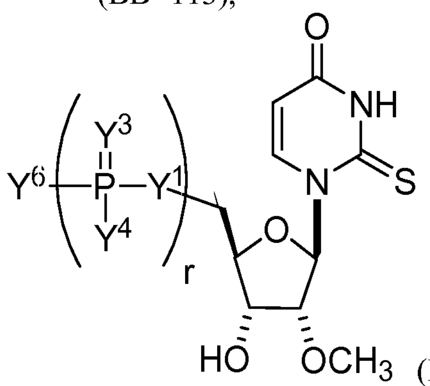
(BB- 112),



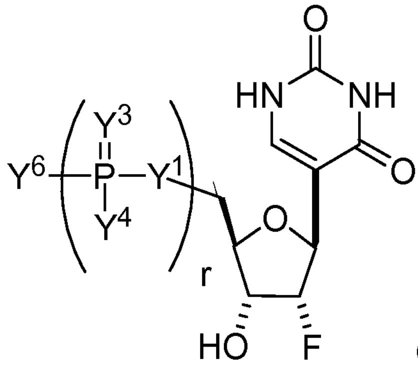
(BB- 113),

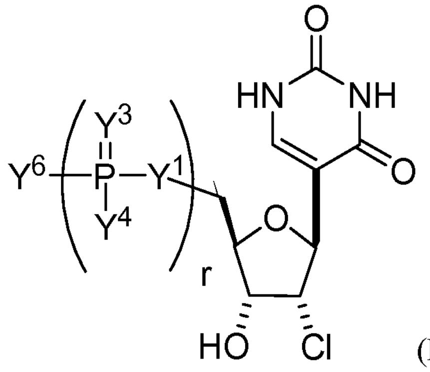


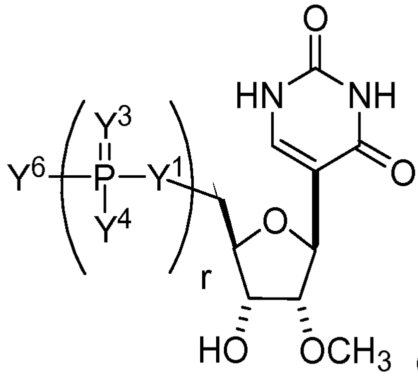
(BB- 114),

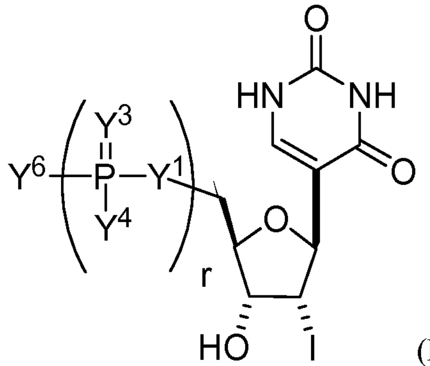


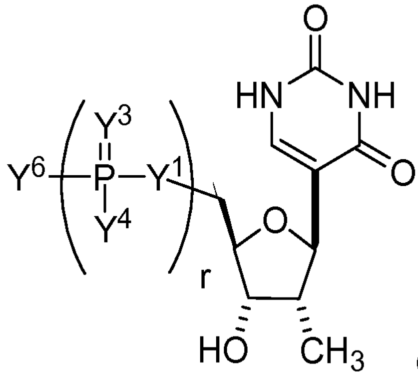
(BB- 115),

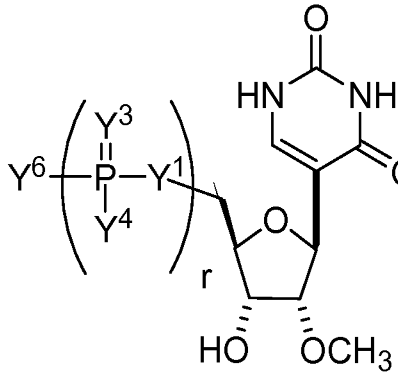
 (BB- 116),

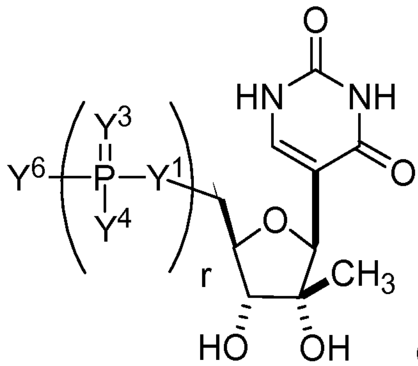
 (BB- 117),

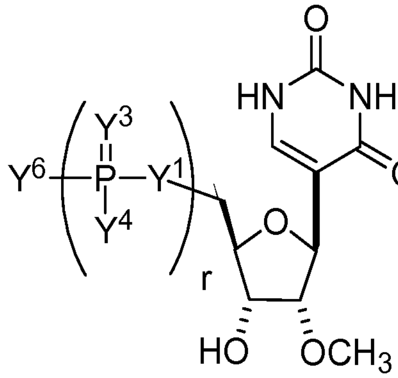
 (BB- 118),

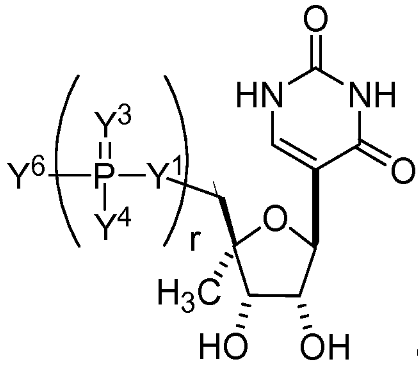
 (BB- 119),

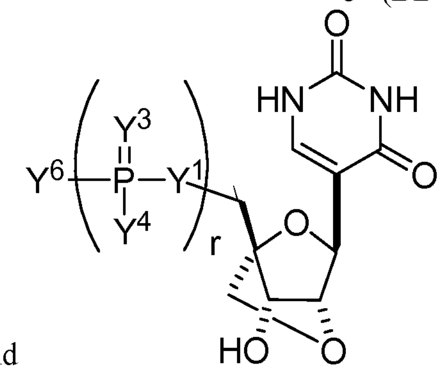
 (BB- 120),

 (BB- 121),

 (BB- 122),

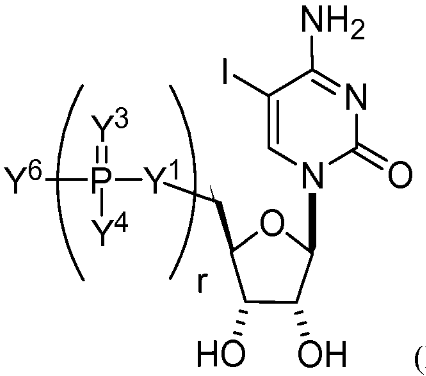
 (BB- 123),

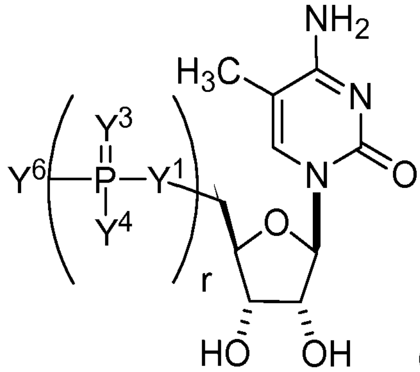
 (BB- 124), and

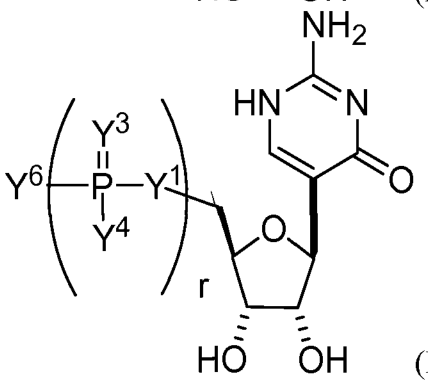
 (BB- 125),

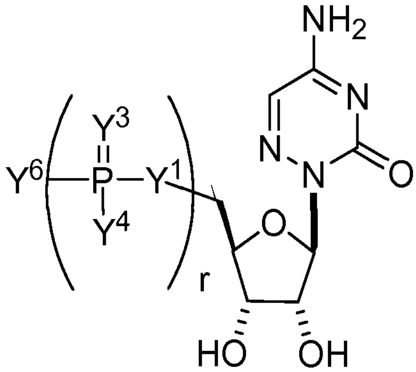
[00275] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein Y 1, Y3, Y4, Y6, and r are as described herein (e.g., each r is, independently, an integer from 0 to 5, such as from 0 to 3, from 1 to 3, or from 1 to 5)).

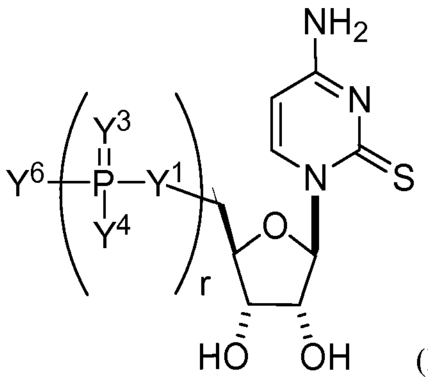
[00276] In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, is a modified cytidine (e.g., selected from the group consisting of:

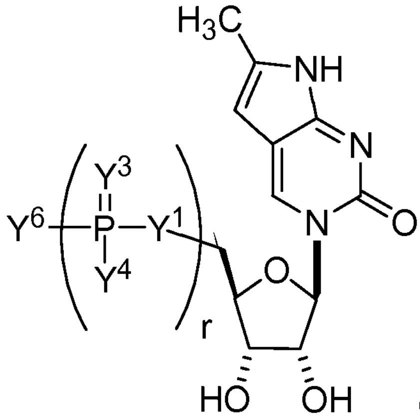

(BB- 126),

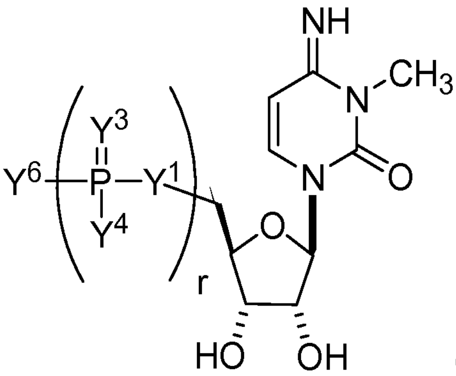

(BB- 127),


(BB- 128),

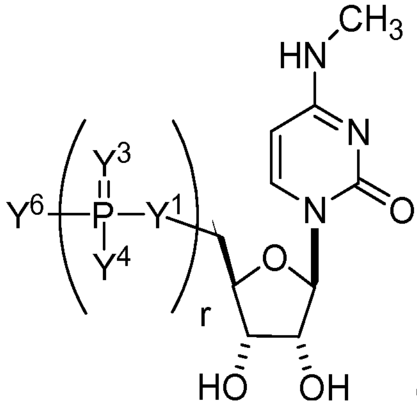

(BB- 129),


(BB- 130),

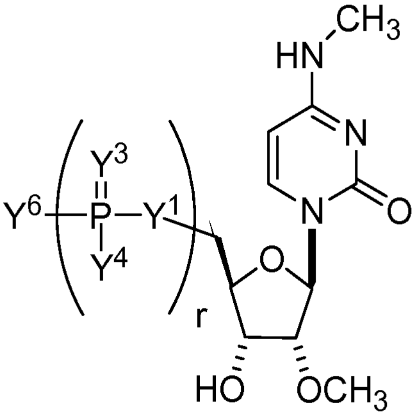

(BB- 131),



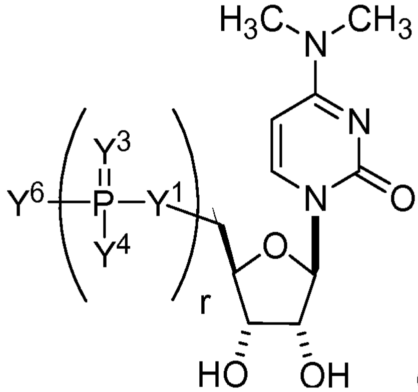
(BB- 132),



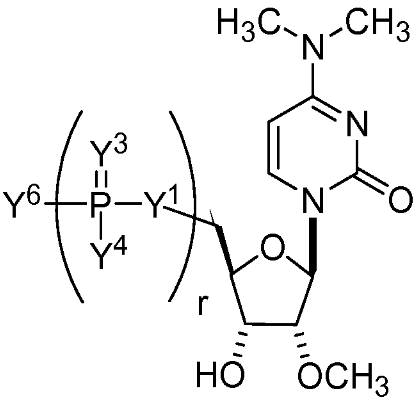
(BB- 133),



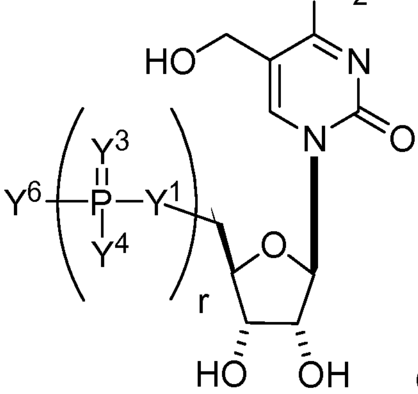
(BB- 134),



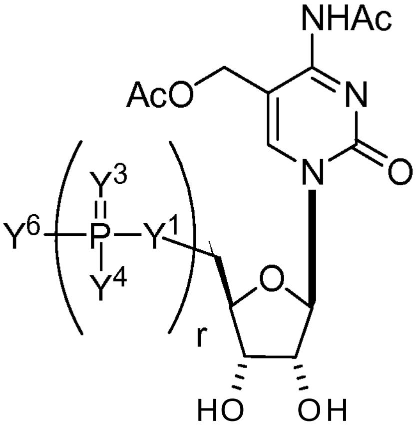
(BB- 135),



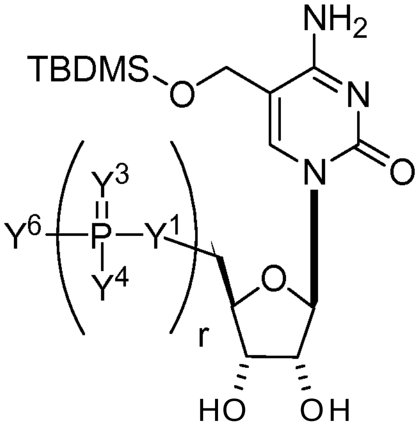
(BB- 136),



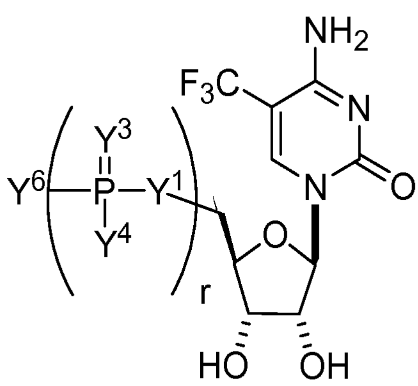
(BB- 137),

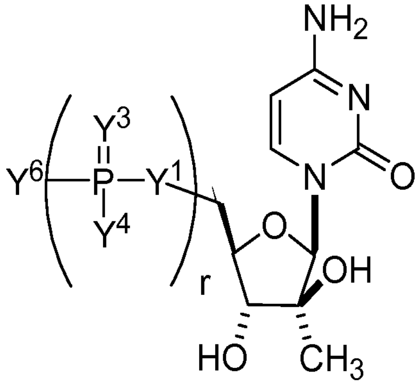


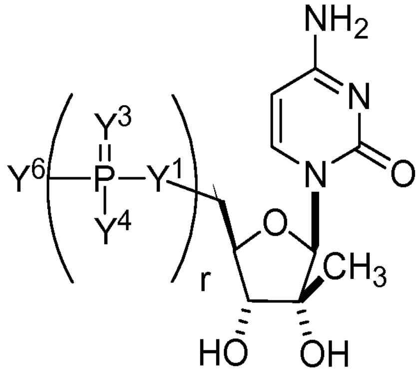
(BB- 138),

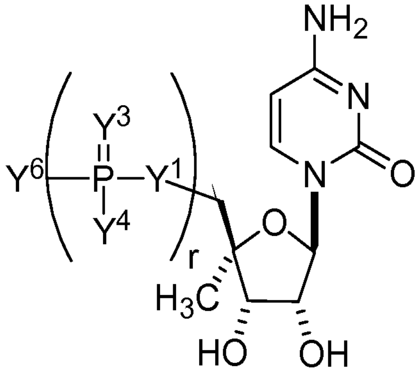


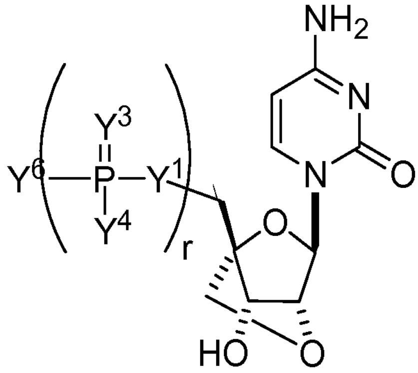
(BB- 139),

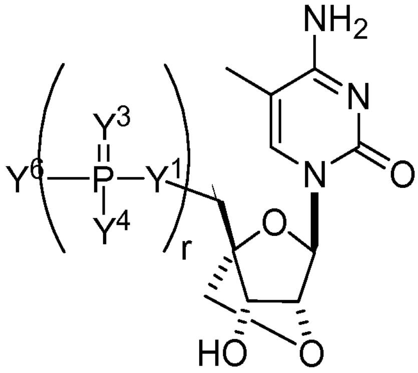

(BB- 140),

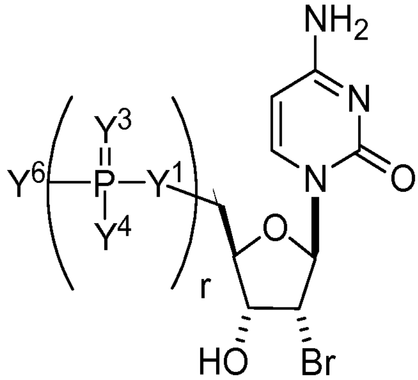

(BB- 141),

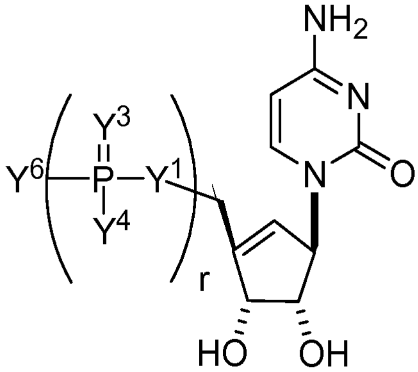

(BB- 142),

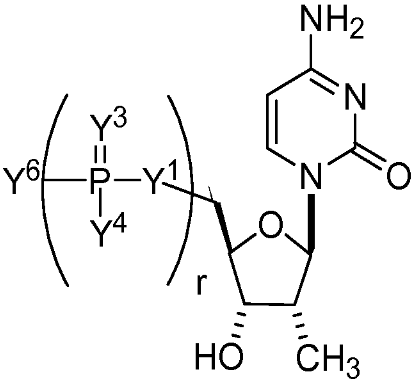

(BB- 143),

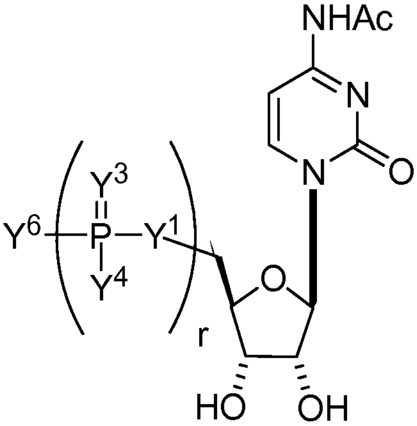

(BB- 144),

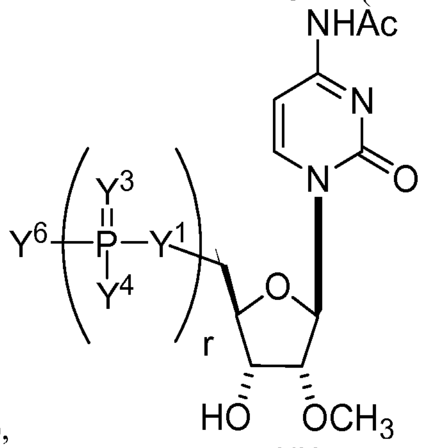

(BB- 145),

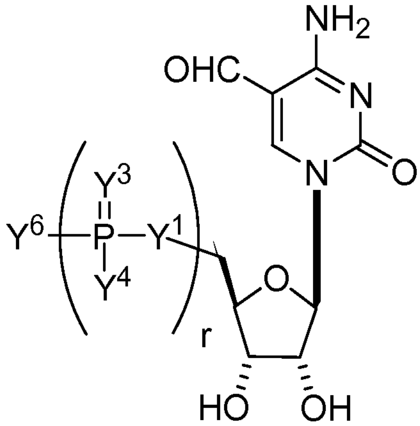

(BB- 146),

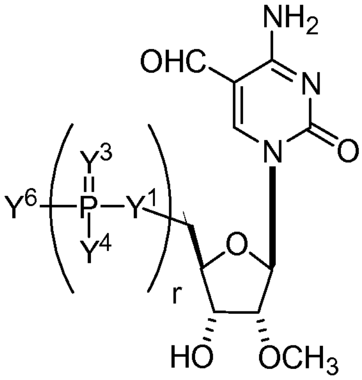

(BB- 147),

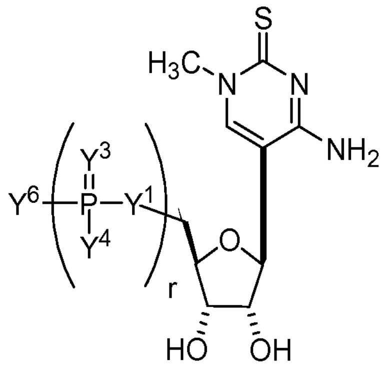

(BB- 148),

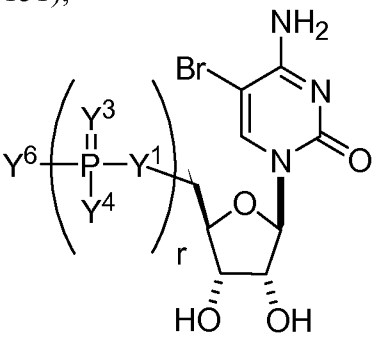

(BB-
149),

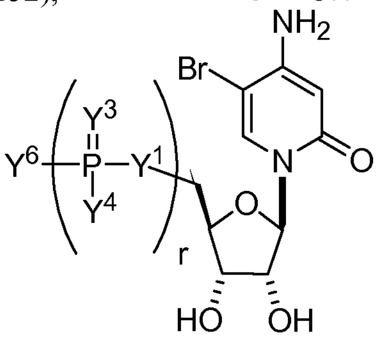

(BB- 150),

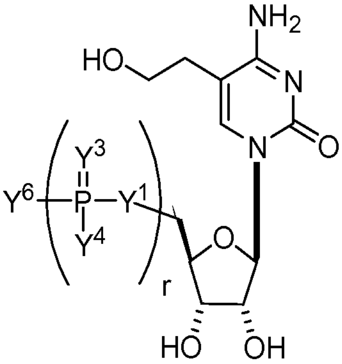
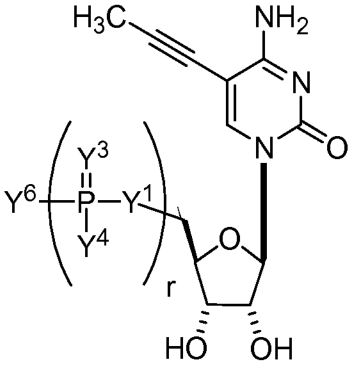

