

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

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
30 March 2017 (30.03.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/053297 A1

- (51) International Patent Classification:
A23L 33/13 (2016.01) A61K 31/712 (2006.01)
A61K 31/7115 (2006.01)
- (21) International Application Number:
PCT/US2016/052670
- (22) International Filing Date:
20 September 2016 (20.09.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/221,248 21 September 2015 (21.09.2015) US
- (71) Applicant: **TRILINK BIOTECHNOLOGIES, INC.**
[US/US]; 9955 Mesa Rim Road, San Diego, California
92109 (US).
- (72) Inventors: **HOGREFE, Richard I.**; 5235 Castle Hills
Drive, San Diego, California 92109 (US). **LEBEDEV, Al-**
exandre; 9620 Towne Center Drive, Unit 24, San Diego,
California 92121 (US). **MCCAFFREY, Anton P.**; 4944
Cass Street, Apt. 706, San Diego, California 92109 (US).
SHIN, Dongwong; 10956 Caminito Arcada, San Diego,
California 92131 (US).
- (74) Agent: **REITER, Stephen E.**; Kilpatrick Townsend &
Stockton LLP, 1100 Peachtree Street, Suite 2800, Atlanta,
Georgia 30309 (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,
KW, KZ, LA, LC, LK, LR, LS, 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, ST, 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).

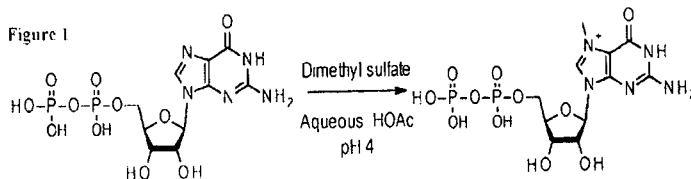
Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

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))

(54) Title: COMPOSITIONS AND METHODS FOR SYNTHESIZING 5'-CAPPED RNAS



(57) Abstract: Provided herein are methods and compositions for synthesizing 5'-capped RNAs wherein the initiating capped oligo-nucleotide primers have the general form ${}^{m7}\text{Gppp}[N]_n$ wherein ${}^{m7}\text{G}$ is N7-methylated guanosine or any guanosine analog, N is any natural, modified or unnatural nucleoside, "n" can be any integer from 0 to 4 and "m" can be an integer from 1 to 9.

COMPOSITIONS AND METHODS FOR SYNTHESIZING 5'-CAPPED RNAs

CROSS-REFERENCE TO RELATED APPLICATION

[0001] The present application claims priority from US Provisional Application No. 62/221,248, filed 21 September 2015, the contents of which are hereby incorporated by reference herein in their entirety.

SEQUENCE LISTING

[0001.1] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on November 16, 2016, is named 095109-000500PC-1022543_SL.txt and is 785 bytes in size.

FIELD OF THE INVENTION

[0002] The present invention relates to methods and compositions for the synthesis of 5'-Capped RNAs. In particular aspects, the present invention relates to novel Cap containing initiating oligonucleotide primers having natural or modified 5'-Cap 0, Cap 1, Cap 2 or trimethylguanosine-Cap (TMG-Cap) structures. In additional aspects, the present invention relates to methods for efficiently generating and using the same for preparing 5'-Capped RNAs are provided.

BACKGROUND OF THE INVENTION

[0003] The following description is provided to assist the understanding of the reader. None of the information provided or references cited is admitted to be prior art in the present invention.

[0004] Messenger RNA (mRNA), encoding physiologically important proteins for therapeutic applications, has shown significant advantages over DNA-based plasmid and viral vectors for delivering genetic material. The most important of these advantages are:

- (i) high level of safety (reduces potential genome damage from viral or plasmid integration),
- (ii) mRNA delivery results in immediate protein expression (unlike delayed responses

that generally occur with plasmids),

(iii) mRNA allows for robust dose-dependent control over expression of proteins and

(iv) the simplicity of large scale synthesis of mRNAs compared to manufacturing of plasmid and viral vectors.

[0005] Messenger RNAs can be encoded for virtually any known protein and can be delivered to specific tissues and organs by a variety of methods known to those skilled in the art. Once delivered, these mRNAs direct ribosomal protein expression within targeted tissues resulting in the production of many hundreds of proteins per mRNA molecule.

[0006] Several structural elements, present in each active mRNA molecule, are utilized to translate the encoded proteins efficiently. One of these elements is a Cap structure on the 5'-end of mRNAs, which is present in all eukaryotic organisms (and some viruses). Naturally occurring Cap structures comprise a ribo-guanosine residue that is methylated at position N7 of the guanine base. This 7-methylguanosine (⁷mG) is linked via a 5'- to 5'-triphosphate chain at the 5'-end of the mRNA molecule. Throughout this application, 7m and m7 are used interchangeably with equivalent meaning. The presence of the ⁷mGppp fragment on the 5'-end is essential for mRNA maturation, it:

protects the mRNAs from degradation by exonucleases,

facilitates transport of mRNAs from the nucleus to the cytoplasm and

plays a key role in assembly of the translation initiation complex (*Cell* 9:645-653, (1976); *Nature* 266:235, (1977); *Federation of Experimental Biologists Society Letter* 96:1-11, (1978); *Cell* 40:223-24, (1985); *Prog. Nuc. Acid Res.* 35:173-207, (1988); *Ann. Rev. Biochem.* 68:913-963, (1999); *J. Biol. Chem.* 274:30337-3040, (1999)).

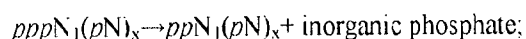
[0007] Only those mRNAs that carry the Cap structure are active in Cap dependent translation; "decapitation" of mRNA results in an almost complete loss of their template activity for protein synthesis (*Nature*, 255:33-37, (1975); *J. Biol. Chem.*, vol. 253:5228-5231, (1978); and *Proc. Natl. Acad. Sci. USA*, 72:1189-1193, (1975)).

[0008] Another element of eukaryotic mRNA is the presence of 2'-O-methyl nucleoside residues at transcript position 1 (Cap 1), and in some cases, at transcript positions 1 and 2 (Cap 2). The 2'-O-methylation of mRNA is required for higher efficacy of mRNA translation *in vivo* (*Proc. Natl. Acad. Sci. USA*, 77:3952-3956 (1980)) and further improves nuclease

stability of the 5'-capped mRNA. The mRNA with Cap 1 (and Cap 2) is a distinctive mark that allows cells to recognize the *bona fide* mRNA 5' end, and in some instances, to discriminate against transcripts emanating from infectious genetic elements (*Nucleic Acid Research* 43: 482-492 (2015)).

[0009] A primary mRNA transcript carries a 5'-triphosphate group (5'-pppmRNA) resulting from initiation of RNA synthesis starting with NTP (typically GTP) *in vivo*. Conversion of 5'-triphosphorylated end of mRNA transcript into a Cap structure (Cap 0) occurs via several enzymatic steps (*J. Biol. Chem.* 250:9322, (1975); *J. Biol. Chem.* 271:11936, (1996); *J. Biol. Chem.* 267:16430, (1992)). These enzymatic reactions steps include:

Step 1: RNA triphosphatase converts 5'-triphosphate of mRNA to a 5'-diphosphate,



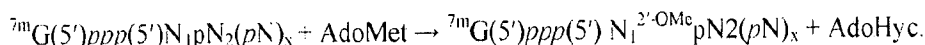
Step 2: RNA guanylyltransferase uses GTP to transfer a GMP residue to the 5'-diphosphate of the mRNA, $ppN_1(pN)_x + GTP \rightarrow G(5')ppp(5')N_1(pN)_x + \text{inorganic pyrophosphate};$ and