(BB-
151),

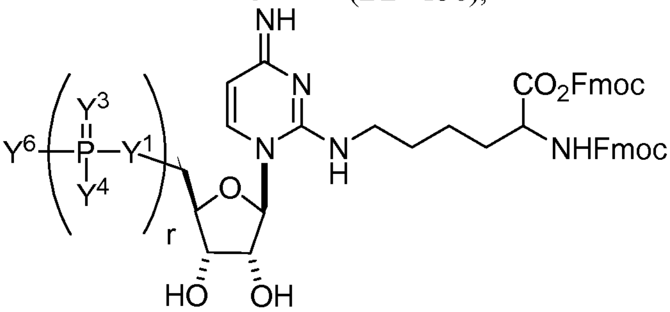

(BB- 152),

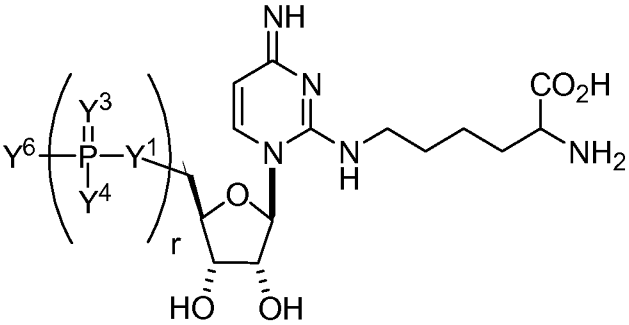

(BB- 153),


(BB- 154),

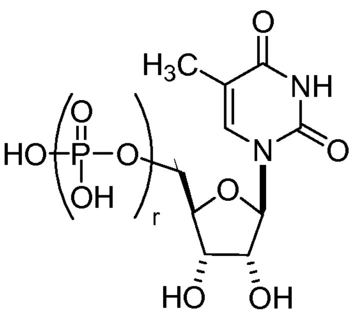
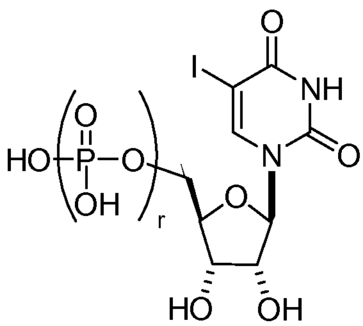

(BB- 155),

 (BB- 156),  (BB- 157),

 (BB- 158), and

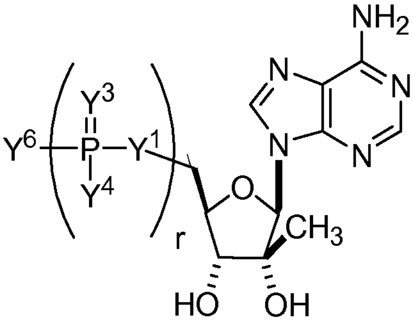
 (BB- 159),

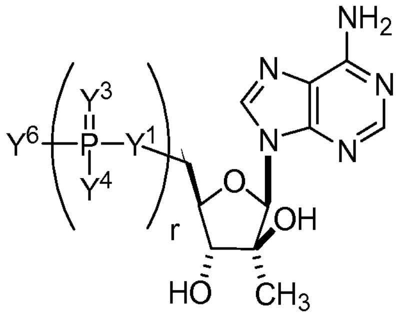
[00277] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein Yl, Y3, Y4, Y6, and r are as described herein (e.g., each r is, independently, an integer from 0 to 5, such as from 0 to 3, from 1 to 3, or from 1 to 5)). For example, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, can be:

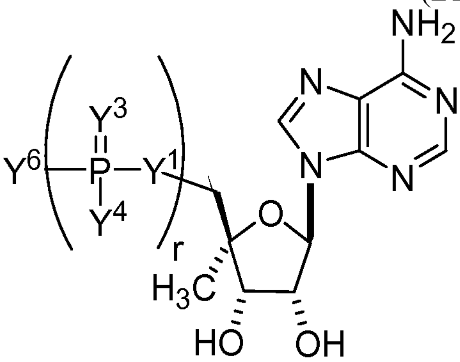
 (BB- 160) or  (BB- 161),

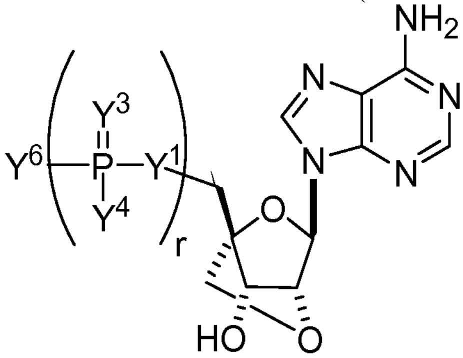
[00278] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each r is, independently, an integer from 0 to 5 (e.g., from 0 to 3, from 1 to 3, or from 1 to 5).

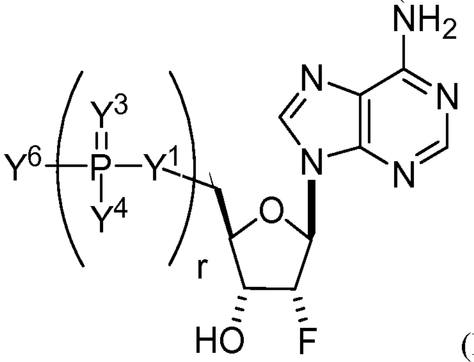
[00279] In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, is a modified adenosine (e.g., selected from the group consisting of:

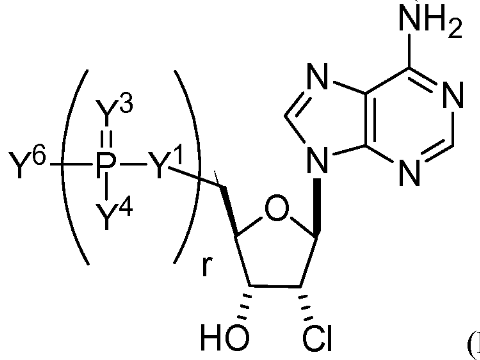

(BB- 162),

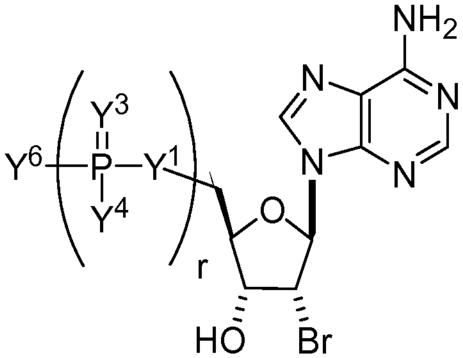

(BB- 163),

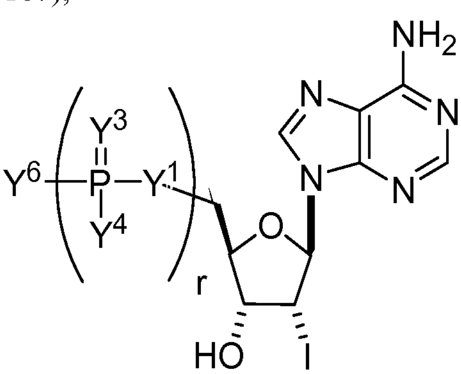

(BB- 164),

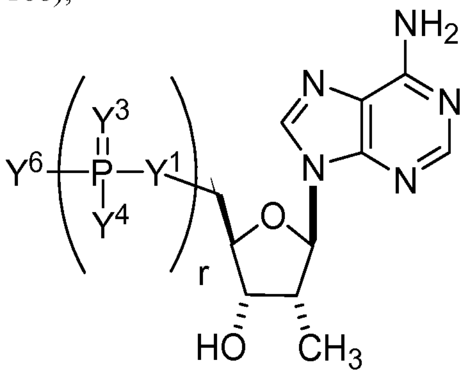

(BB-165),

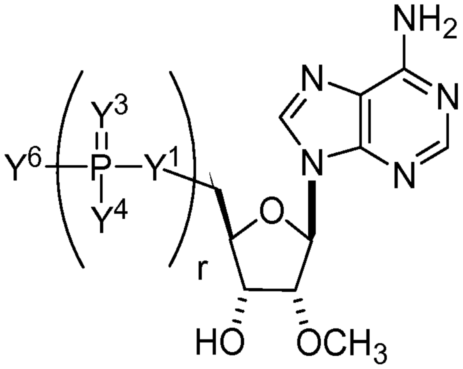
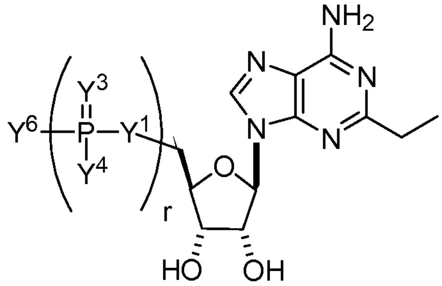
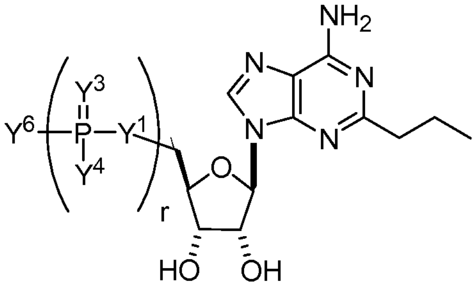
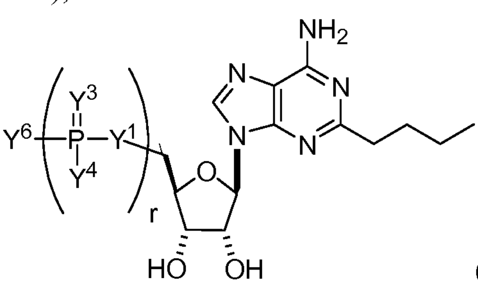
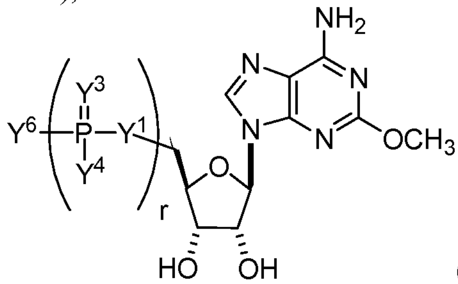

(BB- 166),

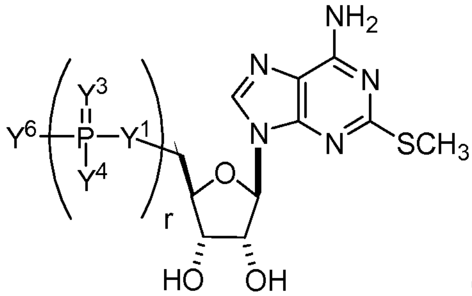
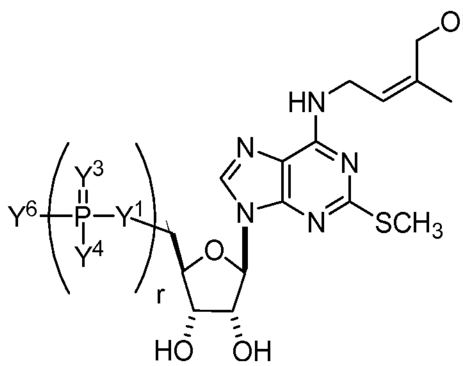
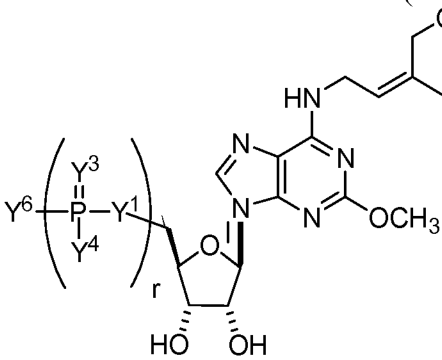
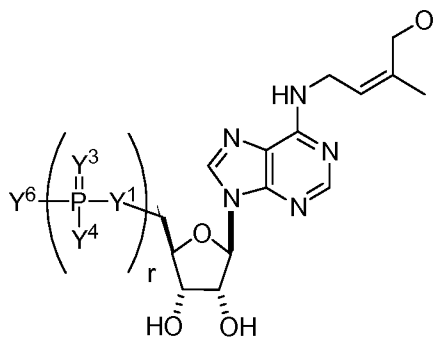

(BB-167),


(BB- 168),

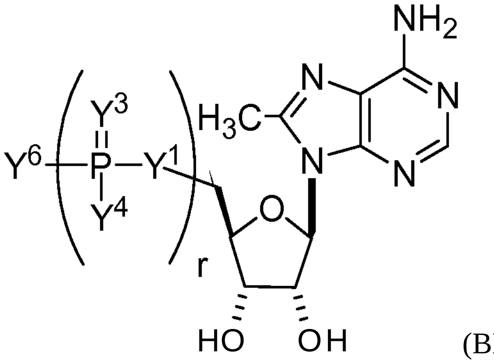

(BB- 169),


(BB-

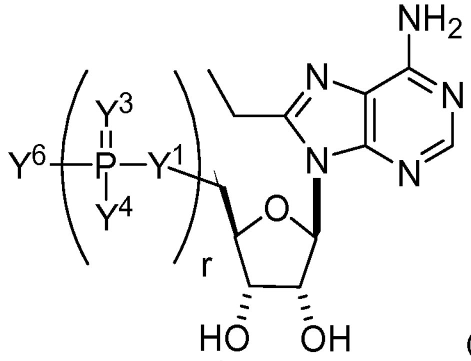
170),  (BB- 171),  (BB- 172),  (BB- 173),  (BB- 174),  (BB- 175),