Step 3: guanine-7-methyltransferase, uses S-adenosyl-methionine (AdoMet) as a cofactor and transfers the methyl group from AdoMet to the 7-nitrogen of the guanine base, $G(5')ppp(5')N_1(pN)_x + \text{AdoMet} \rightarrow {}^mG(5')ppp(5')N_1(pN)_x + \text{AdoHyc}.$

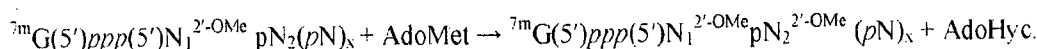
The RNA that results from these enzymatic activities is referred to as "5' capped RNA" or "capped RNA", and the combination of enzymes involved in this process that results in formation of "capped RNA" are referred to as "capping enzymes". Capping enzymes, including cloned forms of such enzymes, have been identified and purified from many sources and are well known in the art (*Prog. Nucleic Acid Res. Mol. Biol.* 66:1-40, (2001); *Prog. Nucleic Acid Res. Mol. Biol.* 50:101-129, (1995); and *Microbiol. Rev.* 44:175, (1980)). The capped RNA that results from the addition of the cap nucleotide to the 5'-end of primary RNA by capping enzymes has been referred to as capped RNA having a "Cap 0 structure" (*J. Biol. Chem.* 269:14974-14981, (1994); *J. Biol. Chem.* 271:11936-11944, (1996)). Capping enzymes have been used to synthesize capped RNA having a Cap 0 structure *in vitro* (*J. Biol. Chem.* 255:11588, (1980); *Proc. Natl. Acad. Sci. USA* 94:9573, (1997); *J. Biol. Chem.* 267:16430, (1992); *J. Biol. Chem.* 269:14974, (1994); and *J. Biol. Chem.* 271:11936, (1996)).

[0010] Capped RNA having a 5'-Cap 0 structure can be further transformed *in vivo* to a "Cap 1" structure by the action of (nucleoside-2'-O-) methyltransferase (*J. Biol. Chem.* 269:14974-

14981, (1994); *J. Biol. Chem.* 271:11936-11944, (1996); and *EMBO* 21:2757-2768, (2002)). For example, vaccinia mRNA (nucleoside-2'-O) methyltransferase can catalyze methylation of the 2'-hydroxyl group of the 5'-penultimate nucleotide of 5'-capped RNA having a Cap 0 structure by the following reaction:



[0011] Dimethylated capped RNAs having a Cap 1 structure have been reported to be translated more efficiently than capped RNAs having a Cap 0 structure (*Nucleic Acids Res.* 26:3208, (1998)). Eukaryotic cells utilize another (nucleoside-2'-O) methyltransferase (for example hMTR2 in human cells (*Nucleic Acids Res.* 39:4756 (2011))) to catalyze methylation of the 2'-hydroxyl group of the second transcribed nucleotide of 5'-capped RNA to convert the Cap 1 structure to a Cap 2 structure by the following reaction:



Approximately 50% of eukaryotic mRNAs have a Cap 2 structure.

[0012] In order to produce long functional RNAs for various biological studies, a method of *in vitro* enzymatic synthesis of primary RNA was developed in the mid-1980s (*Methods Enzymol.* 180:51-62, (1989); *Nucl. Acids Res.*, 10:6353-6362, (1982); *Meth. Enzymol.*, 180:42-50 (1989); *Nucl. Acids Res.*, 12:7035-7056, (1984) and *Nucleic Acid Research* 15: 8783-8798, (1987)).

[0013] After *in vitro* transcription, the primary mRNA transcript carrying 5'-triphosphate group can be further capped with capping enzymes. However, *in vitro* enzymatic 5'-capping is expensive, laborious, inefficient and difficult to control.

[0014] In view of these disadvantages, another method was developed for the *in vitro* synthesis of capped mRNA where a chemically synthesized dinucleotide $^7\text{mG}(5')\text{ppp}(5')\text{G}$ (also referred to as mCAP) is used to initiate transcription (*RNA* 1: 957-967, (1995)). The mCAP dinucleotide contains the 5'-Cap 0 structure of mature mRNA but does not have 2'-O-methyl nucleosides characteristic for Cap 1 and Cap 2 structures.