 (BB- 176),  (BB- 177),  (BB- 178),  (BB-

179),

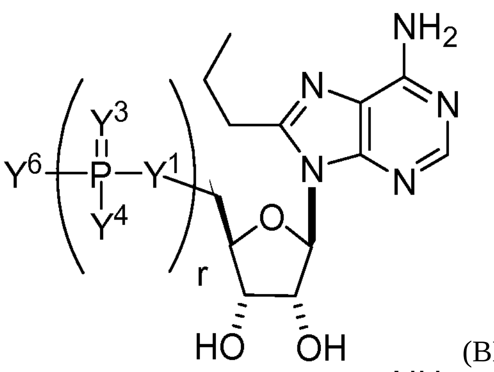


(BB- 180),

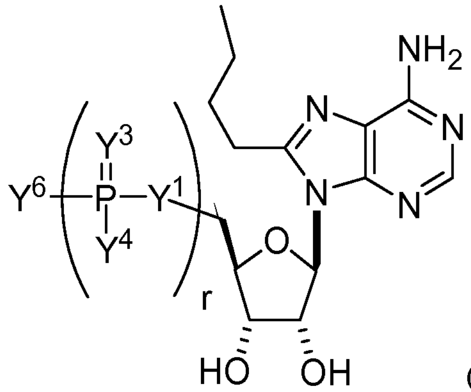


(BB-

181),

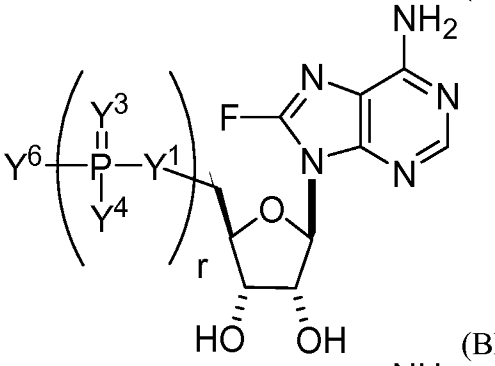


(BB- 182),

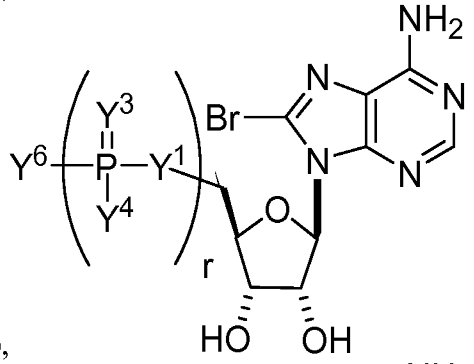


(BB-

183),

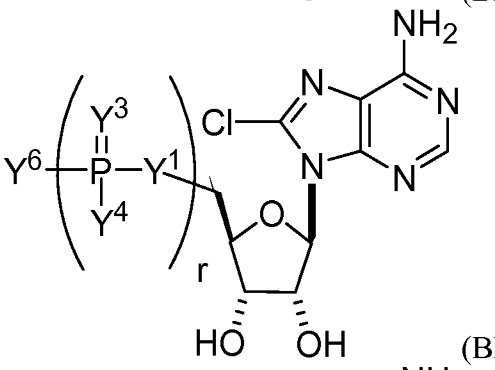


(BB- 184),

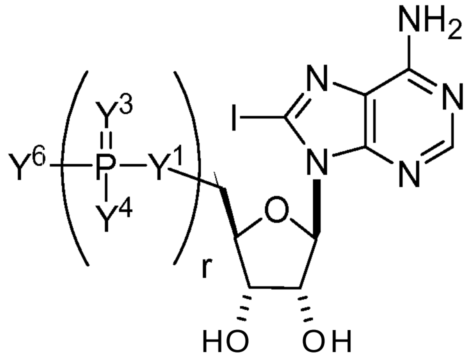


(BB-

185),

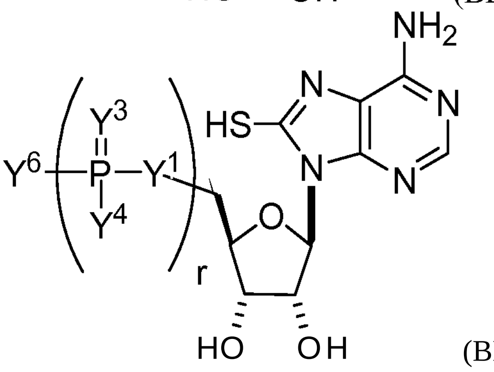


(BB- 186),

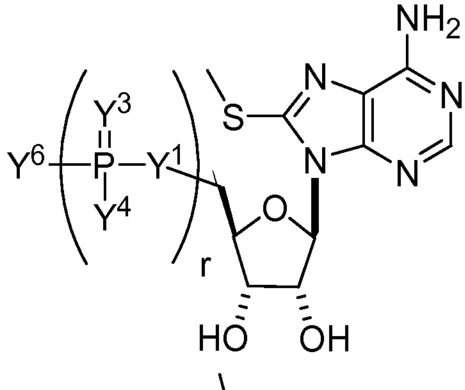


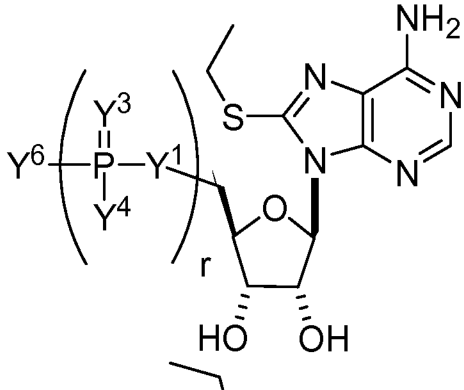
(BB-

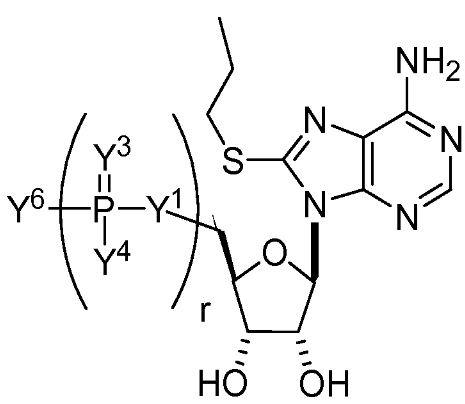
187),

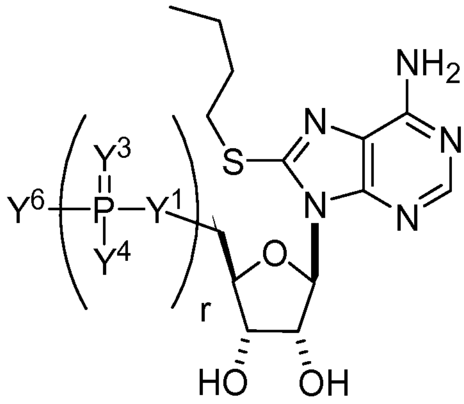


(BB- 188),

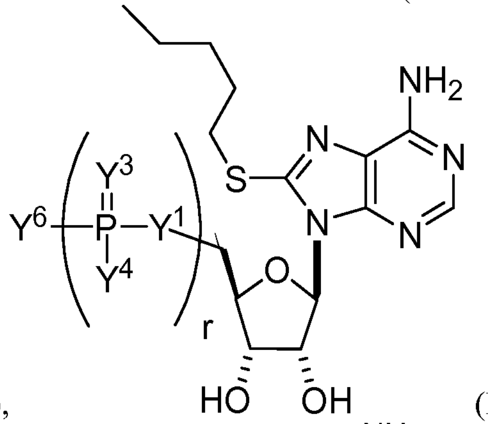
 (BB- 189),

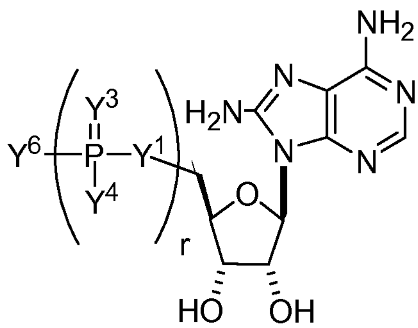
 (BB- 190),

 (BB- 191),

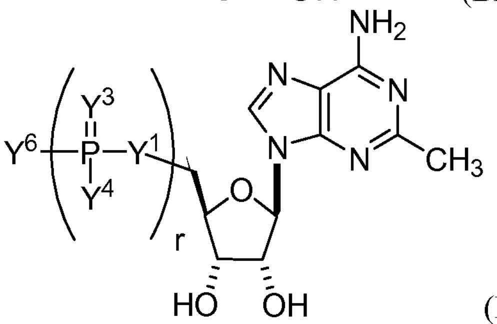
 (BB-

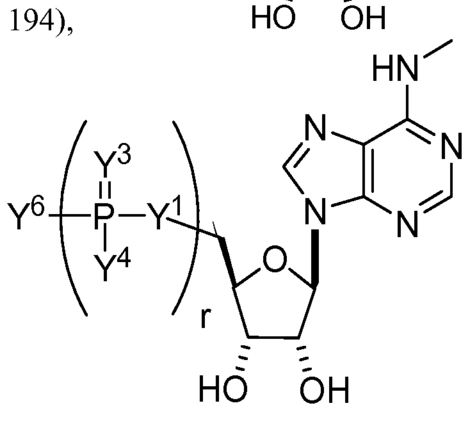
192),

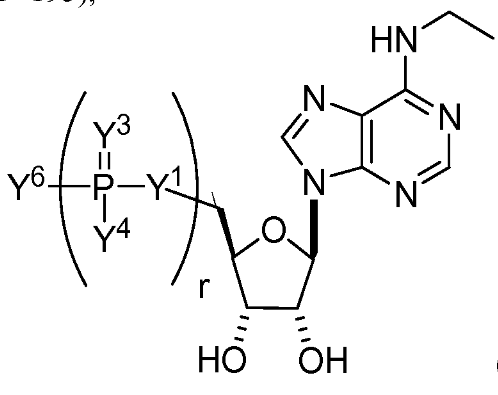
 (BB- 193),

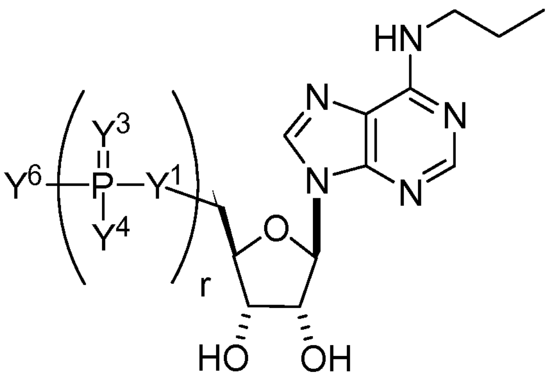
 (BB-

194),

 (BB- 195),

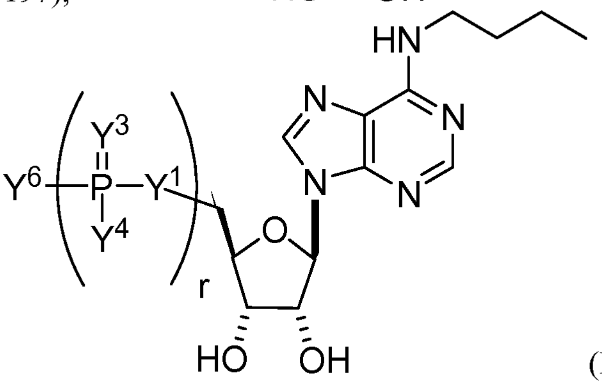
 (BB- 196),

 (BB-

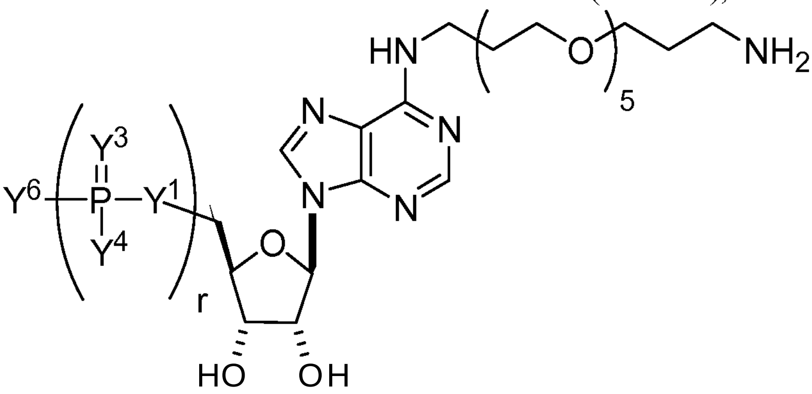


197),

(BB- 198),



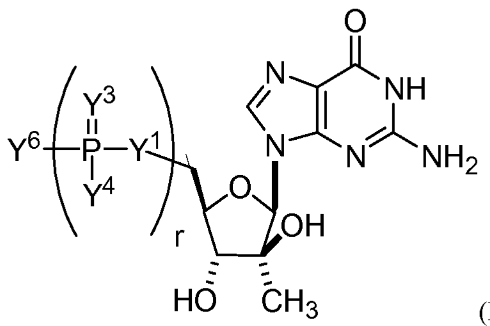
(BB- 199), and

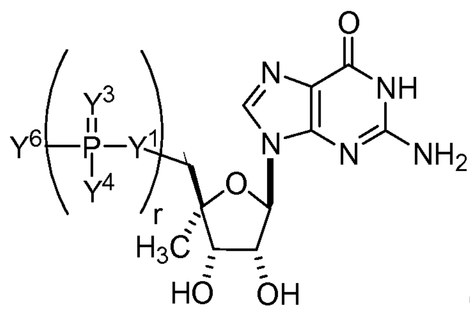


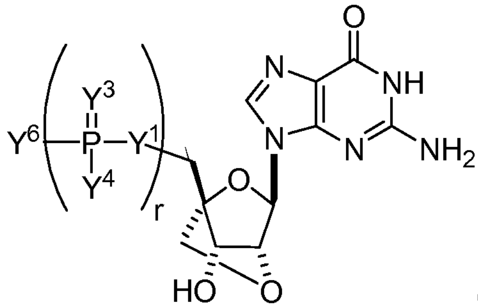
(BB- 200)

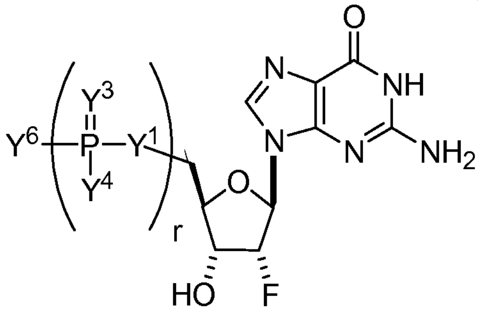
[00280] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein Yl, Y3, Y4, Y6, and r are as described herein (e.g., each r is, independently, an integer from 0 to 5, such as from 0 to 3, from 1 to 3, or from 1 to 5)).

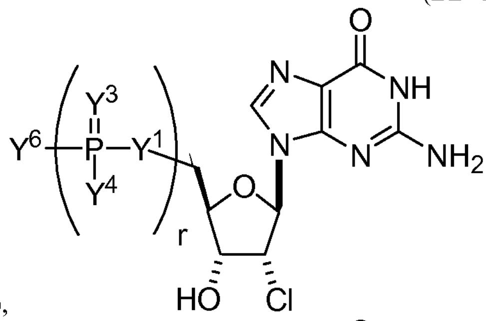
[00281] In some embodiments, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, is a modified guanosine (e.g., selected from the group consisting of:

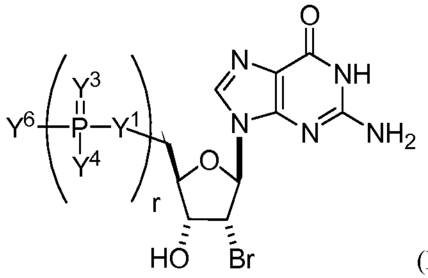

(BB- 201),

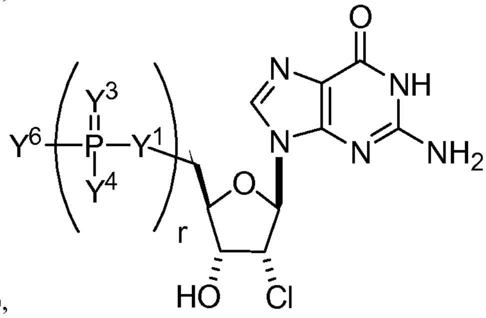

(BB- 202),

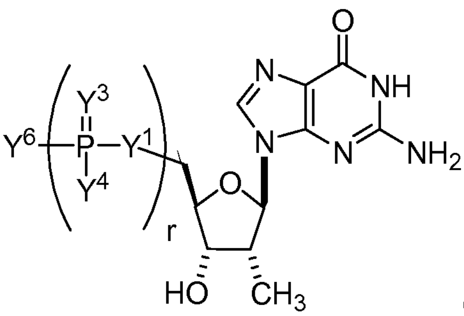
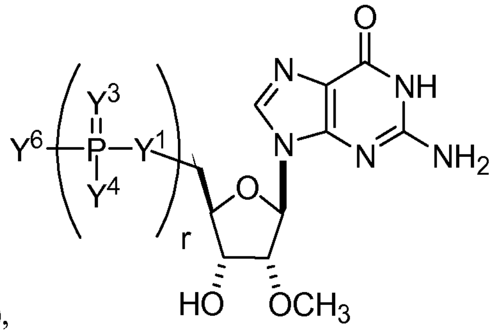

(BB- 203),

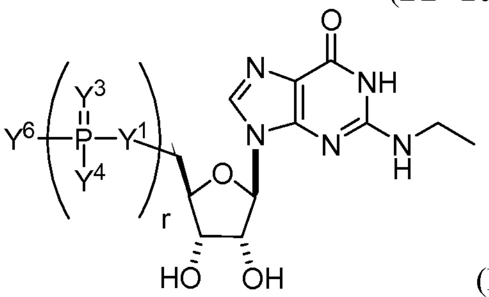

(BB-
204),

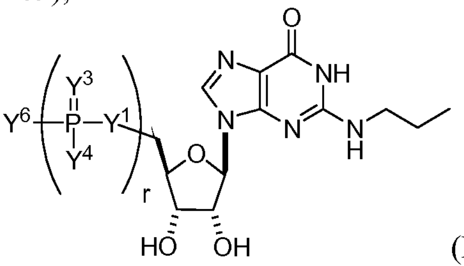
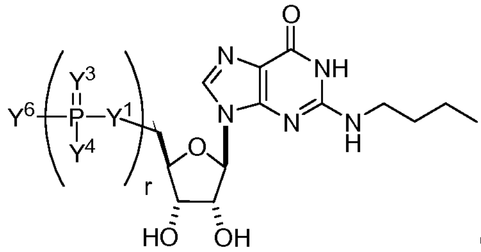

(BB- 205),

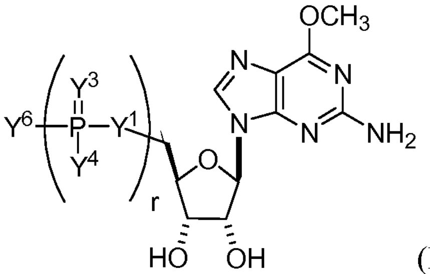
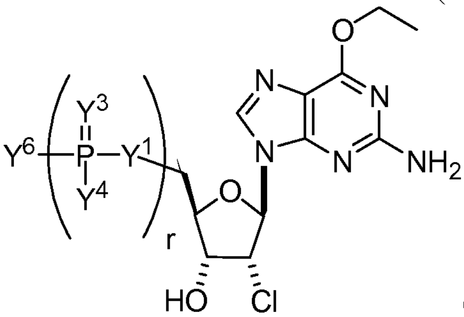

(BB-
206),

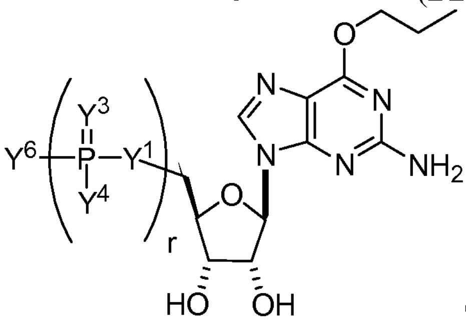
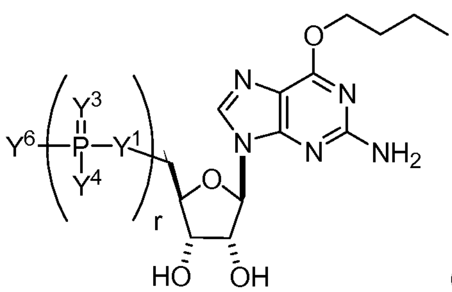

(BB- 207),

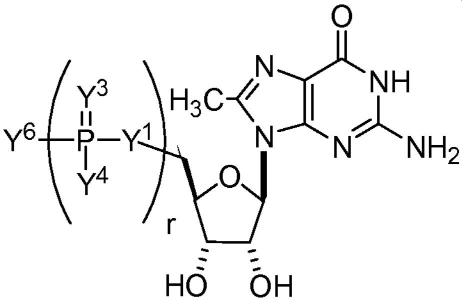
 (BB- 208),  (BB-

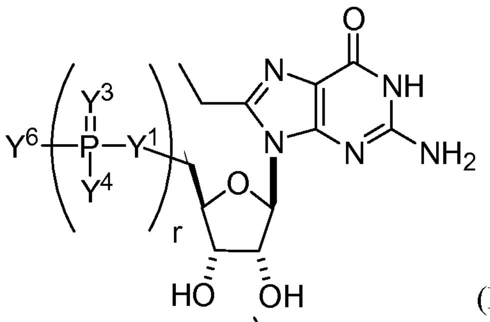
209),  (BB- 210),

 (BB- 211),  (BB-

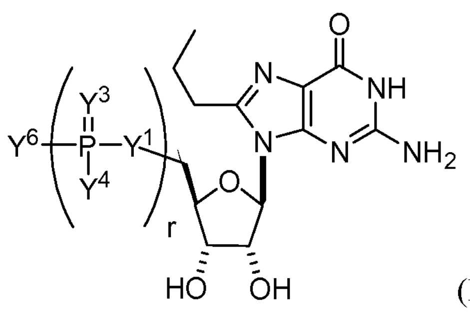
212),  (BB- 213),  (BB-

214),  (BB- 215),  (BB-

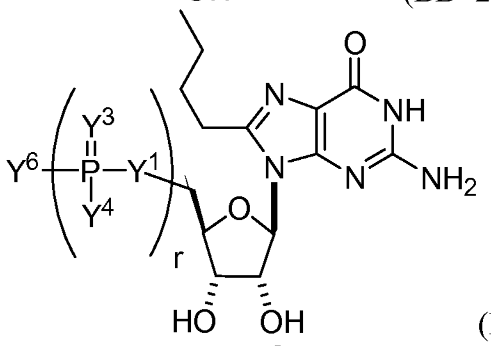
216),  (BB- 217),



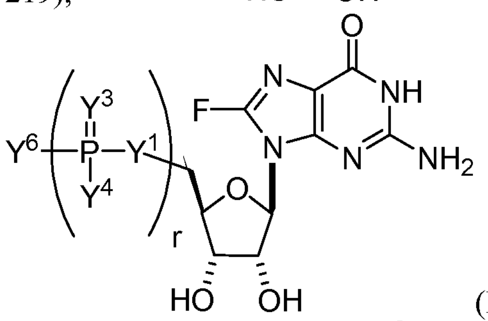
(BB- 218),



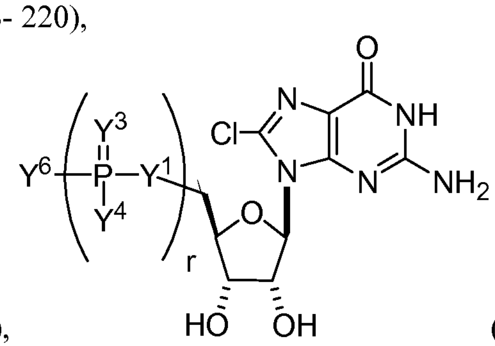
(BB- 219),



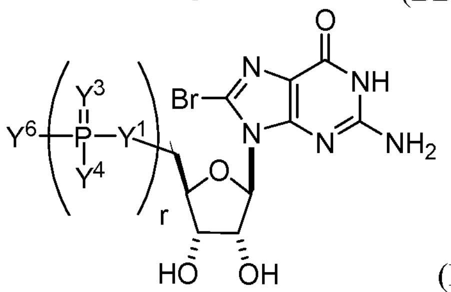
(BB- 220),



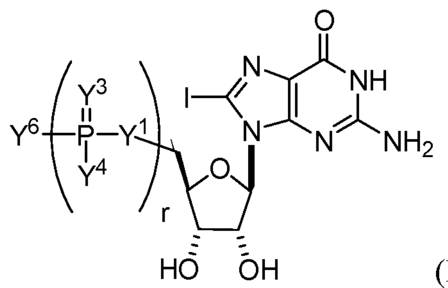
(BB- 221),



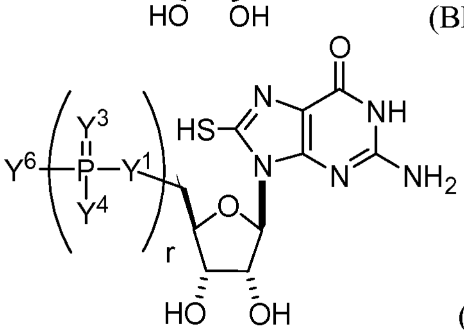
(BB- 222),



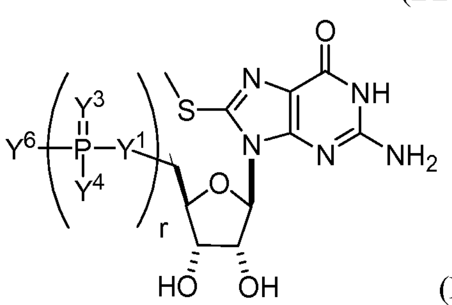
(BB- 223),



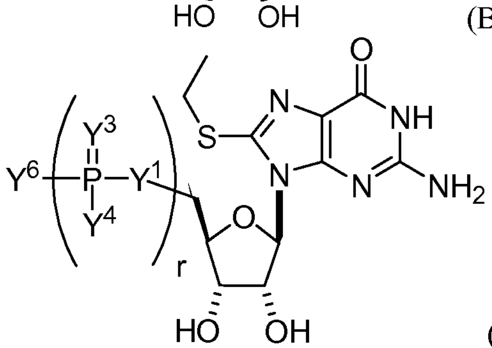
(BB- 224),



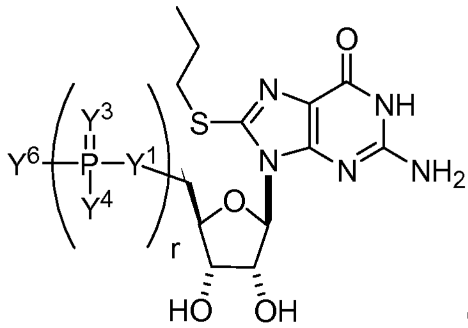
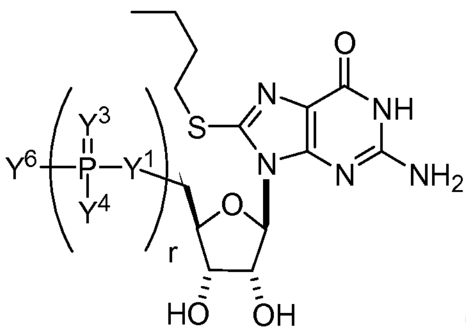
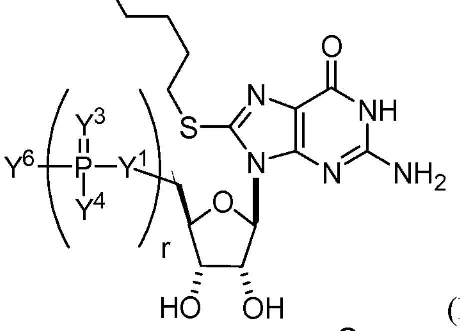
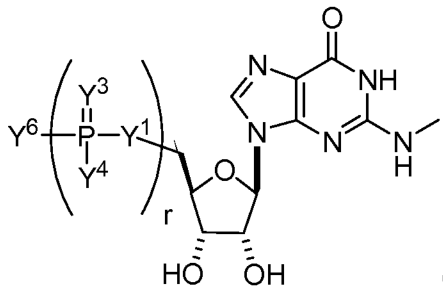
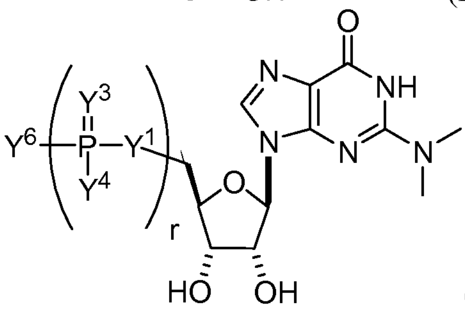
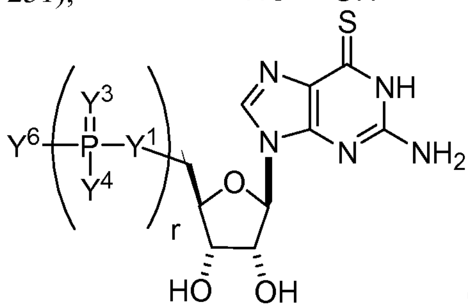
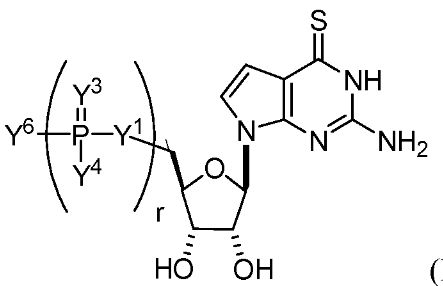
(BB- 225),

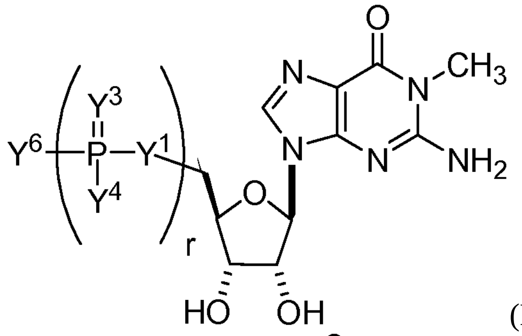


(BB- 226),

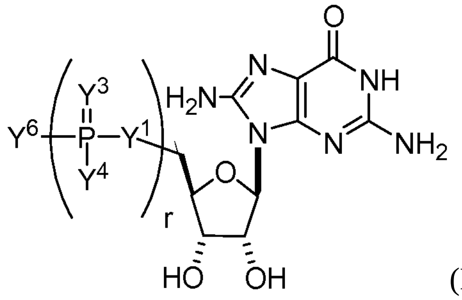


(BB- 227),


(BB- 228),

(BB- 229),

(BB- 230),

(BB- 231),

(BB- 232),

(BB- 233),

(BB- 234),



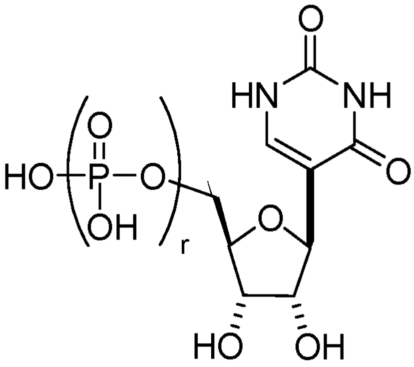
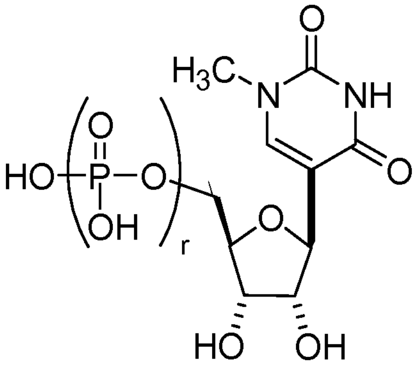
(BB- 235), and

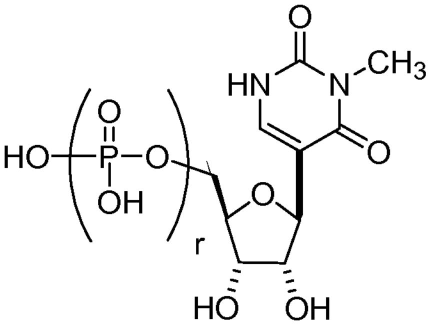
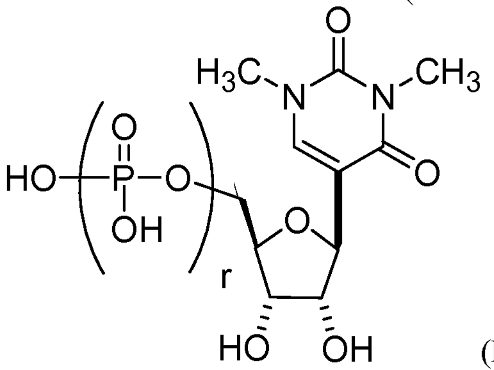


(BB- 236),

[00282] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein Yl, Y3, Y4, Y6, 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)).

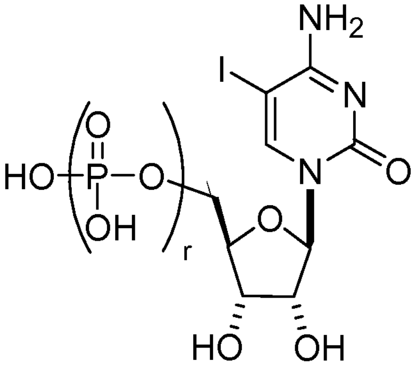
[00283] In some embodiments, the chemical modification can include replacement of C group at C-5 of the ring (e.g., for a pyrimidine nucleoside, such as cytosine or uracil) with N (e.g., replacement of the >CH group at C-5 with >NRN1 group, wherein RNl is H or optionally substituted alkyl). For example, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, can be:

 (BB- 237) or  (BB- 238) or

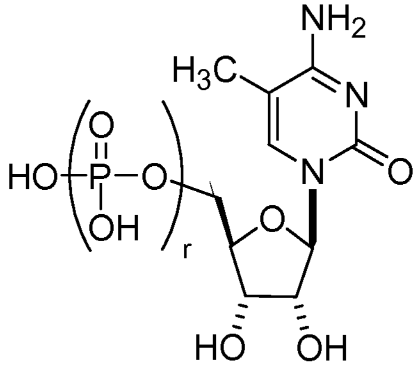
 (BB- 239) or  (BB-240),

[00284] 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).

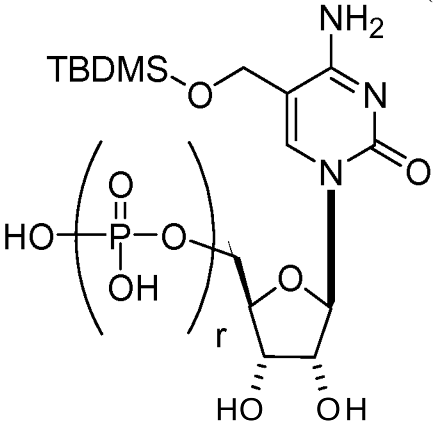
[00285] In another embodiment, the chemical modification can include replacement of the hydrogen at C-5 of cytosine with halo (e.g., Br, CI, F, or I) or optionally substituted alkyl (e.g., methyl). For example, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, can be:



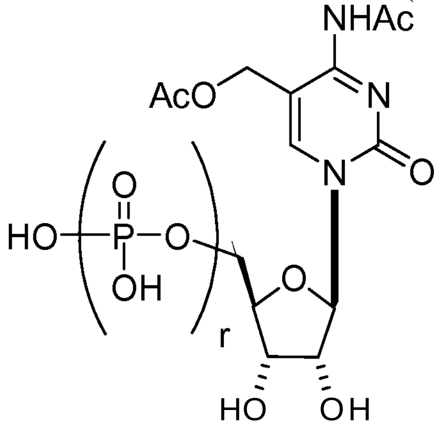
(BB- 241) or



(BB- 242) or



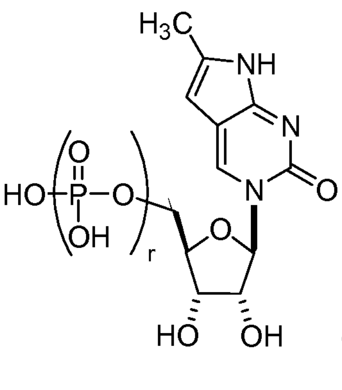
(BB- 243) or



(BB- 244),

[00286] 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).

[00287] In yet a further embodiment, the chemical modification can include a fused ring that is formed by the NH2 at the C-4 position and the carbon atom at the C-5 position. For example, the building block molecule, which may be incorporated into a polynucleotide, primary construct, or mmRNA, can be:



(BB- 245),

[00288] 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

[00289] The modified nucleosides and nucleotides (e.g., building block molecules), which may be incorporated into a polynucleotide, primary construct, or mmRNA (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 CI-6 alkyl; optionally substituted CI-6 alkoxy; optionally substituted C6-10 aryloxy; optionally substituted C3-8 cycloalkyl; optionally substituted C3-8 cycloalkoxy; optionally substituted C6-10 aryloxy; optionally substituted C6-10 aryl-Cl-6 alkoxy, optionally substituted Cl-12 (heterocyclyl)oxy; a sugar (e.g., ribose, pentose, or any described herein); a polyethyleneglycol (PEG), -0(CH2CH20)nCH2CH20R, 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 C1-6 alkylene or CI-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

[00290] 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 a-L-threofuranosyl-(3' $\rightarrow$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, primary construct, or mrnRNA molecule can include nucleotides containing, e.g., arabinose, as the sugar.

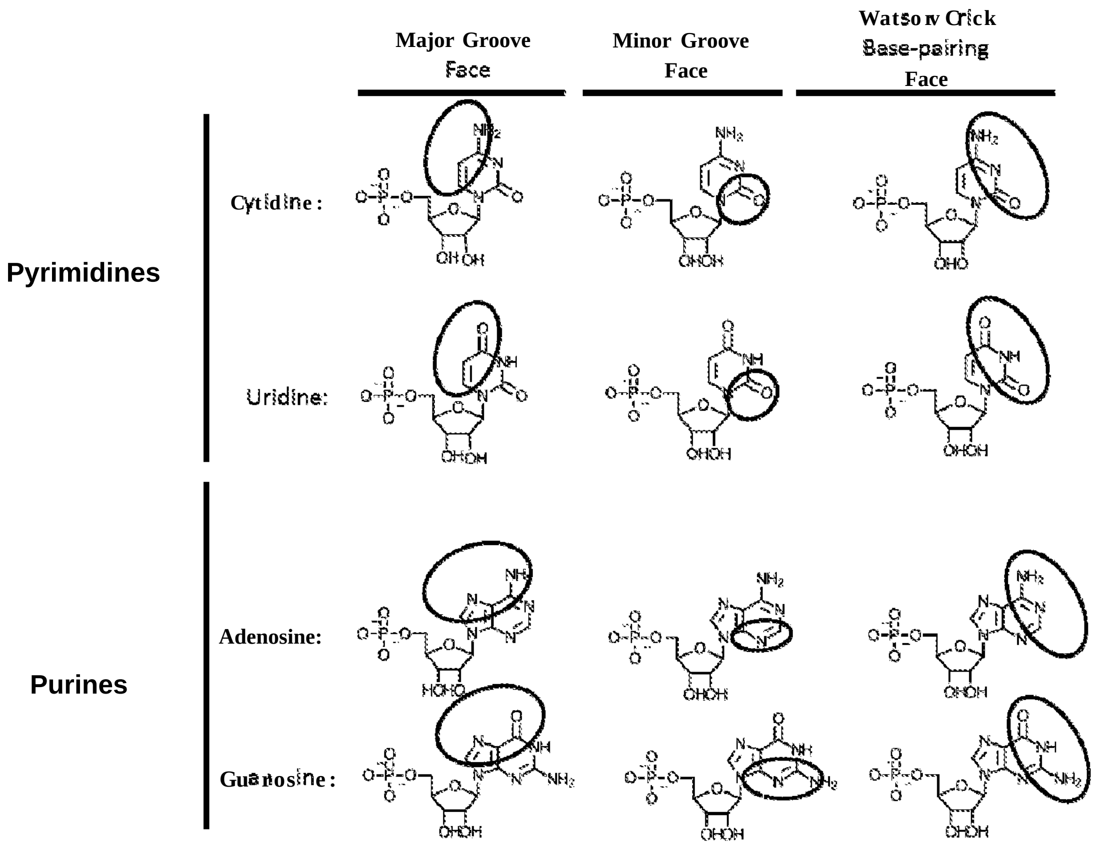
Modifications on the Nucleobase

[00291] 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 a 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. 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).

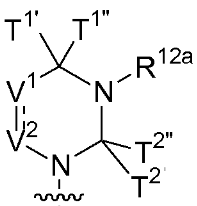
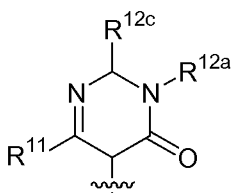
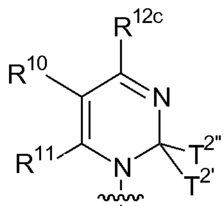
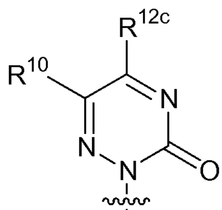
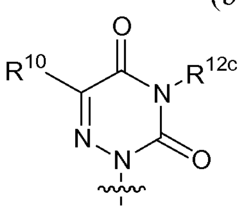
[00292] The modified nucleotide base pairing encompasses not only the standard adenosine-thymine, adenosine-uracil, or guanosine-cytosine base pairs, but also base pairs formed between nucleotides and/or modified nucleotides comprising non-standard or modified bases, wherein the arrangement of hydrogen 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.

[00293] 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 polynucleotides, primary constructs, or mmRNA molecules having enhanced properties, e.g., resistance to nucleases through disruption of the binding of a major groove binding partner. Table 8 below identifies the chemical faces of each canonical nucleotide. Circles identify the atoms comprising the respective chemical regions.

Table 8



[00294] In some embodiments, B is a modified uracil. Exemplary modified uracils include those having Formula (bl)-(b5):


(b1),

(b2),

(b3),

(b4), or

(b5),

[00295] or a pharmaceutically acceptable salt or stereoisomer thereof,

[00296] Wherein ⩦ is a single or double bond; each of T1', Tl", T2', and T2" is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of T1' and Tl" or the combination of T2' and T2" join together (e.g., as in T2) to form O (oxo), S (thio), or Se (seleno); each of VI and V2 is, independently, O, S, N(RVb)nv, or C(RVb)nv, wherein nv is an integer from **0** to 2 and each RVb 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 alkynyloxy (e.g., optionally substituted with any substituent described herein, such as those selected from (1)-(21) for alkyl);

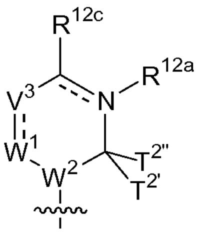
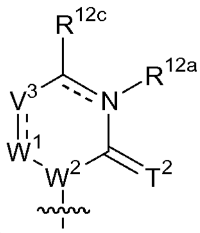
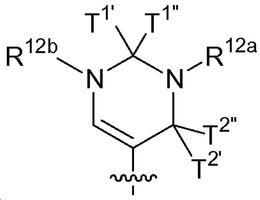
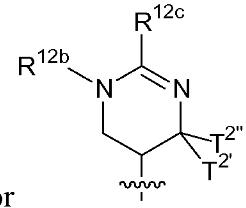
[00297] RIO is H, halo, optionally substituted amino acid, hydroxy, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aminoalkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, 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;

[00298] Rl 1 is H or optionally substituted alkyl;

[00299] R12a 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

[00300] R12c 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.

Other exemplary modified uracils include those having Formula (b6)-(b9):

 (b6),  (b7),  (b8), or  (b9),

[00301] or a pharmaceutically acceptable salt or stereoisomer thereof,

[00302] wherein

[00303] == is a single or double bond;

[00304] each of T1 ', Tl", T2', and T2" is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of T1 ' and Tl" join together (e.g., as in Tl) or the combination of T2' and T2" join together (e.g., as in T2) to form O (oxo), S (thio), or Se (seleno), or each T1 and T2 is, independently, O (oxo), S (thio), or Se (seleno);

[00305] each of W1 and W2 is, independently, N(RWa)nw or C(RWa)nw, wherein nw is an integer from 0 to 2 and each RWa is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy;

[00306] each V3 is, independently, O, S, N(RVa)nv, or C(RVa)nv, wherein nv is an integer from 0 to 2 and each RVa 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 RVa and R12c 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);

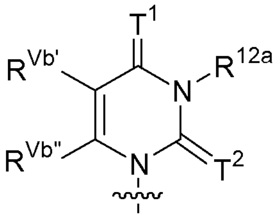
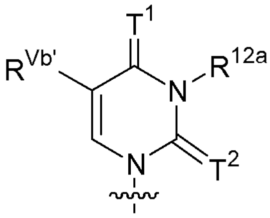
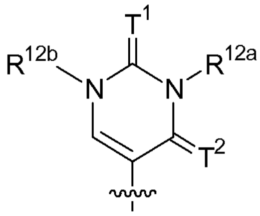
[00307] R12a is H, optionally substituted alkyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally substituted aminoalkynyl, optionally substituted carboxyalkyl (e.g., optionally substituted with hydroxy and/or an O-protecting group), optionally substituted carboxyalkoxy, optionally substituted carboxyaminoalkyl, optionally substituted carbamoylalkyl, or absent;

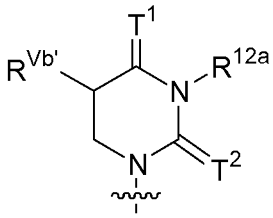
[00308] R12b is H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted aminoalkyl, optionally substituted aminoalkenyl, optionally 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,

[00309] wherein the combination of R12b and T1' or the combination of R12b and R12c can join together to form optionally substituted heterocyclyl; and

[00310] R12c 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.

[00311] Further exemplary modified uracils include those having Formula (b28)-(b31):

 (b28),  (b29),  (b30), or

 (b31),

[00312] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each of T1 and T2 is, independently, O (oxo), S (thio), or Se (seleno); each RVb' and RVb" 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., RVb' 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);

[00313] R12a 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

[00314] R12b 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),

[00315] 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.

[00316] In particular embodiments, Tl is O (oxo), and T2 is S (thio) or Se (seleno). In other embodiments, Tl is S (thio), and T2 is O (oxo) or Se (seleno). In some embodiments, RVb' is H, optionally substituted alkyl, or optionally substituted alkoxy.

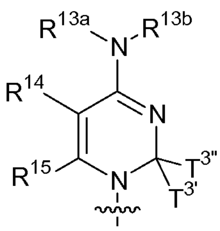
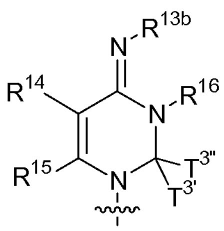
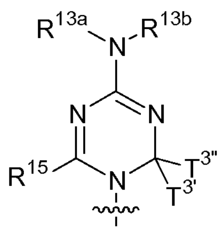
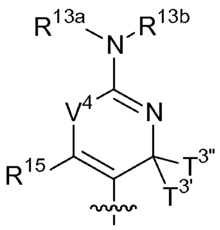
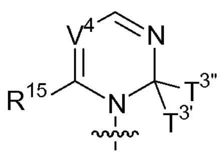
[00317] In other embodiments, each R12a and R12b is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, or optionally substituted hydroxyalkyl. In particular embodiments, R12a is H. In other embodiments, both R12a and R12b are H.

[00318] In some embodiments, each RVb' of R12b 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, R12a and RVb" are both H. In particular embodiments, Tl is O (oxo), and T2 is S (thio) or Se (seleno).

[00319] In some embodiments, RVb' is optionally substituted alkoxycarbonylalkyl or optionally substituted carbamoylalkyl.

[00320] In particular embodiments, the optional substituent for R12a, R12b, R12c, or RVa is a polyethylene glycol group (e.g., -(CH2)s2(OCH2CH2)sl(CH2)s30R', wherein si 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 Cl-20 alkyl); or an amino-polyethylene glycol group (e.g., -NRNl(CH2)s2(CH2CH20)sl(CH2)s3NRNl, wherein si 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 RN1 is, independently, hydrogen or optionally substituted CI-6 alkyl).

[00321] In some embodiments, B is a modified cytosine. Exemplary modified cytosines include compounds of Formula (bl0)-(bl4):


(b10),

(b11),

(b12),

(b13), or

(b14),

[00322] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each of T3' and T3" is, independently, H, optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy, or the combination of T3' and T3" join together (e.g., as in T3) to form O (oxo), S (thio), or Se (seleno); each V4 is, independently, O, S, N(RVc)nv, or C(RVc)nv, wherein nv is an integer from 0 to 2 and each RVc 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 R13b and RVc can be taken together to form optionally substituted heterocyclyl;

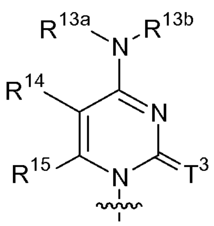
[00323] each V5 is, independently, N(RVd)nv, or C(RVd)nv, wherein nv is an integer from 0 to 2 and each RVd is, independently, H, halo, optionally substituted amino acid, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally 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., V5 is -CH or N);

[00324] each of R13a and R13b is, independently, H, optionally substituted acyl, optionally substituted acyloxyalkyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the combination of R13b and R14 can be taken together to form optionally substituted heterocyclyl;

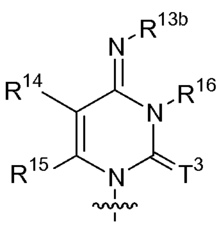
[00325] each R14 is, independently, H, halo, hydroxy, thiol, optionally substituted acyl, optionally substituted amino acid, optionally substituted alkyl, optionally substituted haloalkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl (e.g., substituted 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 aminoalkyl; and

[00326] each of R15 and R16 is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl.

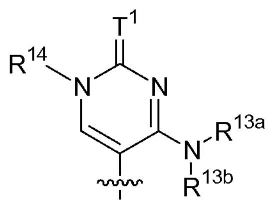
[00327] Further exemplary modified cytosines include those having Formula (b32)-(b35):



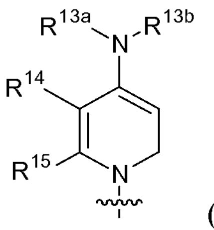
(b32),



(b33),



(b34), or



(b35),

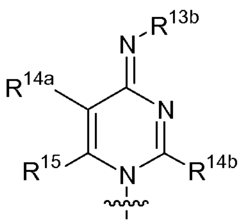
[00328] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each of T1 and T3 is, independently, O (oxo), S (thio), or Se (seleno); each of R13a and R13b is, independently, H, optionally substituted acyl, optionally substituted acyloxyalkyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the combination of R13b and R14 can be taken together to form optionally substituted heterocyclyl;

[00329] each R14 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

[00330] each of R15 and R16 is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl (e.g., R15 is H, and R16 is H or optionally substituted alkyl).

[00331] In some embodiments, R15 is H, and R16 is H or optionally substituted alkyl. In particular embodiments, R14 is H, acyl, or hydroxyalkyl. In some embodiments, R14 is halo. In some embodiments, both R14 and R15 are H. In some embodiments, both R15 and R16 are H. In some embodiments, each of R14 and R15 and R16 is H. In further embodiments, each of R13a and R13b is independently, H or optionally substituted alkyl.

[00332] Further non-limiting examples of modified cytosines include compounds of Formula (b36):



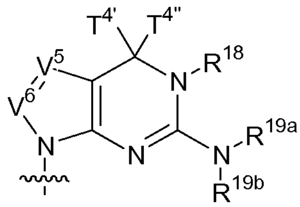
(b36)

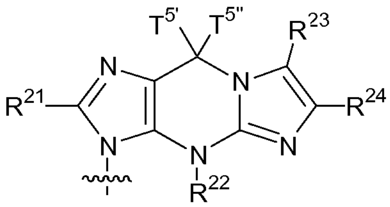
[00333] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each R13b is, independently, H, optionally substituted acyl, optionally substituted acyloxyalkyl, optionally substituted alkyl, or optionally substituted alkoxy, wherein the combination of R13b and R14b can be taken together to form optionally substituted heterocyclyl; each R14a and R14b 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

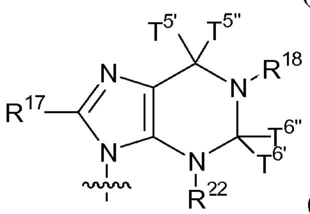
[00334] each of R15 is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl.

[00335] In particular embodiments, R14b is an optionally substituted amino acid (e.g., optionally substituted lysine). In some embodiments, R14a is H.