[0015] However, there are two main disadvantages attributed to initiation of *in vitro* transcription using synthetic mCAP dinucleotide. The first is a strong competition of mCAP and pppG for initiation of mRNA synthesis. When mRNA is initiated with pppG the resultant ppp-mRNA is inactive in translation and immunogenic due to the presence of the 5'-

triphosphate. Correspondingly, when mRNA is initiated with mCAP the resultant 5'-capped-mRNA is active in translation and is not as immunogenic.

[0016] In order to improve the ratio of 5'-capped to 5'-uncapped (or 5'-triphosphorylated; pppmRNA) mRNAs, an excess of $^7\text{mGpppG}$ over pppG (from 4:1 to 10:1) must be used to favor the production of the 5'-Cap structure mRNA transcripts (up to 80-90%). The negative side of this approach is a significant reduction of overall yield of mRNA due to a fast depletion of GTP supply during transcription and a requirement for large quantities of a synthetic mCAP dimer which can be expensive. After transcription an additional treatment of crude mixture, containing both 5'-capped mRNA and 5'-pppmRNA, with alkaline phosphatase is necessary to remove uncapped 5'-triphosphate groups from pppmRNA in order to reduce immunogenicity of synthesized mRNA. The uncapped 5'-OH form of mRNA obtained after phosphatase treatment is inactive and does not participate in translation process.

[0017] Another disadvantage, a bi-directional initiation, can arise when using a non-symmetrical mCAP dinucleotide. There is a tendency of the 3'-hydroxyl group of either the G or ^7mG moiety of $^7\text{mGpppG}$ to serve as initiation point for transcriptional elongation with a nearly equal probability. It typically leads to a synthesis of two isomeric RNAs of the form $^7\text{mG}(5')\text{pppG}(\text{pN})_n$ and $\text{G}(5')\text{ppp}^7\text{mG}(\text{pN})_n$, in approximately equal proportions, depending on conditions of the transcription reaction (*RNA* 1: 957-967, (1995)).

[0018] To eliminate bi-directional initiation of mRNA synthesis with mCAP dinucleotide a novel modified mCAP analog in which the 3'-OH group of ^7mG residue is replaced with OCH_3 ("OMe"): $^7\text{mG}(3'\text{-O-Me})\text{pppG}$ (also known as Anti-Reverse Cap Analog (ARCA)) was developed. ARCA initiates mRNA synthesis only in the correct forward orientation (*RNA* 7:1486-1495 (2001)). Several types of ARCA analogs are known in the art (see, for example, U.S. Patent No. 7,074,596). However, a large molar excess of ARCA over pppG is still required to ensure that most mRNA transcript molecules have the 5'-Cap structure. A further disadvantage is that an mRNA with a Cap 1 structure cannot be synthesized using a $^7\text{mGpppG}^{2'\text{-OMe}}$ Cap dimer (*RNA* 1: 957, (1995)) or its ARCA analog.

[0019] Presently, the known routes to production of active long mRNAs containing a Cap 1 structure consist of enzymatic capping and enzymatic 2' O-methylation of the 5'-triphosphorylated mRNA transcript or enzymatic 2'-O-methylation of mCAP-capped or ARCA-capped mRNA precursor (*Nucleosides, Nucleotides, and Nucleic Acids*, 25:337-340,

(2006) and *Nucleosides, Nucleotides, and Nucleic Acids* 25(3):307-14, (2006)). Both approaches are quite laborious, difficult to control and, even with a substantial optimization, neither approach can guarantee a high yield of a capped and methylated mRNA precursor (*J. Gen. Virol.*, 91:112-121, (2010)). Further, methods for preparing mRNAs with a Cap 2 structure are even more difficult and results are less predictable. Enzymes for converting Cap 1 to Cap 2 are not currently commercially available.

[0020] Another significant complication of *in vitro* synthesis of mRNAs, especially in large scale manufacturing, is the necessity to isolate and purify the active mRNA molecules carrying Cap from all uncapped mRNA forms which are inactive and in some cases immunogenic. Unfortunately, these methods are not trivial and often require a synthesis of modified mCAP analogs with conjugated affinity tag moieties allowing for easier isolation and purification of capped RNA transcript. Methods of synthesizing mCAP analogs with affinity tags as a reporter/affinity moiety and novel protocols for isolation of capped RNA from the transcription reaction mixture are known in the art (see, for example, U.S. Patent 8,344,118). While these approaches are efficient, they require use of more expensive mCAP analogs and they allow for preparation and isolation of mRNAs containing the Cap 0 structure only.

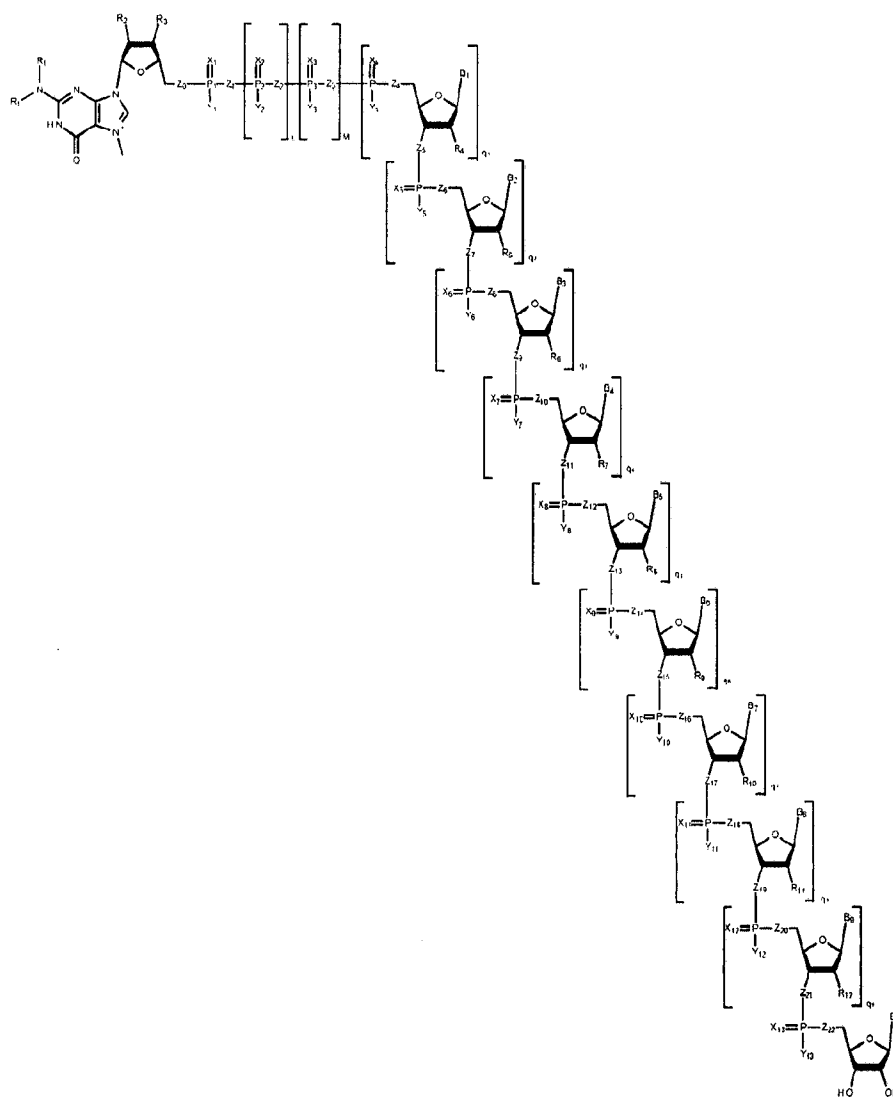
[0021] The *in vitro* synthesis of natural and modified RNAs find use in a variety of applications, including ribozyme, antisense, biophysical and biochemical studies. Additionally, capped mRNA transcripts are used for applications requiring protein synthesis such as *in vivo* expression experiments (using microinjection, transfection and infection), *in vitro* translation experiments and assays as well as various applications in therapeutics, diagnostics, vaccine development, labeling and detection.