[00336] In some embodiments, B is a modified guanine. Exemplary modified guanines include compounds of Formula (bl5)-(bl7):


(b15),


(b16), or


(b17),

[00337] or a pharmaceutically acceptable salt or stereoisomer thereof,

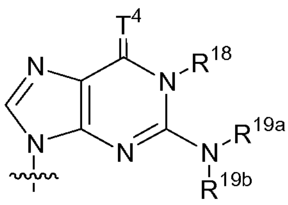
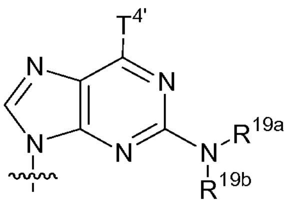
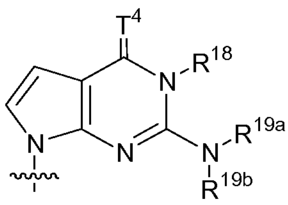
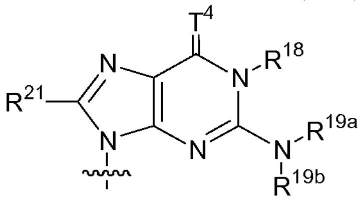
[00338] wherein

[00339] each of T4', T4", T5', T5", T6', and T6" is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy, and wherein the combination of T4' and T4" (e.g., as in T4) or the combination of T5' and T5" (e.g., as in T5) or the combination of T6' and T6" (e.g., as in T6) join together form O (oxo), S (thio), or Se (seleno);

[00340] each of V5 and V6 is, independently, O, S, N(RVd)nv, or C(RVd)nv, wherein nv is an integer from 0 to 2 and each RVd 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, or 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

[00341] each of R17, R18, R19a, R19b, R21, R22, R23, and R24 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.

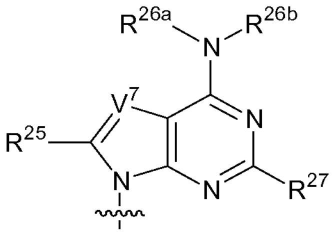
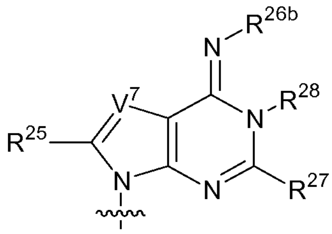
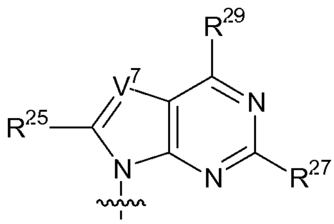
[00342] Exemplary modified guanosines include compounds of Formula (b37)-(b40):


(b37),

(b38),

(b39), or

(b40),

[00343] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each of T4' is, independently, H, optionally substituted alkyl, or optionally substituted alkoxy, and each T4 is, independently, O (oxo), S (thio), or Se (seleno); each of R18, R19a, R19b, and R21 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.

[00344] In some embodiments, R18 is H or optionally substituted alkyl. In further embodiments, T4 is oxo. In some embodiments, each of R19a and R19b is, independently, H or optionally substituted alkyl.

[00345] In some embodiments, B is a modified adenine. Exemplary modified adenines include compounds of Formula (M8)-(b20):

 (b18),  (b19), or 
(b20),

[00346] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each V7 is, independently, O, S, N(RVe)nv, or C(RVe)nv, wherein nv is an integer from 0 to 2 and each RVe 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);

[00347] each R25 is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, or optionally substituted amino;

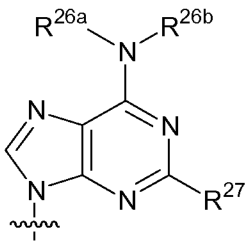
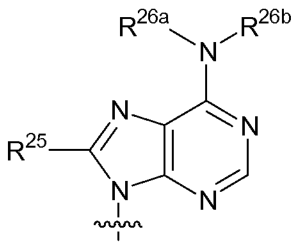
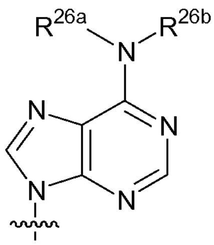
[00348] each of R26a and R26b is, independently, H, optionally substituted acyl, optionally substituted amino acid, optionally substituted carbamoylalkyl, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted hydroxyalkyl, optionally substituted hydroxyalkenyl, optionally substituted hydroxyalkynyl, optionally substituted alkoxy, or polyethylene glycol group (e.g., -(CH2)s2(OCH2CH2)sl(CH2)s30R\ wherein si 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 CI-20 alkyl); or an amino-polyethylene glycol group (e.g., -NRNl(CH2)s2(CH2CH20)sl(CH2)s3NRNl, wherein si 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 RNl is, independently, hydrogen or optionally substituted CI-6 alkyl);

[00349] each R27 is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted thioalkoxy or optionally substituted amino;

[00350] each R28 is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl; and

[00351] each R29 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.

[00352] Exemplary modified adenines include compounds of Formula (b41)-(b43):


(b41),

(b42), or

(b43),

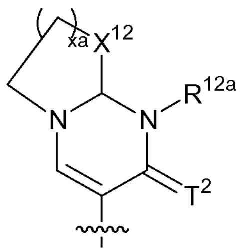
[00353] or a pharmaceutically acceptable salt or stereoisomer thereof, wherein each R25 is, independently, H, halo, thiol, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted thioalkoxy, or optionally substituted amino; each of R26a and R26b 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., -(CH2)s2(OCH2CH2)sl(CH2)s30R\ wherein si 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 CI-20 alkyl); or an amino-polyethylene glycol group (e.g., -NRNl(CH2)s2(CH2CH20)sl(CH2)s3NRNl, wherein si 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 RN1 is, independently, hydrogen or optionally substituted CI-6 alkyl); and

[00354] each R27 is, independently, H, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted thioalkoxy, or optionally substituted amino.

[00355] In some embodiments, R26a is H, and R26b is optionally substituted alkyl. In some embodiments, each of R26a and R26b is, independently, optionally substituted alkyl. In particular embodiments, R27 is optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy. In other embodiments, R25 is optionally substituted alkyl, optionally substituted alkoxy, or optionally substituted thioalkoxy.

[00356] In particular embodiments, the optional substituent for R26a, R26b, or R29 is a polyethylene glycol group (e.g., -(CH2)s2(OCH2CH2)sl(CH2)s30R\ wherein si 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 Cl-20 alkyl); or an amino-polyethylene glycol group (e.g., -NRNl(CH2)s2(CH2CH20)sl(CH2)s3NRNl, wherein si 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 RN1 is, independently, hydrogen or optionally substituted CI-6 alkyl).

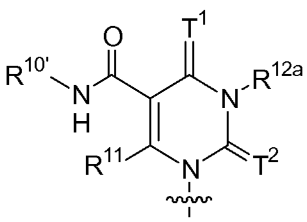
[00357] In some embodiments, B may have Formula (b21):



(b21),

[00358] wherein X12 is, independently, O, S, optionally substituted alkylene (e.g., methylene), or optionally substituted heteroalkylene, xa is an integer from 0 to 3, and R12a and T2 are as described herein.

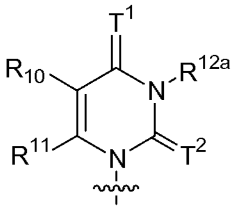
[00359] In some embodiments, B may have Formula (b22):



(b22),

[00360] wherein R10' 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 Rl 1, R12a, Tl, and T2 are as described herein.

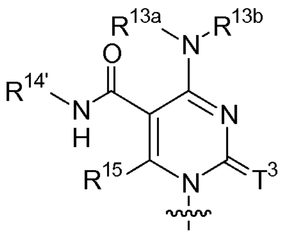
[00361] In some embodiments, B may have Formula (b23):



(b23),

[00362] wherein R10 is optionally substituted heterocyclyl (e.g., optionally substituted furyl, optionally substitued thienyl, or optionally substitued pyrrolyl), optionally substituted aryl (e.g., optionally substituted phenyl or optionally substituted naphthyl), or any substituent described herein (e.g., for R10) ;and wherein Rl 1 (e.g., H or any substituent described herein), R12a (e.g., H or any substituent described herein), T1 (e.g., oxo or any substituent described herein), and T2 (e.g., oxo or any substituent described herein) are as described herein.

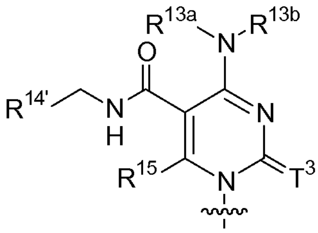
[00363] In some embodiments, B may have Formula (b24):



(b24),

[00364] wherein R14' 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 alkoxycarbonylalkenyl, optionally substituted alkoxycarbonylalkynyl, optionally substituted alkoxycarbonylalkyl, optionally substituted alkoxycarbonylalkoxy, optionally substituted carboxyalkoxy, optionally substituted carboxyalkyl, or optionally substituted carbamoylalkyl, and R13a, R13b, R15, and T3 are as described herein.

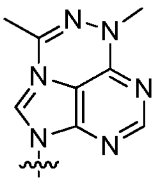
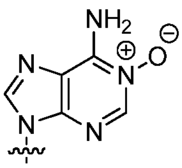
[00365] In some embodiments, B may have Formula (b25):



(b25),

[00366] wherein R14' is optionally substituted heterocyclyl (e.g., optionally substituted furyl, optionally substitued thienyl, or optionally substitued pyrrolyl), optionally substituted aryl (e.g., optionally substituted phenyl or optionally substituted naphthyl), or any substituent described herein (e.g., for R14 or R14'); and wherein R13a (e.g., H or any substituent described herein), R13b (e.g., H or any substituent described herein), R15 (e.g., H or any substituent described herein), and T3 (e.g., oxo or any substituent described herein) are as described herein.

[00367] 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).

[00368] 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 (s2U), 4-thio-uridine (s4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), 3-methyl-uridine (m3U), 5-methoxy-uridine (mo5U), uridine 5-oxyacetic acid (cmo5U), uridine 5-oxyacetic acid methyl ester (mcmo5U), 5-carboxymethyl-uridine (cm5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm5U), 5-carboxyhydroxymethyl-uridine methyl ester (mchm5U), 5-methoxycarbonylmethyl-uridine (mcm5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm5s2U), 5-aminomethyl-2-thio-uridine (nm5s2U), 5-methylaminomethyl-uridine (mnm5U), 5-methylaminomethyl-2-thio-uridine (mnm5s2U), 5-methylaminomethyl-2-seleno-uridine (mnm5se2U), 5-carbamoylmethyl-uridine (ncm5U), 5-carboxymethylaminomethyl-uridine (cmnm5U), 5-carboxymethylaminomethyl-2-thio-uridine (cmnm5s2U), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (xm5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine **(Tm5s2U),** 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-uridine (m5U, i.e., having the nucleobase deoxythymine), 1-methylpseudouridine (ηιΐψ), 5-methyl-2-thio-uridine (m5s2U), l-methyl-4-thio-pseudouridine (m1s4ψ), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine (ηι3ψ), 2-thio-l-methyl-pseudouridine, 1-methyl-l-deaza-pseudouridine, 2-thio-l -methyl- 1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, Nl-methyl-pseudouridine (also known as 1-methylpseudouridine (m1ψ)), 3-(3-amino-3-carboxypropyl)uridine (acp3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp3 ψ), 5-(isopentenylaminomethyl)uridine (inm5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm5s2U), a-thio-uridine, 2'-0-methyl-uridine (Um), 5,2'-0-dimethyl-uridine (m5Um), 2'-O-methyl-pseudouridine (ψmι), 2-thio-2'-0-methyl-uridine (s2Um), 5-methoxycarbonylmethyl-2'-0-methyl-uridine (mcm5Um), 5-carbamoylmethyl-2'-0-methyl-uridine (ncm5Um), 5-carboxymethylaminomethyl-2'-0-methyl-uridine (cmnm5Um), 3,2'-0-dimethyl-uridine (m3Um), 5-(isopentenylaminomethyl)-2'-0-methyl-uridine (inm5Um), 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.

[00369] 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 (m3C), N4-acetyl-cytidine (ac4C), 5-formyl-cytidine (f5C), N4-methyl-cytidine (m4C), 5-methyl-cytidine (m5C), 5-halo-cytidine (e.g., 5-iodo-cytidine), 5-hydroxymethyl-cytidine (hm5C), 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine (s2C), 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 (k2C), a-thio-cytidine, 2'-0-methyl-cytidine (Cm), 5,2'-0-dimethyl-cytidine (m5Cm), N4-acetyl-2'-0-methyl-cytidine (ac4Cm), N4,2'-0-dimethyl-cytidine (m4Cm), 5-formyl-2'-0-methyl-cytidine (f5Cm), N4,N4,2'-0-trimethyl-cytidine (m42Cm), 1-thio-cytidine, 2' F ara cytidine, 2' F cytidine, and 2' OH ara cytidine.

[00370] 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 (ml A), 2-methyl-adenine (m2A), N6-methyl-adenosine (m6A), 2-methylthio-N6-methyl-adenosine (ms2m6A), N6-isopentenyl-adenosine (i6A), 2-methylthio-N6-isopentenyl-adenosine (ms2i6A), N6-(cis-hydroxyisopentenyl)adenosine (io6A), 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine (ms2io6A), N6-glycinylcarbamoyl-adenosine (g6A), N6-threonylcarbamoyl-adenosine (t6A), N6-methyl-N6-threonylcarbamoyl-adenosine (m6t6A), 2-methylthio-N6-threonylcarbamoyl-adenosine (ms2g6A), N6,N6-dimethyl-adenosine (m62A), N6-hydroxynorvalylcarbamoyl-adenosine (hn6A), 2-methylthio-N6-hydroxynorvalylcarbamoyl-adenosine (ms2hn6A), N6-acetyl-adenosine (ac6A), 7-methyl-adenine, 2-methylthio-adenine, 2-methoxy-adenine, a-thio-adenosine, 2'-0-methyl-adenosine (Am), N6,2'-0-dimethyl-adenosine (m6Am), N6,N6,2'-0-trimethyl-adenosine (m62Am), l,2'-0-dimethyl-adenosine (ml Am), 2'-0-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.

[00371] In some embodiments, the modified nucleobase is a modified guanine. Exemplary nucleobases and nucleosides having a modified guanine include inosine (I), 1-methyl-inosine (mil), wyosine (imG), methylwyosine (mimG), 4-demethyl-wyosine (imG-14), isowyosine (imG2), wybutosine (yW), peroxywybutosine (o2yW), hydroxywybutosine (OHyW), undermodified hydroxywybutosine (OHyW*), 7-deaza-guanosine, queuosine (Q), epoxyqueuosine (oQ), galactosyl-queuosine (galQ), mannosyl-queuosine (manQ), 7-cyano-7-deaza-guanosine (preQO), 7-aminomethyl-7-deaza-guanosine (preQl), 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 (m7G), 6-thio-7-methyl-guanosine, 7-methyl-inosine, 6-methoxy-guanosine, 1-methyl-guanosine (mlG), N2-methyl-guanosine (m2G), N2,N2-dimethyl-guanosine (m22G), N2,7-dimethyl-guanosine (m2,7G), N2, N2,7-dimethyl-guanosine (m2,2,7G), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, l-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, N2,N2-dimethyl-6-thio-guanosine, a-thio-guanosine, 2'-0-methyl-guanosine (Gm), N2-methyl-2'-0-methyl-guanosine (m2Gm), N2,N2-dimethyl-2'-0-methyl-guanosine (m22Gm), 1-methyl-2'-0-methyl-guanosine (mlGm), N2,7-dimethyl-2'-0-methyl-guanosine (m2,7Gm), 2'-0-methyl-inosine (Im), 1,2'-O-dimethyl-inosine (mllm), and 2'-0-ribosylguanosine (phosphate) (Gr(p)).

[00372] 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[l,5-a]l,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).

Modifications on the Internucleoside Linkage

[00373] The modified nucleotides, which may be incorporated into a polynucleotide, primary construct, or mmRNA 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).

[00374] The a-thio substituted phosphate moiety is provided to confer stability to RNA and DNA polymers through the unnatural phosphorothioate backbone linkages. Phosphorothioate DNA and RNA have increased nuclease resistance and subsequently a longer half-life in a cellular environment. Phosphorothioate linked polynucleotides, primary constructs, or mmRNA molecules are expected to also reduce the innate immune response through weaker binding/activation of cellular innate immune molecules.

[00375] In specific embodiments, a modified nucleoside includes an alpha-thio-nucleoside (e.g., 5'-0-(l-thiophosphate)-adenosine, 5'-0-(l-thiophosphate)-cytidine (a-thio-cytidine), 5'-0-(l-thiophosphate)-guanosine, 5'-0-(l-thiophosphate)-uridine, or 5'-0-(l-thiophosphate)-pseudouridine).

[00376] 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

[00377] The polynucleotides, primary constructs, and mmRNA 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 (la), (Ia-l)-(Ia-3), (Ib)-(If), (Ila)-(IIp), (IIb-1), (IIb-2), (IIc-l)-(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 (bl)-(b43) or any other described herein).

Cytotoxic Nucleosides

[00378] In one embodiment, the polynucleotides, primary constructs or mmRNA of the present invention may incorporate one or more cytotoxic nucleosides. For example, cytotoxic nucleosides may be incorporated into polynucleotides, primary constructs or mmRNA such as bifunctional modified RNAs or mRNAs. Cytotoxic nucleoside anti-cancer agents include, but are not limited to, adenosine arabinoside, cytarabine, cytosine arabinoside, 5-fluorouracil, fludarabine, floxuridine, FTORAFUR® (a combination of tegafur and uracil), tegafur ((RS)-5-fluoro-l-(tetrahydrofuran-2-yl)pyrimidine-2,4(lH,3H)-dione), and 6-mercaptopurine.

[00379] A number of cytotoxic nucleoside analogues are in clinical use, or have been the subject of clinical trials, as anticancer agents. Examples of such analogues include, but are not limited to, cytarabine, gemcitabine, troxacitabine, decitabine, tezacitabine, 2'-deoxy-2'-methylidenecytidine (DMDC), cladribine, clofarabine, 5-azacytidine, 4'-thio-aracytidine, cyclopentenylcytosine and 1-(2-C-cyano-2-deoxy-beta-D-arabino-pentofuranosyl)-cytosine. Another example of such a compound is fludarabine phosphate. These compounds may be administered systemically and may have side effects which are typical of cytotoxic agents such as, but not limited to, little or no specificity for tumor cells over proliferating normal cells.

[00380] A number of prodrugs of cytotoxic nucleoside analogues are also reported in the art. Examples include, but are not limited to, N4-behenoyl-l-beta-D-arabinofuranosylcytosine, N4-octadecyl- 1-beta-D-arabinofuranosylcytosine, N4-palmitoyl- 1-(2-C-cyano-2-deoxy-beta-D-arabino-pentofuranosyl) cytosine, and P-4055 (cytarabine 5'-elaidic acid ester). In general, these prodrugs may be converted into the active drugs mainly in the liver and systemic circulation and display little or no selective release of active drug in the tumor tissue. For example, capecitabine, a prodrug of 5'-deoxy-5-fluorocytidine (and eventually of 5-fluorouracil), is metabolized both in the liver and in the tumor tissue. A series of capecitabine analogues containing "an easily hydrolysable radical under physiological conditions" has been claimed by Fujiu et al. (U.S. Pat. No. 4,966,891) and is herein incorporated by reference. The series described by Fujiu includes N4 alkyl and aralkyl carbamates of 5'-deoxy-5-fluorocytidine and the implication that these compounds will be activated by hydrolysis under normal physiological conditions to provide 5'-deoxy-5-fluorocytidine.

[00381] A series of cytarabine N4-carbamates has been by reported by Fadl et al (Pharmazie. 1995, 50, 382-7, herein incorporated by reference) in which compounds were designed to convert into cytarabine in the liver and plasma. WO 2004/041203, herein incorporated by reference, discloses prodrugs of gemcitabine, where some of the prodrugs are N4-carbamates. These compounds were designed to overcome the gastrointestinal toxicity of gemcitabine and were intended to provide gemcitabine by hydrolytic release in the liver and plasma after absorption of the intact prodrug from the gastrointestinal tract. Nomura et al (Bioorg Med. Chem. 2003, 11, 2453-61, herein incorporated by reference) have described acetal derivatives of l-(3-C-ethynyl -β-D-ribo-pentofaranosyl) cytosine which, on bioreduction, produced an intermediate that required further hydrolysis under acidic conditions to produce a cytotoxic nucleoside compound.

[00382] Cytotoxic nucleotides which may be chemotherapeutic also include, but are not limited to, pyrazolo [3,4-D]-pyrimidines, allopurinol, azathioprine, capecitabine, cytosine arabinoside, fluorouracil, mercaptopurine, 6-thioguanine, acyclovir, ara-adenosine, ribavirin, 7-deaza-adenosine, 7-deaza-guanosine, 6-aza-uracil, 6-aza-cytidine, thymidine ribonucleotide, 5-bromodeoxyuridine, 2-chloro-purine, and inosine, or combinations thereof.

[00383]

[00384] The The modified nucleosides and nucleotides used in the synthesis of polynucleotides, primary constructs, and mmRNA 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.

[00385] 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., 1H or 13C) infrared spectroscopy, spectrophotometry (e.g., UV-visible), or mass spectrometry, or by chromatography such as high performance liquid chromatography (HPLC) or thin layer chromatography.

[00386] Preparation of polypeptides, primary constructs, and mmRNA 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.

[00387] 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.

[00388] 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.

[00389] 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.

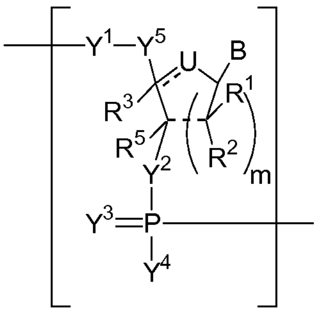
[00390] The polypeptides, primary constructs, and mmRNA of the invention may or may not be uniformly modified along the entire length of the molecule. For example, one or more or all types of nucleotide (e.g., purine or pyrimidine, or any one or more or all of A, G, U, C) may or may not be uniformly modified in a polynucleotide of the invention, or in a given predetermined sequence region thereof (e.g. one or more of the sequence regions represented in Figure 1). 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.

[00391] Different sugar modifications, nucleotide modifications, and/or internucleoside linkages (e.g., backbone structures) may exist at various positions in the polynucleotide, primary construct, or mmRNA. 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, primary construct, or mmRNA such that the function of the polynucleotide, primary construct, or mmRNA is not substantially decreased. A modification may also be a 5' or 3' terminal modification. The polynucleotide, primary construct, or mmRNA 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%).

[00392] In some embodiments, the polynucleotide, primary construct, or mmRNA 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, primary construct, or mmRNA 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, primary construct, or mmRNA 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

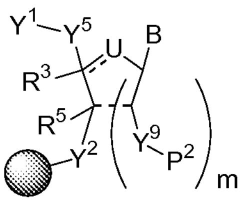
1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 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).