[0022] Consequently, there is a need in the industry for compositions and methods that allow for large scale synthesis of mRNAs that are (a) less laborious than conventional methods, (b) eliminate or reduce bi-directional initiation during transcription, (c) result in higher yields of mRNA, at a (d) reduced cost compared to current methods, (e) reduces production of heterogeneous products with different 5'-sequences and (f) does not require additional enzymatic reactions to incorporate Cap 1 and Cap 2 structures into the synthesized mRNA. There is also a need for the synthesis of various mRNAs containing modified and/or unnatural nucleosides, carrying specific modifications and/or affinity tags such as fluorescent dyes, a radioisotope, a mass tag and/or one partner of a molecular binding pair such as biotin

at or near the 5' end of the molecule.

SUMMARY OF THE INVENTION

[0023] Provided herein are methods and compositions for synthesizing 5'-capped RNAs. In one aspect of the present invention, the initiating capped oligonucleotide primers comprise the general structure of Formula I:



wherein

each of B₁ through B₁₀ is independently a natural, modified or unnatural nucleoside base;
 M is 0 or 1;
 L is 0 or 1;
 q₁ is 1;
 each of q₂ through q₉ is independently 0 or 1;
 R₁ is H or methyl;
 R₂ and R₃ are independently H, OH, alkyl, O-alkyl, halogen, amine, azide, a linker or a detectable marker;
 each of X₁ through X₁₃ is independently O or S;
 each of Y₁ through Y₁₃ is independently OH, SH, BH₃, aryl, alkyl, O-alkyl or O-aryl;
 each of Z₁ through Z₂₂ is independently O, S, NH, CH₂, C(halogen)₂ or CH(halogen);
 and
 each of R₄ through R₁₂ are independently H, OH, OMe, linker or a detectable marker.

[0024] In another aspect of the present invention, there are provided RNA molecules comprising a Cap containing initiating oligonucleotide primer of Formula I, pharmaceutical compositions comprising such RNA, cells containing such RNA and cells containing a protein or peptide translated from such RNA.

[0025] In yet another aspect of the present invention, there are provided methods for synthesizing RNA molecules comprising a Cap containing initiating oligonucleotide primer of Formula I into a mixture comprising a polynucleotide template and an RNA polymerase under conditions conducive to transcription by the RNA polymerase of the polynucleotide template, and thereafter incubating the resulting mixture for a time sufficient to allow for transcription of said template.

BRIEF DESCRIPTION OF THE FIGURES

[0026] Figure 1 shows an exemplary method of preparing 7-methylguanosine 5'-diphosphate (pp^{7m}G) from guanosine 5'-diphosphate;

[0027] Figure 2 shows an exemplary method of preparing 7-methylguanosine 5'-diphosphate imidazolide (Im-pp^{7m}G) from pp^{7m}G;

[0028] Figure 3 shows an exemplary method of preparing 3'-O-methylguanosine 5'-phosphate (pG_{3'-Ome}) from 3'-O-methylguanosine;

[0029] Figures 4 shows an exemplary method of preparing 3'-O-methylguanosine 5'-phosphorimidazolide (Im-pG_{3'-Ome}) from pG_{3'-Ome};

[0030] Figure 5 shows an exemplary method of preparing of 3'-*O*-methylguanosine 5'-diphosphate (ppG_{3'-Ome}) from Im-pG_{3'-Ome};