[00393] In some embodiments, the present disclosure provides methods of synthesizing a polynucleotide, primary construct, or mmRNA (e.g., the first region, first flanking region, or second flanking region) including n number of linked nucleosides having Formula (Ia-1):



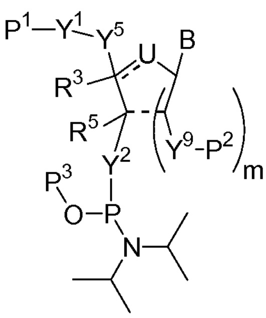
(Ia-1), comprising:

[00394] a) reacting a nucleotide of Formula (IV-1):



(IV-1),

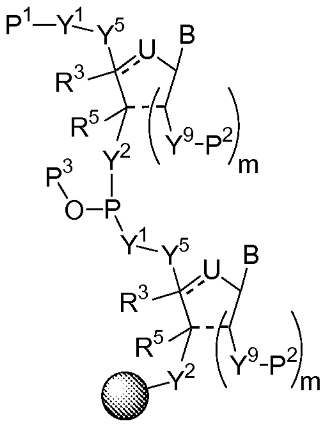
[00395] with a phosphoramidite compound of Formula (V-1):



(V-1),

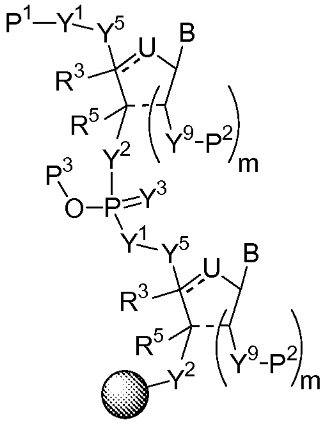
[00396] wherein Y9 is H, hydroxy, phosphoryl, pyrophosphate, sulfate, amino, thiol, optionally substituted amino acid, or a peptide (e.g., including from 2 to 12 amino acids); and each PI, P2, and P3 is, independently, a suitable protecting group; and  denotes a solid support;

[00397] to provide a polynucleotide, primary construct, or mmRNA of Formula (VI-1):



(VI-1), and

[00398] b) oxidizing or sulfurizing the polynucleotide, primary construct, or mmRNA of Formula (V) to yield a polynucleotide, primary construct, or mmRNA of Formula (VII-1):



(VII-1), and

[00399] c) removing the protecting groups to yield the polynucleotide, primary construct, or mmRNA of Formula (la).

[00400] 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 (e.g., mmRNA molecule) 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, primary construct, or mmRNA is translatable.

[00401] Other components of polynucleotides, primary constructs, and mmRNA 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, primary constructs, and mmRNA containing a Kozak sequence.

Combinations of Nucleotides in mmRNA

[00402] Further examples of modified nucleotides and modified nucleotide combinations are provided below in Table 9. These combinations of modified nucleotides can be used to form the polypeptides, primary constructs, or mmRNA of the invention. Unless otherwise noted, the modified nucleotides may be completely substituted for the natural nucleotides of the modified nucleic acids or mmRNA 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 9

Modified Nucleotide	Modified Nucleotide Combination
α-thio-cytidine	α-thio-cytidine/5-iodo-uridine
	α-thio-cytidine/N1-methyl-pseudouridine
	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
pyrrolo-cytidine	pyrrolo-cytidine/5-iodo-uridine

	pyrrolo-cytidine/N 1-methyl-pseudouridine
	pyrrolo-cytidine/a-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/N 1-methyl-pseudouridine
	5-methyl-cytidine/a-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 /N 1-methyl-pseudouridine
	N4-acetyl-cytidine /a-thio-uridine
	N4-acetyl-cytidine 15-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

[00403] Further examples of modified nucleotide combinations are provided below in Table 10. These combinations of modified nucleotides can be used to form the polypeptides, primary constructs, or mmRNA of the invention.

Table 10

Modified Nucleotide	Modified Nucleotide Combination
modified cytidine having	modified cytidine with (b10)/pseudouridine

one or more nucleobases of Formula (blO)	modified cytidine with (blO)/Nl-methyl-pseudouridine
	modified cytidine with (bl0)/5-methoxy -uridine
	modified cytidine with (bl0)/5-methyl-uridine
	modified cytidine with (bl0)/5-bromo-uridine
	modified cytidine with (b 10)/2-thio-uridine
	about 50% of cytidine substituted with modified cytidine (b 10)/ 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)/Nl-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 (bl)	modified uridine with (bl)/ N4-acetyl-cytidine
	modified uridine with (bl)/ 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

[00404] In some embodiments, at least *25%* of the cytosines are replaced by a compound of Formula (M0)-(bl4) (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%o,** 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%).

[00405] In some embodiments, at least 25%> of the uracils are replaced by a compound of Formula (bl)-(b9) (e.g., at least about 30%>, at least about 35%>, at least about 40%>, at least about **45%o,** 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%).

[00406] In some embodiments, at least 25% of the cytosines are replaced by a compound of Formula (bl0)-(bl4), and at least 25% of the uracils are replaced by a compound of Formula (bl)-(b9) (e.g., at least about 30%, at least about 35%, at least about 40%, at least about 45%o, 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%).

EXAMPLES

[00407] Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

[00408] The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, biochemistry, recombinant DNA techniques and pharmacology, within the skill of the art. Such techniques are explained fully in the literature. See, e.g., T.E. Creighton, *Proteins: Structures and Molecular Properties* (W.H. Freeman and Company, 1993); A.L. Lehninger, *Biochemistry* (Worth Publishers, Inc., current addition); Sambrook, et al, *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *Methods In Enzymology* (S. Colowick and N. Kaplan eds., Academic Press, Inc.); *Remington's Pharmaceutical Sciences,* 18th Edition (Easton, Pennsylvania: Mack Publishing Company, 1990); Carey and Sundberg *Advanced Organic Chemistry 3rd Ed.* (Plenum Press) Vols A and B(l 992).

Example 1: DNA template with GCSF gene

[00409] The plasmid pJ344:91543-TC-GCSF includes the coding sequence for GCSF (granulyte colony stimulating factor) and 141 nucleotide sequence coding for Poly A Tail and an XBal restriction endonuclease recognition site immediately downstream of the poly A tail sequence tract. A plasmid map is shown in Figure 5, and a plasmid DNA sequence is shown in Figure 6.

[00410] An E. coli strain DH10B harboring above plasmid was grown in 250 ml animal-free broth with ampicillin. This process generated 1107μg of plasmid DNA. The plasmid

DNA was isolated and purified according to DNA2.0 standard operating procedures. Plasmid DNA yield and DNA homogeneity was determined using agarose gel electrophoresis. The insert (but not vector backbone) was sequence verified in both orientations using a capillary electrophoresis DNA analyzer.

Example 2: Linearization of DNA template.

[00411] The plasmid DNA template was linearized using the restriction endonuclease Xbal. The Xbal restriction digest reaction conditions were as follows:

Water	25573.6 µL
Plasmid DNA (1145 µg/µL	1879.9 µï.
BSA (IOOX)	430.5 µL
Buffer 4 (1OX)	4305.0 µL
Xbal (20000 U/ml	861 µL
Total vol	43050 µï.

[00412] The reaction proceeded at 37°C overnight. The linearized plasmid was diafiltered into lOmM Tris HCl pH 7.5 and concentrated to 959ng/uL prior to use using lOOkDa MWCO Amicon spin filters (EMD Millipore).

Example 3: In vitro transcription

[00413] The non-amplified, linearized DNA plasmid was used as a template for *in vitro* transcription. A lmL transcription reaction was performed utilizing 250 µg of plasmid template. Nucleotides ATP, GTP, 5mCTP (5-methyl cytosine triphosphate) and N-l-methylpseudouridine triphosphate were added at 7.5 mM each. RNase inhibitor (1000 U), inorganic pyrophosphatase (1U) and T7 RNA polymerase (7000U) were added. The final buffer conditions (IX) were as follows: 40 mM Tris HCl pH8, 40 mM magnesium acetate, 5 mM dithiothreitol (DTT), and 1 mM spermidine.

[00414] The *in vitro* transcription reaction proceeded for 4 hours at 37°C under constant mixing. The total reaction yield was 5.4 mg RNA transcript. A 400uL (40%) portionof the reaction was diafiltered into water and concentrated to 622 ng^L using lOOkDa MWCO Amicon spin filters (EMD Millipore).

[00415] The results demonstrate successful, milligram scale production of RNA transcript with *in vitro* transcription using a non-PCT amplified, linearized plasmid DNA template.

Example 4: Oligo dT removal of DNA template.

[00416] The full length poly A containing RNA transcript was purified from truncated RNA, DNA template, and residual enzymes using oligo dT chromatography. A 20 mer polythymidine Sepharose (3 ml) was packed in a 5mL SPE column on a solid phase extraction vacuum manifold. The RNA transcript (2.18 mg) was applied to column, followed by washing and elution. The oligo dT purified RNA transcript was diafiltered into water and concentrated to 1.22 mg/mL using 100 kDa MWCO Amicon spin filters (EMD Millipore). Approximately 1.82 mg of 2.18 mg was recovered (83%) as determined by Bioanalyzer gel elctrophoresis.

Example 5: Characterization of uncapped RNA transcript, pre and post oligo dT.

[00417] The RNA transcript was analyzed using a Bioanalyzer gel electrohphoresis. Sample taken both before and after oligo dT purification were analyzed. The results are shown in the electopherograms in Figures 7A and 7B.

[00418] The results demonstrate recovery of >80% RNA transcript using oligo dT purification and virtual elimination of impurities.

Example 6: Capping of RNA transcript.

[00419] A 5' cap was enzymatically added to the RNA transcript. Approximately 0.98 mg of oligo dT purified RNA transcript (1225ng/uL) was capped to obtain a 5' Cap 1 structure. GTP was added at a final reaction concentration of 0.9 mM. SAM (S-adenosyl-L-methionine) was added at a final reaction concentration of 0.4 mM. 907 units of RNase inhibitor, 3630 units of 2'O-methyltransferase, and 363 units of vaccinia guanylyltransferase were added to the reaction. The final 1x buffer conditions consist of the following: 50mM Tris HCl pH 8, 5mM KCl, ImM MgC12, and ImM dithiothreitol. Final reaction volume was 1070 µL. The reaction proceeded at 37°C for 2 hours under constant mixing. Reaction setup for capped RNA (lot 12-04-155-C) is shown below.

mRNA Capping Worksheet (Molar Calc)										
				10X Capping Buffer		Conc (mM)				
Total RNA	1.5	mg		Tris HCl pH8		500				
mRNA length	900	Bases		KCl		50				
mRNA MW	306,000	Da		MgCl2		10				
mRNA (amt)	0.0049	umol		DTT		10				
mRNA amount	4.90	nmol								
Target Final mRNA Conc in RXN	3.64	uM								
Target RXN Volume	1348	uL								
mRNA Volume	971	uL								
Target mRNA Conc:	1.55	mg/mL								
Target mRNA Conc:	1546	ng/uL								
Acceptable Range (+/- 5%)	1468	ng/uL	to 1623	ng/uL						
	Stock conc		Volume		Stoichiometry				Effective Conc:	
10X Capping Buffer			134.8 uL							
GTP	100	mM	13.5 uL		275.2	:1 mol/mol			1.00	mM
SAM	20	mM	33.8 uL		137.8	:1 mol/mol			0.50	mM
	Stock conc		Volume				Total Units		Effective Conc:	
Rnase Inhibitor	40000	U/mL	33.7 uL		275	U/nmol	1348.0	Units	1000	U/mL
2'Ome Transferase	50000	U/mL	107.9 uL		1100.2	U/nmol	5393.1	Units	4001	U/mL
Vaccinia Cap	10000	U/mL	53.9 uL		109.9	U/nmol	538.7	Units	400	U/mL
	Mastermix Volume		378 uL							
	mRNA Volume		971 uL							
	Total Volume		1348 uL							

[00420] The projected sequence of the RNA transcript is below:

mRNA Sequence : 12-04-101-I and 12-G4-111-I (Capped: 12-G4-15;4-C and 12-04-155-C)
Parent Gene ID 91543 and 103294

5' G*GGGAAAUAAG AGAGAAAAGA AGAGUAAGAA GAAAUAUAAG AGCCACCAUG GCCGGUCCCG
CGACCCAAAG CCCCAUGAAA CUUAUGGCCC UGCAGUUGCU GCUUUGGCAC UCGGCCCUCU
GGACAGUCCA AGAAGCGACU CCUCUCGGAC CUGCCUCAUC GUUGCCGCAG UCAUUCCUUU
UGAAGUGUCU GGAGCAGGUG CGAAAGAUUC AGGGCGAUGG AGCCGCACUC CAAGAGAAGC
UCUGCGCGAC AUACAAACUU UGCCAUCCCG AGGAGCUCGU ACUGCUCGGG CACAGCUUGG
GGAUUCCCUG GGCUCCUCUC UCGUCCUGUC CGUCGCAGGC UUUGCAGUUG GCAGGGUGCC
UUUCCCAGCU CCACUCCGGU UUGUUCUUGU AUCAGGGACU GCUGCAAGCC CUUGAGGGAA
UCUCGCCAGA AUUGGGCCCG ACGCUGGACA CGUUGCAGCU CGACGUGGCG GAUUUCGCAA
CAACCAUCUG GCAGCAGAUG GAGGAACUGG GGAUGGCACC CGCGCUGCAG CCCACGCAGG
GGGCAAUGCC GGCCUUUGCG UCCGCGUUUC AGCGCAGGGC GGGUGGAGUC CUCGUAGCGA
GCCACCUUCA AUCAUUUUUG GAAGUCUCGU ACCGGGUGCU GAGACAUCUU GCGCAGCCGU
GAUAAGCUGC CUUCUGCGGG GCUUGCCUUC UGGCCAUGCC CUUCUUCUCU CCCUUGCACC
UGUACCUCUU GGUCUUUGAA UAAAGCCUGA GUAGGAAGGC GGCCGCUCGA GCAUGCAAAA
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
AAAAAAAAAA AAAAAAAUCU AG 3'

G* = cap

Example 7: Oligo dT purification of capped RNA transcript.

[00421] The capped RNA transcript was purified using 2mL of 20mer polythymidine Sepharose packed in a 5mL SPE column on a solid phase extraction vacuum manifold. 0.82 mg was recovered in the elution (84% recovery) The RNA was diafiltered into water and harvested at a concentration of 366 ng/uL using lOOkDa MWCO Amicon spin filters (EMD Millipore). The results are shown in Figure 8.

Example 8: Protein expression using capped RNA transcript produced by the method of the invention.

[00422] Assessment of protein expression was conducted via forward transfection of RNA into HeLa cells. GCSF ELISA was used to quantify expressed protein ELISA data; data are shown in Figure 9.

[00423] 12-04-155-C produced using a linearized Poly A:T tract DNA template and enzymatic capping yielded ~ 1,150 ng/mL GCSF expression, significantly higher than the four lots that utilized PCR product as a DNA template: 12-04-89-C, 12-04-106-C, 12-04-1 14-C, and 12-04-122-C.

Example 9: Interferon-a induction assessment using capped RNA transcript produced by the method of the invention

[00424] RNAs were transfected into Peripheral blood mononuclear cells (PBMC) to assay in vitro cytokine induction. The results are shown in Figure 10. The RNA transcript 12-04-155-C produced using a linearized Poly A:T tract DNA template and enzymatic capping resulted in undetectable levels of interferon-a as determined by IFN-a ELISA. In contrast, two of four samples that utilized PCR product as a DNA template: 12-04-89-C, 12-04-106-C showed some level of interferon-α induction. Poly IC (Polyinosinic:polycytidylic acid) is an immune stimulatory positive control and R-848 is an imidazoquinoline compound that is also used as a positive control.

Example 10: DNA template with Factor IX gene.

[00425] The plasmid pJ204 109475 includes the coding sequence for Factor IX and 141 nucleotide sequence coding for Poly A Tail and an Sapl restriction endonuclease recognition site immediately downstream of the poly A tail sequence tract.

[00426] A plasmid map and DNA sequence are shown in Figures 11A and 1IB.

[00427] While the invention has been particularly shown and described with reference to a preferred embodiment and various alternate embodiments, it will be understood by persons skilled in the relevant art that various changes in form and details can be made therein without departing from the spirit and scope of the invention.

[00428] All references, issued patents and patent applications cited within the body of the instant specification are hereby incorporated by reference in their entirety, for all purposes.

CLAIMS

1. A method for producing a composition comprising an RNA transcript for a gene of interest, the method comprising:

obtaining a sample comprising a linear, non-amplified DNA template, the DNA template comprising an RNA polymerase promoter sequence operatively linked to a target sequence coding for the gene of interest;

contacting the sample with a RNA polymerase and ribonucleotides to form a reaction;

maintaining the reaction under a set of conditions sufficient for vitro transcription to occur,

thereby producing the composition comprising the RNA transcript.

2. The method of claim 1, wherein at least 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, or at least 95% of the RNA transcript in the composition is full-length RNA transcript .

3. The method of claim 2, wherein the percent full length RNA transcript is determined using reverse phase HPLC and measured by peak area of full length relative to total peak area.

4. The method of any of claims 1-3, wherein treatment of the composition with DNase is excluded.

5. The method of any of claims 1-4, wherein the DNA template is separated from the composition via affinity based or ion exchange chromatography.

6. The method of any of claims 1-5, wherein a polymerase chain reaction (PCR) amplification step is excluded.

7. The method of any of claims 1-6, wherein the DNA template comprises 500-25000 or 3000-7000 basepairs.

8. The method of any of claims 1-7, wherein the DNA template further comprises a sequence coding for a poly A tail, an endonuclease recognition site sequence immediately downstream of the poly A tail sequence, and optionally a 3' untranslated region (UTR).

9. The method of claim 8, wherein the poly A tail is 5-300 nucleotides or 60-160 nucleotides in length.

10. The method of claim 8, wherein the endonuclease recognition site sequence is recognized by Xbal or Sapl.

11. The method of claim 8, wherein the DNA template is produced as a circular plasmid and linearized by a method consisting of contacting the DNA template with an endonuclease that recognizes the endonuclease recognition site sequence.

12. The method of claim 11, wherein the DNA template is produced at a microgram scale or a milligram scale or a gram scale.

13. The method of any of claims 1-12, wherein the RNA polymerase is a T7 polymerase.

14. The method of any of claims 1-13, wherein at least one nucleotide is a modified nucleotide.

15. The method of any of claims 1-14, further comprising capping the RNA transcript by contacting the RNA transcript with Vaccinia guanylyltransferase, s-adenosyl-L-methionine (SAM)and guanosine triphosphate (GTP) in a reaction and maintaining the reaction under conditions sufficient to cap the RNA transcript, wherein the RNA transcript is capped with a CapO structure.

16. The method of claim 15, wherein the reaction further comprises 2'-0-methyltransferase wherein the RNA transcript is capped with a Capl structure..

17. The method of any of claims 1-16, further comprising an affinity chromatography step.

18. The method of claim 17, wherein the RNA transcript comprises a poly A tail and the affinity chromatography step is a poly A capture based affinity purification.

19. The method of claim 18, wherein the affinity chromatography step is oligo dT based purification.

20. The method of any of claims 1-19, further comprising an anion exchange chromatography step.

2 1. The method of any of claims 1-20, further comprising a filtration step.

22. The method of claim 21, wherein the filtration step is tangential flow ultrafiltration/ diafiltration.

23. The method of claim 1 comprising:

obtaining a sample comprising a linear, non-amplified DNA template, the DNA template comprising an RNA polymerase promoter sequence operatively linked to a target sequence coding for the gene of interest, a poly A tail sequence of 60-160 nucleotides, an endonuclease recognition site sequence immediately downstream of the poly A tail sequence, optionally a 5' untranslated region (UTR) and/or a 3' UTR, wherein the DNA template is produced at a milligram scale as a circular plasmid and linearized by a method consisting of contacting the DNA template with an endonuclease that recognizes the endonuclease recognition site sequence;

contacting the sample with a T7 RNA polymerase and ribonucleotides to form a first reaction under conditions sufficient to produce the composition comprising the RNA transcript wherein at least 80% of the RNA transcript is full-length RNA transcript

purifying the RNA transcript using oligo dT affinity purification;

capping the RNA transcript by contacting the RNA transcript with guanylyltransferase, s-adenosyl-L-methionine (SAM); guanosine triphosphate (GTP), and optionally 2'-0-methyltransferase in a second reaction and maintaining the second reaction under the following conditions;

purifying the RNA transcript using an anion exchange chromatography step; and

filtering the RNA transcript using an ultrafiltration or tangential flow filtration membrane thereby producing the composition comprising the RNA transcript.