[0031] Figure 6 shows an exemplary method for preparing 7-methyl-3'-*O*-methylguanosine 5-diphosphate (pp^{7m}G_{3'-Ome}) from ppG_{3'-Ome};

[0032] Figure 7 shows an exemplary method for preparing 7-methyl-3'-*O*-methylguanosine 5-diphosphate imidazolide (Im-pp^{7m}G_{3'-Ome}) from pp^{7m}G_{3'-Ome};

[0033] Figure 8 shows a general procedure for the preparation of pN_{2'-OR}pN oligonucleotides (R₁ = H or Me);

[0034] Figure 9 shows a general procedure for the synthesis of initiating oligonucleotides with Cap 0, Cap 1 or Cap 2 structures;

[0035] Figure 10A shows the structure of initiating capped oligonucleotide primers used in examples according to Formula I wherein: B₁ is guanine; M is 0; L is 1; q₁ through q₉ are 0; R₁ is H; R₂ is H; R₃ is *O*-methyl; X₁ is O; X₂ is O; X₁₃ is O; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₂₂ is O;

[0036] Figure 10B shows the structure of initiating capped oligonucleotide primers used in examples according to Formula I wherein: B₁ is guanine; B₁₀ is guanine; M is 0; L is 1; q₁ is 1; q₂ through q₉ are 0; R₁ is H; R₂ is H; R₃ is H; X₁ is O; X₂ is O; X₄ is O; X₁₃ is O; Y₁ is OH; Y₂ is OH; Y₄ is OH; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₄ is O; Z₅ is O; Z₂₂ is O; R₄ is *O*-methyl;

[0037] Figure 10C shows the structure of initiating capped oligonucleotide primers used in examples according to Formula I wherein: B₁ is guanine; B₁₀ is guanine; M is 0; L is 1; q₁ is 1; q₂ through q₉ are 0; R₁ is H; R₂ is H; R₃ is *O*-methyl; X₁ is O; X₂ is O; X₄ is O; X₁₃ is O; Y₁ is OH; Y₂ is OH; Y₄ is OH; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₄ is O; Z₅ is O; Z₂₂ is O; R₄ is *O*-methyl;

[0038] Figure 10D shows the structure of initiating capped oligonucleotide primers used in examples according to Formula I wherein: B₁ is adenine; B₁₀ is guanine; M is 0; L is 1; q₁ is 1; q₂ through q₉ are 0; R₁ is H; R₂ is H; R₃ is H; X₁ is O; X₂ is O; X₄ is O; X₁₃ is O; Y₁ is OH; Y₂ is OH; Y₄ is OH; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₄ is O; Z₅ is O; Z₂₂ is O; R₄ is *O*-methyl;

[0039] Figure 10E shows the structure of initiating capped oligonucleotide primers used

in examples according to Formula I wherein: B₁ is adenine; B₁₀ is guanine; M is O; L is 1; q₁ is 1; q₂ through q₉ are O; R₁ is H; R₂ is H; R₃ is O-methyl; X₁ is O; X₂ is O; X₄ is O; X₁₃ is O; Y₁ is OH; Y₂ is OH; Y₄ is OH; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₄ is O; Z₅ is O; Z₂₂ is O; R₄ is O-methyl.

[0040] Figure 10F shows the structure of initiating capped oligonucleotide primers used in examples according to Formula I wherein: B₁ is cytosine; B₁₀ is guanine; M is O; L is 1; q₁ is 1; q₂ through q₉ are O; R₁ is H; R₂ is H; R₃ is H; X₁ is O; X₂ is O; X₄ is O; X₁₃ is O; Y₁ is OH; Y₂ is OH; Y₄ is OH; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₄ is O; Z₅ is O; Z₂₂ is O; R₄ is O-methyl;