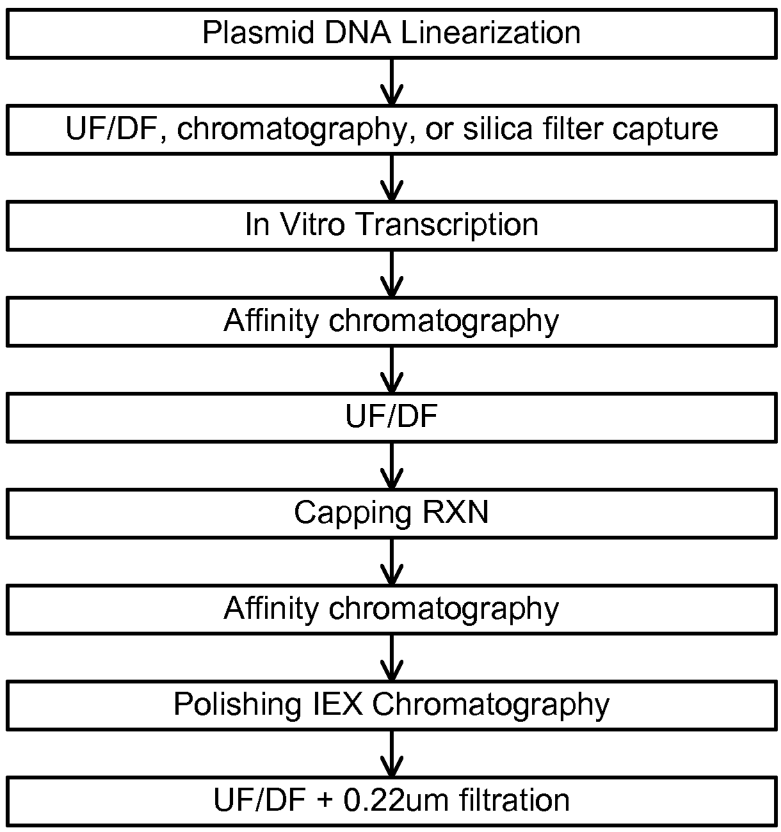

Plasmid DNA Linearization
UF/DF, chromatography, or silica filter capture
In Vitro Transcription
Affinity chromatography
UF/DF
Capping RXN
Affinity chromatography
Polishing IEX Chromatography
UF/DF + 0.22um filtration

FIG. 1


XbaI
5' ...AAAAAAAAAATCTAGA...3'
3' ...TTTTTTTTTTAGATCT...5'
5' ...AAAAAAAAAAT 3'
3' ...TTTTTTTTTTAGATC 5'
5' ...AAAAAAAAAATCTAG 3'
SapI
(class IIs enzyme)
5' ...AAAAAAAAAAGAAGAGC... 3'
3' ...TTTTTTTTTTCTTCTCG... 5'
recognition site
for endonuclease
in plasmid DNA
5' ...AAAAAA-3'
3' ...TTTTTTTTT-5'
linearized
plasmid DNA
5' ...AAAAAAAAA-3'
in vitro
transcribed RNA
Kuhn et. al: RNA Biology 8:1,
35-43; January/February 2011

FIG. 2

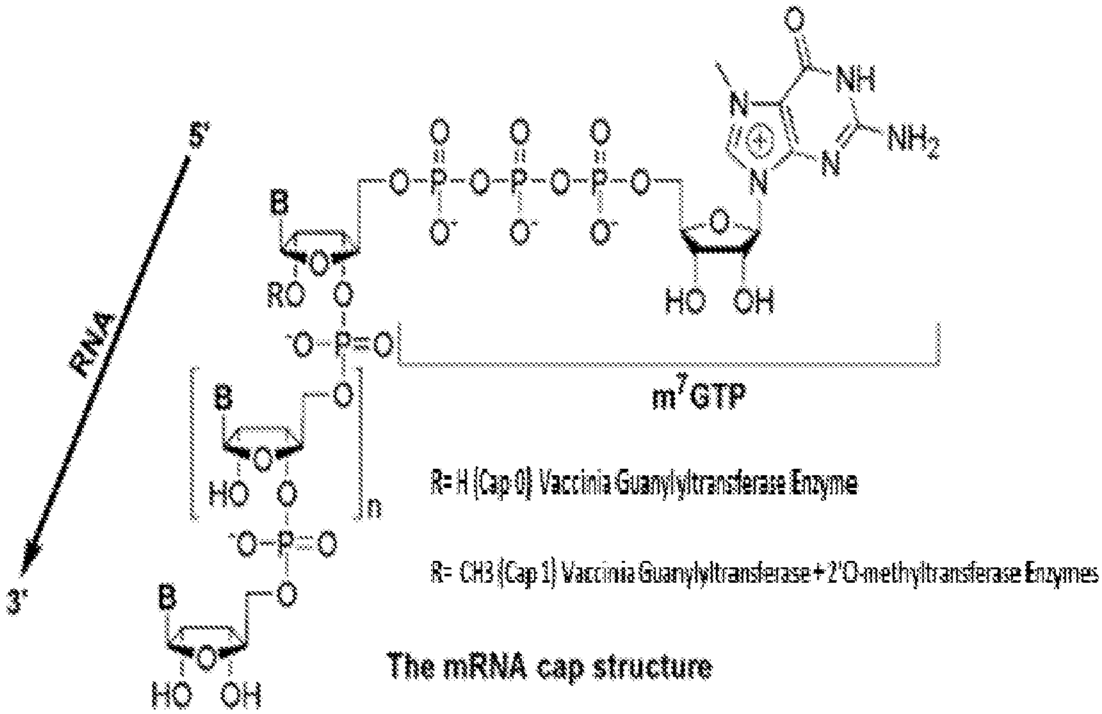

5'
3'
RNA
B
RO
HO
OH
m^7GTP
R= H (Cap 0) Vaccinia Guanylyltransferase Enzyme
R= CH3 (Cap 1) Vaccinia Guanylyltransferase + 2'O-methyltransferase Enzymes
The mRNA cap structure

FIG. 3

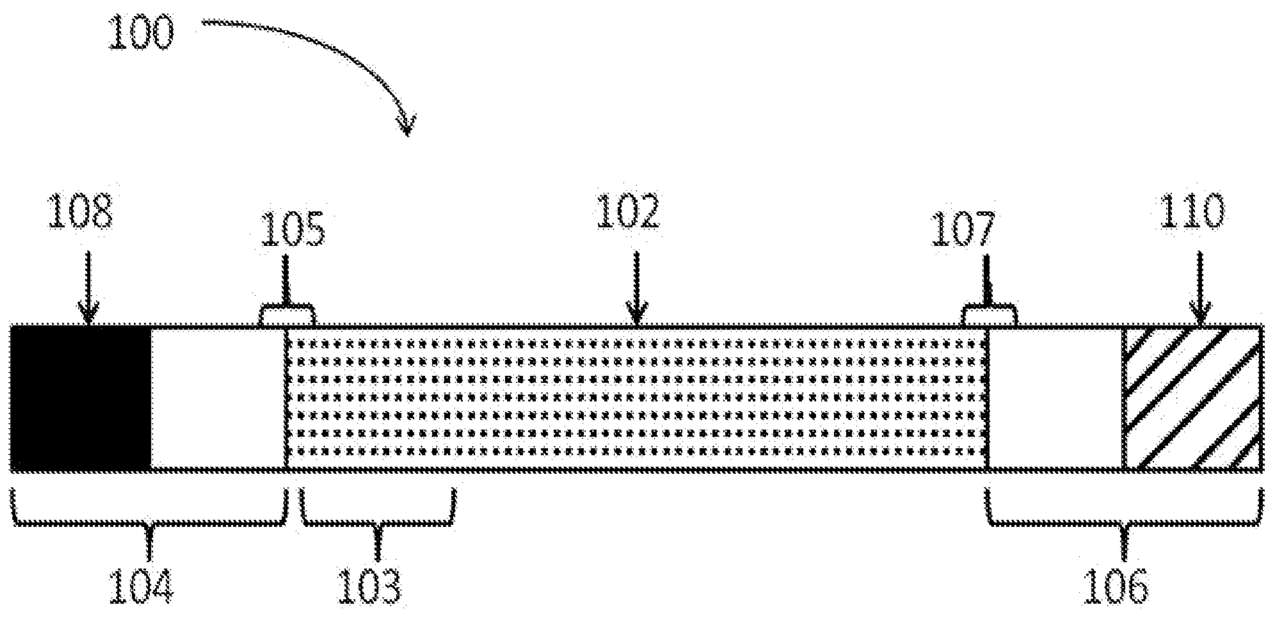

100
108
105
102
107
110
104
103
106

FIG. 4

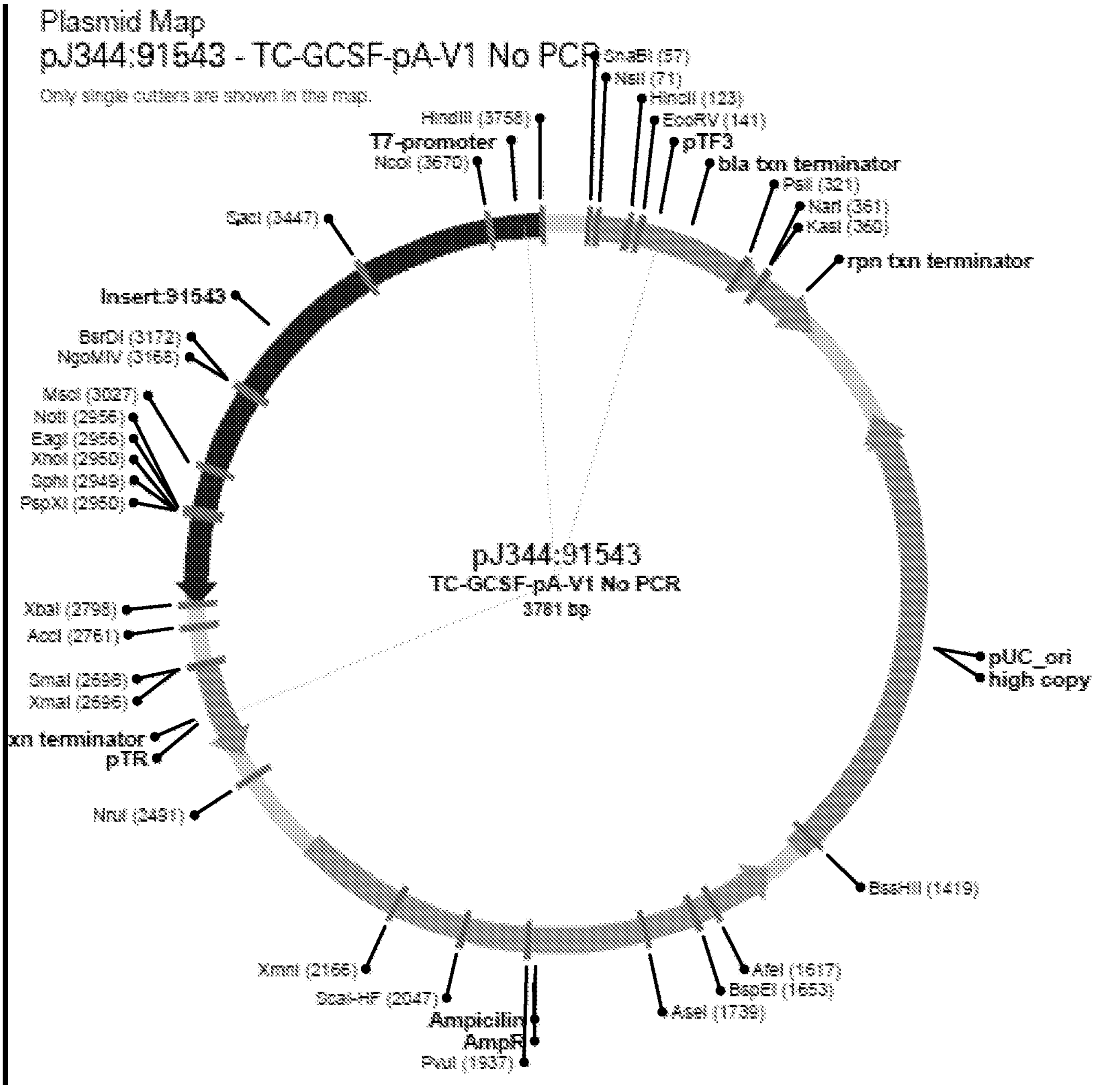

Plasmid Map
pJ344:91543 - TC-GCSF-pA-V1 No PCR
Only single cutters are shown in the map.
SnaBI (57)
NsiI (71)
HincII (123)
EcoRV (141)
pTF3
bla txn terminator
PstI (321)
NarI (361)
KasI (360)
rpn txn terminator
HindIII (3758)
T7-promoter
NcoI (3670)
SacI (3447)
Insert:91543
BsrDI (3172)
NgoMIV (3168)
MscI (3027)
NotI (2956)
EagI (2956)
XhoI (2950)
SphI (2949)
PspXI (2950)
pJ344:91543
TC-GCSF-pA-V1 No PCR
3781 bp
XbaI (2798)
AccI (2761)
SmaI (2698)
XmaI (2696)
txn terminator
pTR
NruI (2491)
pUC_ori
high copy
BssHII (1419)
AfeI (1617)
BspEI (1653)
AseI (1739)
XmnI (2166)
ScaI-HF (2047)
Ampicilin
AmpR
PvuI (1937)

FIG. 5

Sequence

```
   1 GGTCTCCTAT ACCTCTCTTA TACGACATCA CCGATGGGGA ACCCAGACGC TCAGTACGTA TTCTAAATGC
  71 ATAATAAATA CTGATAACAT CTTATAGTTT GTATTATATT TTGTATTATC GTTGACATGT ATAATTTTGA
 141 TATCAAAAAC TGATTTTCCC TTTATTATTT TCGACATTTA TTTTCTTAAT TCTCTTTAAC AAACTAGAAA
 211 TATTGTATAT ACAAAAAATC ATAAATAATA GATGAATAGT TTAATTATAG GTGTTCATCA ATCGAAAAAG
 281 CAACGTATCT TATTTAAAGT GCGTTGCTTT TTTCTCATTT ATAAGGTTAA ATAATTCTCA TATATCAAGC
 351 AAAGTGACAG GCGCCCTTAA ATATTCTGAC AAATGCTCTT TCCCTAAACT CCCCCCATAA AAAAACCCGC
 421 CGAAGCGGGT TTTTACGTTA TTTGCGGATT AACGATTACT CGTTATCAGA ACCGCCCAGG GGGCCCGAGC
 491 TTAACACTGG CCGTCGTTTT ACAACGACGA AAGACTTTGT AGAAACGCAA AAAGGCCATC CGTCAGGGGC
 561 CTTCTGCTTA GTTTGATGCC TGGCAGTTCC CTACTCTCGC CTTCCGCTTC CTCGCTCACT GACTCGCTGC
 631 GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC
 701 AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG
 771 TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA AGTCAGAGGT
 841 GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT
 911 TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC
 981 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG
1051 TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC
1121 GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG
1191 AAGTGGTGGC CTAACTACGG CTACACTAGA AGAACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA
1261 CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT
1331 TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT
1401 GACGCTCAGT GGAACGACGC GCGCGTAACT CACGTTAAGG GATTTTGGTC ATGAGCTTGC GCCGTCCCGT
1471 CAAGTCAGCG TAATGCTCTG CTTTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT
1541 TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG
1611 CCCCAGTGCT GCAATGATAC CGCGAGAACC ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA
1681 GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC
1751 GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT
1821 GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA
1891 TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG
1961 CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT
2031 TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC
2101 CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT
2171 CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
2241 CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
2311 GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATATTCTT CCTTTTTCAA TATTATTGAA
2381 GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG
2451 GGTCAGTGTT ACAACCAATT AACCAATTCT GAACATTATC GCGAGCCCAT TTATACCTGA ATATGGCTCA
2521 TAACACCCCT TGTTTGCCTG GCGGCAGTAG CGCGGTGGTC CCACCTGACC CCATGCCGAA CTCAGAAGTG
2591 AAACGCCGTA GCGCCGATGG TAGTGTGGGG ACTCCCCATG CGAGAGTAGG GAACTGCCAG GCATCAAATA
2661 AAACGAAAGG CTCAGTCGAA AGACTGGGCC TTTCGCCCGG GCTAATTAGG GGGTGTCGAT GACTTCGTGT
2731 TCCCCATGGG TGATGTCGTA TACGAACCAG TATACCAGAC CTATAGGACA CGTATATGGT CTTCTTTTCT
2801 AGATTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
2871 TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
2941 TTTTGCATGC TCGAGCGGCC GCCTTCCTAC TCAGGCTTTA TTCAAAGACC AAGAGGTACA GGTGCAAGGG
3011 AGAGAAGAAG GGCATGGCCA GAAGGCAAGC CCCGCAGAAG GCAGCTTATC ACGGCTGCGC AAGATGTCTC
3081 AGCACCCGGT ACGAGACTTC CAAAAATGAT TGAAGGTGGC TGGCTACGAG GACTCCAGCC GCCCTGCGCT
3151 GAAACCCCGA CGCAAAGGCC GGCATTGCCC CCTGCCTGGG CTGCAGCGCG GGTGCCATCC CCAGTTCCTC
3221 CATCTGCTGC CAGATGGTTG TTGCGAAATC CGCCACCGTG AGCTGCAACG TGTCCAGGGT CGGGGCCAAT
3291 TCTGGCGAGA TTCCCTCAAG GGCTTGCAGC AGTGGCTGAT ACAAGAACAA ACCGGAGTGG AGCTGGGAAA
3361 GGCACCGTGC CAACTGCAAA GGCTGCGACG GACAGGACCA GAGAGGAGCC CACGGAATGC CCAAGCTGTG
3431 CCCGAGCAGT ACGAGCTCCT CGGGATGGCA AAGTTTGTAT GTGTGGGAGA GCTTCTCTTG GAGTGCGGCT
3501 CCATCGGCCT GAATCTTTGG CACCTGGTCC AGACACTTCA AAAGGAATGA CTGCGGGAAC GATGAGGCAG
3571 GTCCCAGAGG AGTCGCTTCT TGCACTGTCC AGAGGGGCGA GTGCCAAAGC AGCAACTGCA GGGGCATAAG
3641 TTTCATGGGG CTTTGGTTCG CGGGACCGGC CATGGTGGCT CTTATATTTC TTCTTACTCT TCTTTTCTCT
3711 CTTATTTCCC TATAGTGAGT CGTATTAGCT TCTGTACGAC GGTCCAAAAG CTTGACCCCG AAGACACTAT
3781 A
```

FIG. 6

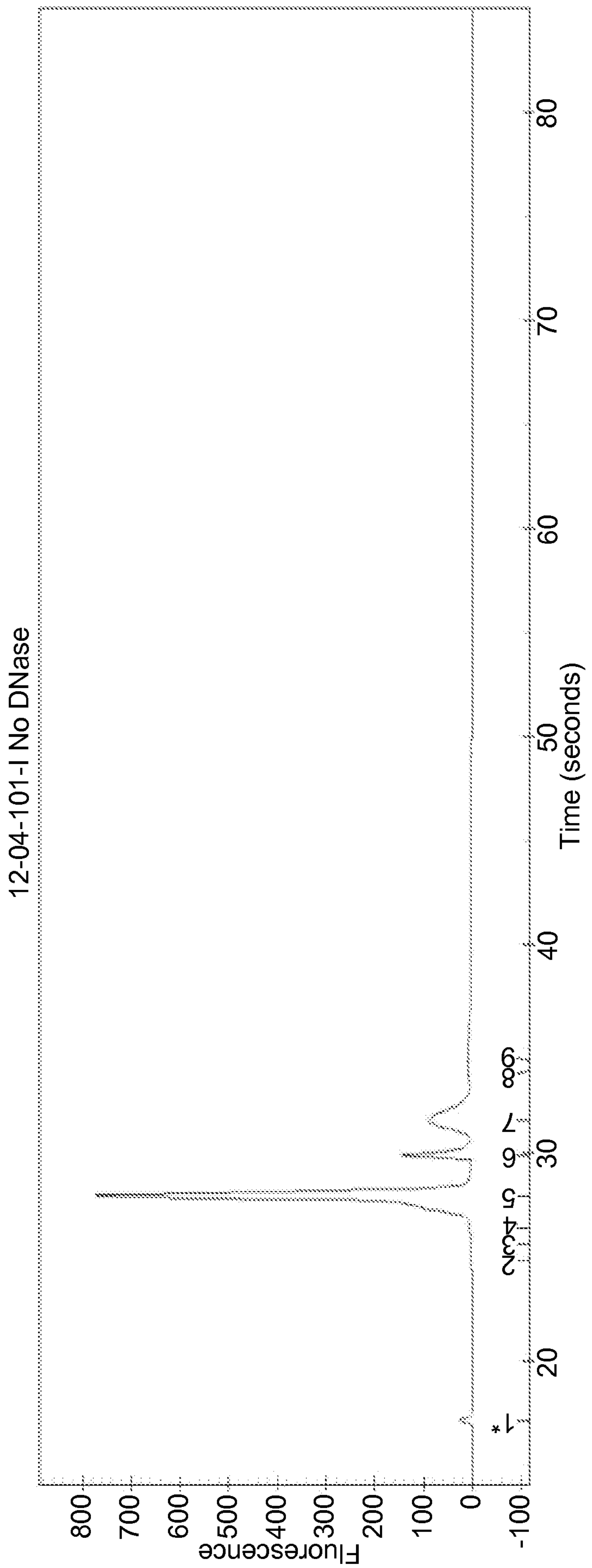

12-04-101-I No DNase
Fluorescence
800
700
600
500
400
300
200
100
0
-100
Time (seconds)
20
30
40
50
60
70
80
1*
2
3
4
5
6
7
8
9

FIG. 7A

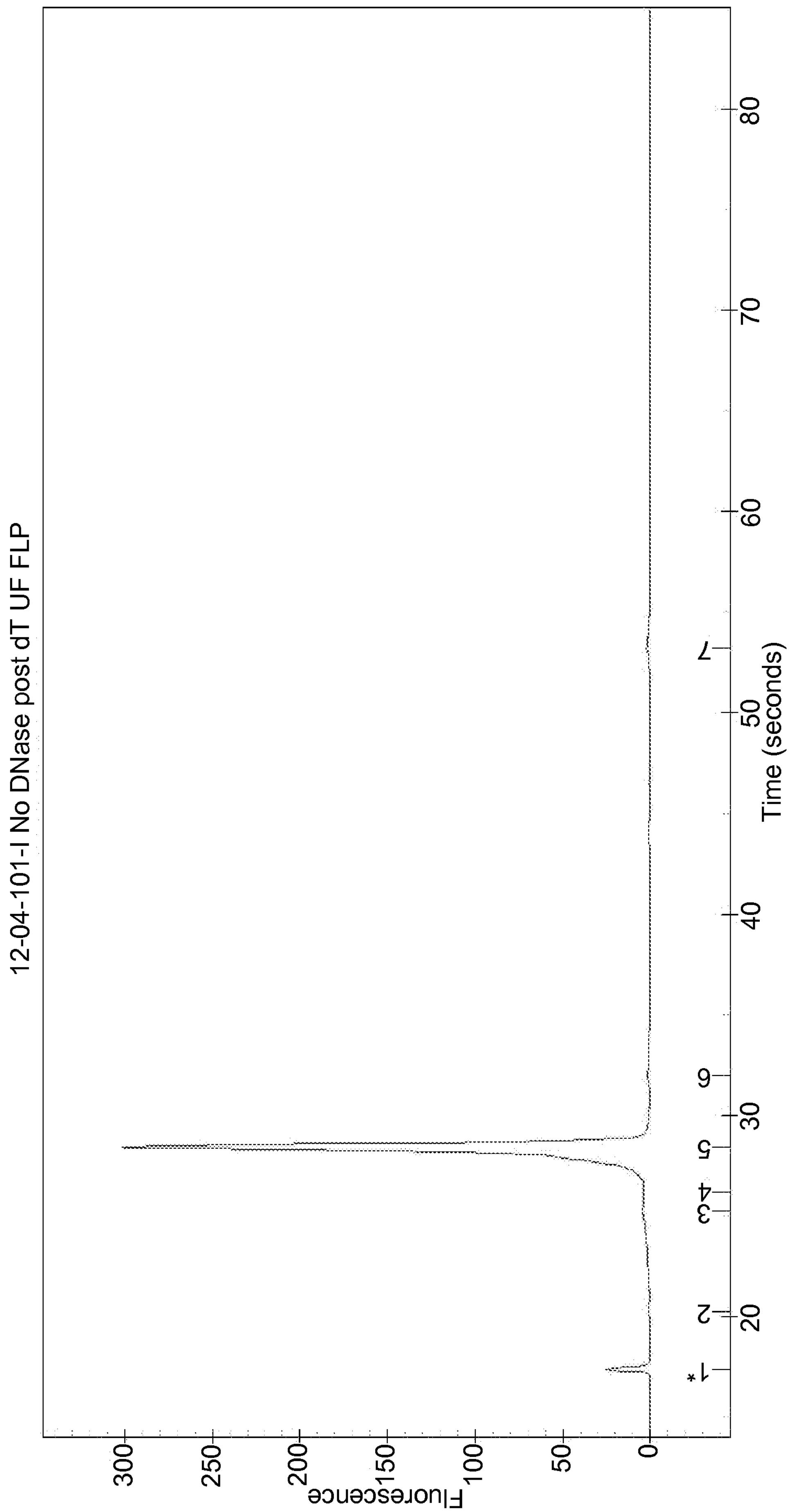

12-04-101-I No DNase post dT UF FLP
Fluorescence
300
250
200
150
100
50
0
20
30
40
50
60
70
80
Time (seconds)
1*
2
3
4
5
6
7

FIG. 7B

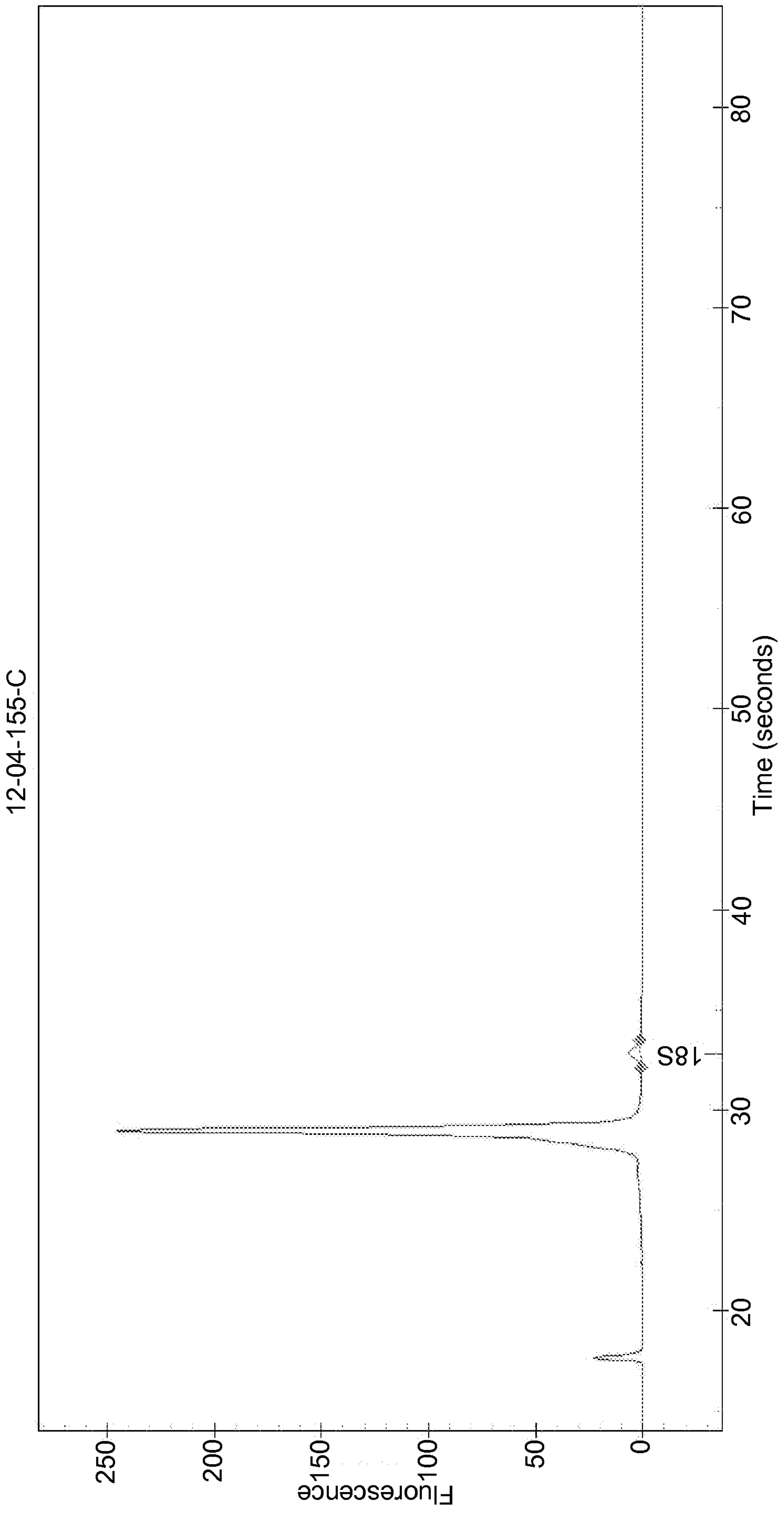

12-04-155-C
Fluorescence
250
200
150
100
50
0
18S
20
30
40
50
60
70
80
Time (seconds)

FIG. 8

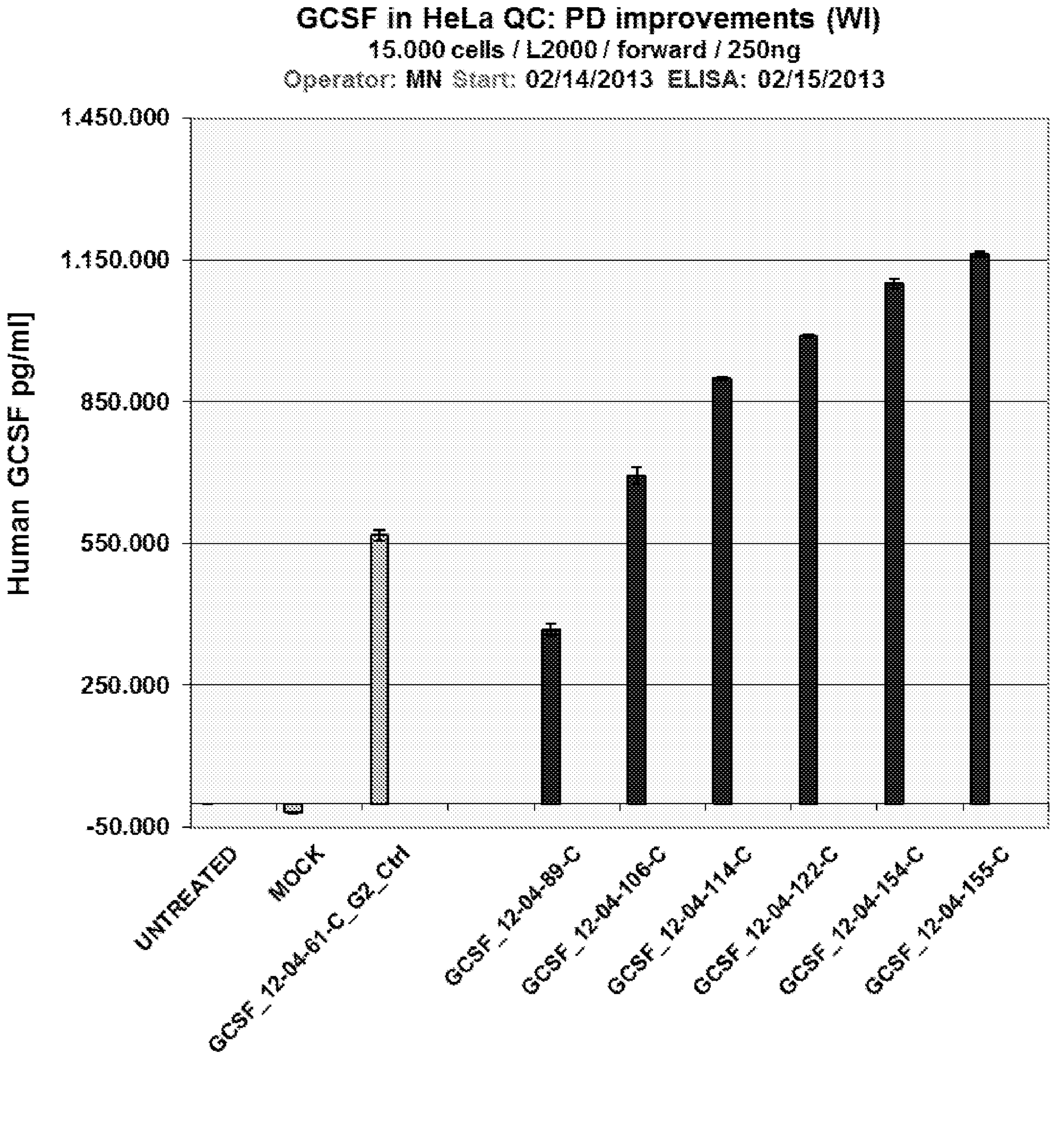

GCSF in HeLa QC: PD improvements (WI)
15.000 cells / L2000 / forward / 250ng
Operator: MN Start: 02/14/2013 ELISA: 02/15/2013
Human GCSF pg/ml]
1.450.000
1.150.000
850.000
550.000
250.000
-50.000
UNTREATED
MOCK
GCSF_12-04-61-C_G2_Ctrl
GCSF_12-04-89-C
GCSF_12-04-106-C
GCSF_12-04-114-C
GCSF_12-04-122-C
GCSF_12-04-154-C
GCSF_12-04-155-C

FIG. 9

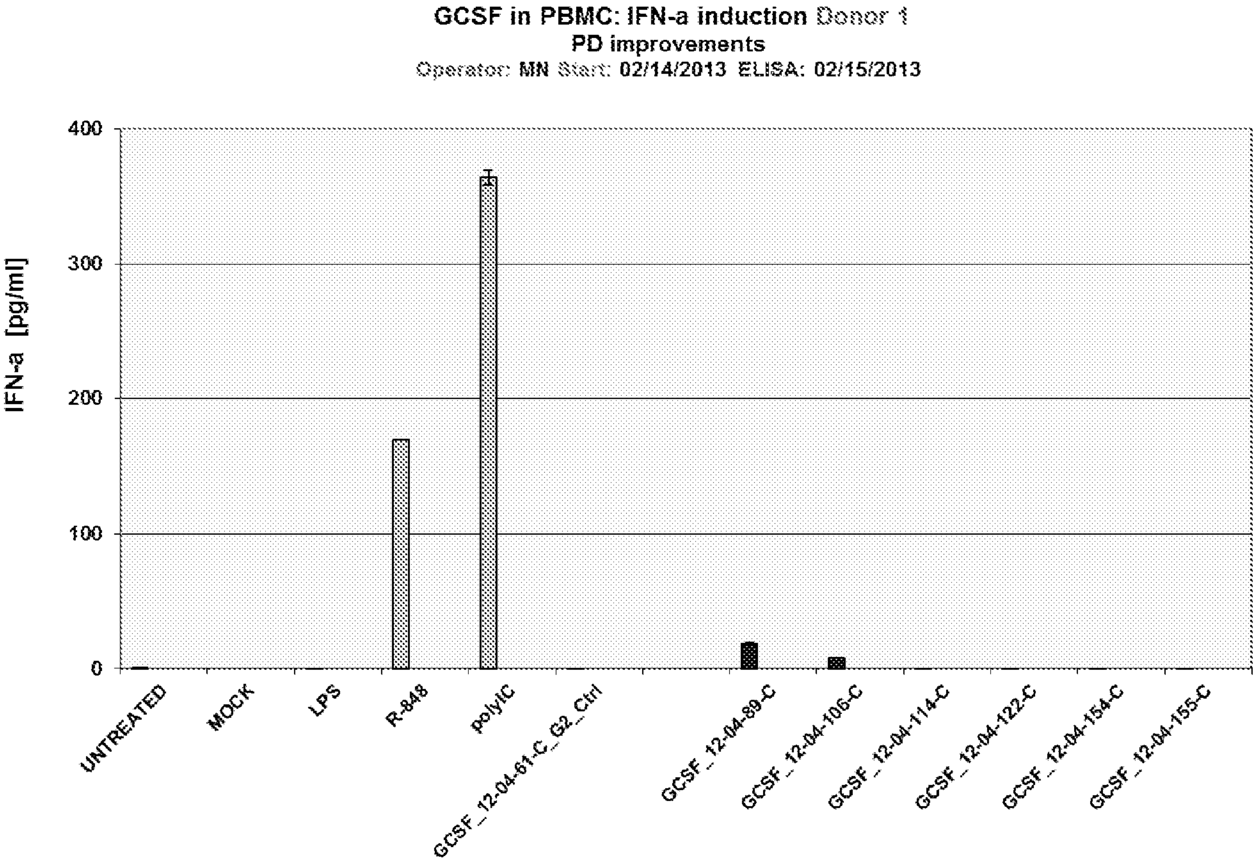

GCSF in PBMC: IFN-a induction Donor 1
PD improvements
Operator: MN Start: 02/14/2013 ELISA: 02/15/2013
IFN-a [pg/ml]
400
300
200
100
0
UNTREATED
MOCK
LPS
R-848
poly IC
GCSF_12-04-61-C_G2_Ctrl
GCSF_12-04-89-C
GCSF_12-04-106-C
GCSF_12-04-114-C
GCSF_12-04-122-C
GCSF_12-04-154-C
GCSF_12-04-155-C

FIG. 10


pJ204:109475
TC_FIX_Hs_3_pA140_Sap
bla txn terminator
pTF3
T7-promoter
Insert:109475
pTR
rrnB1 B2 T1 txn terminator

FIG.11A

Sequence

1 AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA
61 CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT
121 TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC
181 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA
241 TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA
301 GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC
361 CCAGCTTGGA GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA
421 GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA
481 CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG
541 GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC
601 TATGGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT GGCCTTTTGC TGGCCTTTTG
661 CTCACATGTT CTTTCCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG
721 AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG
781 AAGCGGAAGG CGAGAGTAGG GAACTGCCAG GCATCAAACT AAGCAGAAGG CCCCTGACGG
841 ATGGCCTTTT TGCGTTTCTA CAAACTCTTT CTGTGTTGTA AAACGACGGC CAGTCTTAAG
901 CTCGGGCCCC CTGGGCGGTT CTGATAACGA GTAATCGTTA ATCCGCAAAT AACGTAAAAA
961 CCCGCTTCGG CGGGTTTTTT TATGGGGGGA GTTTAGGGAA AGAGCATTTG TCAGAATATT
1021 TAAGGGCGCC TGTCACTTTG CTTGATATAT GAGAATTATT TAACCTTATA AATGAGAAAA
1081 AAGCAACGCA CTTTAAATAA GATACGTTGC TTTTTCGATT GATGAACACC TATAATTAAA
1141 CTATTCATCT ATTATTTATG ATTTTTTGTA TATACAATAT TTCTAGTTTG TTAAAGAGAA
1201 TTAAGAAAAT AAATCTCGAA AATAATAAAG GGAAAATCAG TTTTTGATAT CAAAATTATA
1261 CATGTCAACG ATAATACAAA ATATAATACA AACTATAAGA TGTTATCAGT ATTTATTATG
1321 CATTTAGAAT AAATTTTGTG TCGCCCTTCG CTGAATCAAG CTTTTGGACC CTCGTACAGA
1381 AGCTAATACG ACTCACTATA GGGAAATAAG AGAGAAAAGA AGAGTAAGAA GAAATATAAG
1441 AGCCACCATG CAGCGCGTCA ACATGATTAT GGCCGAATCG CCGGGACTCA TCACAATCTG
1501 CCTCTTGGGT TATCTCTTGT CGGCAGAATG TACCGTGTTC TTGGATCACG AAAACGCGAA
1561 CAAAATTCTT AATCGCCCGA AGCGGTATAA CTCCGGGAAA CTTGAGGAGT TTGTGCAGGG
1621 CAATCTTGAA CGAGAGTGCA TGGAGGAGAA ATGCTCCTTT GAGGAGGCGA GGGAAGTGTT
1681 TGAAAACACA GAGCGAACAA CGGAGTTTTG GAAGCAATAC GTAGATGGGG ACCAGTGTGA
1741 GTCGAATCCG TGCCTCAATG GGGGATCATG TAAAGATGAC ATCAATAGCT ATGAATGCTG
1801 GTGCCCGTTT GGGTTTGAAG GGAAGAACTG TGAGCTGGAT GTGACGTGCA ACATCAAAAA
1861 CGGACGCTGT GAGCAGTTTT GTAAGAACTC GGCTGACAAT AAGGTAGTAT GCTCGTGCAC
1921 AGAGGGATAC CGGCTGGCGG AGAACCAAAA ATCGTGCGAG CCCGCAGTCC CGTTCCCTTG
1981 TGGGAGGGTG AGCGTGTCAC AGACTAGCAA GTTGACGAGA GCGGAGACTG TATTCCCCGA
2041 CGTGGACTAC GTCAACAGCA CCGAAGCCGA AACAATCCTC GATAACATCA CGCAGAGCAC
2101 TCAGTCCTTC AATGACTTTA CGAGGGTCGT AGGTGGTGAG GACGCGAAAC CCGGTCAGTT
2161 CCCCTGGCAG GTGGTATTGA ACGGAAAAGT CGATGCCTTT TGTGGAGGTT CCATTGTCAA
2221 CGAGAAGTGG ATTGTCACAG CGGCACACTG CGTAGAAACA GGAGTGAAAA TCACGGTAGT
2281 GGCGGGAGAG CATAACATTG AAGAGACAGA GCACACGGAA CAAAAGCGAA ATGTCATCAG
2341 AATCATTCCA CACCATAACT ATAACGCGGC AATCAATAAG TACAATCACG ACATCGCACT
2401 TTTGGAGCTT GACGAACCTT TGGTGCTTAA TTCGTACGTC ACCCCTATTT GTATTGCCGA
2461 CAAAGAGTAT ACAAACATCT TCTTGAAATT CGGCTCCGGG TACGTATCGG GCTGGGGCAG
2521 AGTGTTCCAT AAGGGTAGAT CCGCACTGGT GTTGCAATAC CTCAGGGTGC CCCTCGTGGA
2581 TCGAGCCACT TGTCTGCGGT CCACCAAATT CACAATCTAC AACAATATGT TCTGTGCGGG
2641 ATTCCATGAA GGTGGGAGAG ATAGCTGCCA GGGAGACTCA GGGGGTCCCC ACGTGACGGA
2701 AGTCGAGGGG ACGTCATTTC TGACGGGAAT TATCTCATGG GGAGAGGAAT GTGCGATGAA
2761 GGGGAAATAT GGCATCTACA CTAAAGTGTC ACGGTATGTC AATTGGATCA AGGAAAAGAC
2821 GAAACTCACG TGATAATAGG CTGGAGCCTC GGTGGCCATG CTTCTTGCCC CTTGGGCCTC
2881 CCCCCAGCCC CTCCTCCCCT TCCTGCACCC GTACCCCCGT GGTCTTTGAA TAAAGTCTGA
2941 GTGGGCGGCA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3001 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3061 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAG AAGAGCCGTC AATCGAGTTC GTACCTAAGG
3121 GCGACACCCC ATAATTAGCC CGGGCGAAAG GCCCAGTCTT TCGACTGAGC CTTTCGTTTT
3181 ATTTGATGCC TGGCAGTTCC CTACTCTCGC ATGGGGAGTC CCCACACTAC CATCGGCGCT
3241 ACGGCGTTTC ACTTCTGAGT TCGGCATGGG GTCAGGTGGG ACCACCGCGC TACTGCCGCC
3301 AGGCAAACAA GGGGTGTTAT GAGCCATATT CAGGTATAAA TGGGCTCGCG ATAATGTTCA
3361 GAATTGGTTA ATTGGTTGTA ACACTGACCC CTATTTGTTT ATTTTTCTAA ATACATTCAA
3421 ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT TCAATAATAT TGAAAAAGGA
3481 AGAATATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC CTTTTTTGCG GCATTTTGCC
3541 TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG
3601 GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC
3661 GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT
3721 TATCCCGTAT TGACGCCGGG CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG
3781 ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG
3841 AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC GGCCAACTTA CTTCTGACAA
3901 CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA CATGGGGGAT CATGTAACTC
3961 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA
4021 CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC
4081 TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC
4141 TGCGCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG
4201 GGTCTCGCGG TATCATTGCA GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA
4261 TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG
4321 GTGCCTCACT GATTAAGCAT TGGTAAAAGC AGAGCATTAC GCTGACTTGA CGGGACGGCG
4381 CAAGCTCATG ACCAAAATCC CTTAACGTGA GTTACGCGCG CGTCGTTCCA CTGAGCGTCA
4441 GACCCCGTAG

FIG. 11B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US14/26835

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 5-22
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 14/26835

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12P 19/34, 19/30, 19/26 (2014.01)
USPC - 435/91.1, 91.21, 183
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8): C12P 19/34, 19/30, 19/26, 19/00 (2014.01)
USPC: 435/91.1, 91.21, 89, 85, 84, 72, 41, 183

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
MicroPatent (US-G, US-A, EP-A, EP-B, WO, JP-bib, DE-C.B, DE-A, DE-T, DE-U, GB-A, FR-A); ProQuest Dialog; PubMed/PubMed Central; Google; Google Scholar: 'target sequence,' 'target gene,' 'gene of interest,' 'sequence of interest,' 'RNA polymerase,' 'RNA,' 'mRNA,' transcript, promoter, 'in vitro transcription,' 'Poly A tail,' polyadenyl, 'oligo dT affinity purification,' 'transcript cap'

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2004/0259097 A1 (DEBACKER, MD et al.) December 23, 2004; paragraphs [0011]-[0013], [0026], [0035]	1, 4/1 ----- 2, 3, 4/2, 4/2, 23
Y	WO 1991/05058 A1 (KAWASAKI, G) April 18, 1991; page 31, lines 3-5	2, 3, 4/2, 4/3, 23
Y	US 6511832 B1 (GUARINO, LA et al.) January 28, 2003; column 3, lines 13-67; column 4, lines 30-35; column 44, lines 38-45; column 47, lines 30-35; column 53, lines 25-30; column 56, lines 35-40; column 63, lines 1-5	23
A	US 8093367 B2 (KORE, A et al.) January 10, 2012; entire document	1-3, 4/1-4/3, 23
A	WO 02/44399 A2 (ZIMAN, M et al.) June 06, 2002; entire document	1-3, 4/1-3, 23
P, Y	WO 2014/028429 A2 (WANG, Y et al.) February 20, 2014; entire document	1-3, 4/1-4/3, 23

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
08 August 2014 (08.08.2014)	28 AUG 2014

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-3201

Authorized officer:
Shane Thomas
PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774