[0041] Figure 10G shows the structure of initiating capped oligonucleotide primers used in examples according to Formula I wherein: B₁ is cytosine; B₁₀ is guanine; M is O; L is 1; q₁ is 1; q₂ through q₉ are O; R₁ is H; R₂ is H; R₃ is O-methyl; X₁ is O; X₂ is O; X₄ is O; X₁₃ is O; Y₁ is OH; Y₂ is OH; Y₄ is OH; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₄ is O; Z₅ is O; Z₂₂ is O; R₄ is O-methyl;

[0042] Figure 10H shows the structure of initiating capped oligonucleotide primers used in examples according to Formula I wherein: B₁ is adenine; B₂ is guanine; B₁₀ is guanine; M is O; L is 1; q₁ is 1; q₂ is 1; q₃ through q₉ are O; R₁ is H; R₂ is H; R₃ is O-methyl; X₁ is O; X₂ is O; X₄ is O; X₅ is O; X₁₃ is O; Y₁ is OH; Y₂ is OH; Y₄ is OH; Y₅ is OH; Y₁₃ is OH; Z₀ is O; Z₁ is O; Z₂ is O; Z₄ is O; Z₅ is O; Z₆ is O; Z₂₂ is O; R₄ is O-methyl; R₅ is O-methyl;

[0043] Figure 11 shows the luciferase activity of mRNAs co-transcriptionally capped mRNAs;

[0044] Figures 12A-12H collectively show the capping efficiency of mRNAs co-transcriptionally capped with 12A) ARCA, 12B) ^{m7}GpppG_{2'}OmePG, 12C) ^{m7}G_{3'}OmepppG_{2'}OmePG, 12D) ^{m7}GpppA_{2'}OmePG, 12E) ^{m7}G_{3'}OmepppA_{2'}OmePG, 12F) ^{m7}GpppC_{2'}OmePG, 12G) ^{m7}G_{3'}OmepppC_{2'}OmePG, 12H) ^{m7}G_{3'}OmepppA_{2'}OmePG_{2'}OmePG. Capping efficiency in Figures 11B-11G is equal to or significantly exceeds that observed in Figure 11A; and

[0045] Figures 13A-13D collectively show the capping efficiency and fidelity of initiation of mRNAs co-transcriptionally capped with ^{m7}GpppA_{2'}OmePG on a transcription template, wherein Figure 13A illustrates the use of 2'-deoxythymidine and 2'-deoxycytidine residues at template positions +1 and +2 using Primer/NTP formulation 2; Figure 13B

illustrates the use of 2'-deoxythymidine and 2'-deoxycytidine residues at template positions +1 and +2 using Primer/NTP formulation 3; Figure 13C illustrates the use of 2'-deoxycytidine residues at template positions +1 and +2 using Primer/NTP formulation 2; and Figure 13D illustrates the use of 2'-deoxycytidine residues at template positions +1 and +2 using Primer/NTP formulation 3.

[0046] Figures 14A and 14B collectively show comparison of translation in differentiated THP-1 cells of mRNA made with m7GpppA2'OmepG initiating capped oligonucleotide on a transcription template with 2'-deoxythymidine and 2'-deoxycytidine residues at template positions +1 and +2 vs. a transcription template with cytidine residues at template positions +1 and +2.

DETAILED DESCRIPTION OF THE INVENTION

[0047] Unless defined otherwise, all terms used herein have the same meaning as are commonly understood by one of skill in the art to which this invention belongs. All patents, patent applications and publications referred to throughout the disclosure herein are incorporated by reference in their entirety. In the event that there is a plurality of definitions for a term herein, those in this section prevail.

[0048] As used herein in connection with numerical values, the term "approximately" or "about" means plus or minus 30% of the indicated value, including all values within the defined range, including the stated value.

[0049] As used herein, the terms "nucleic acid," "nucleotide sequence," or "nucleic acid sequence" refer to an oligonucleotide, polynucleotide, or any fragment thereof, any ribo or deoxyribo derivatives and to naturally occurring or synthetic molecules containing natural and/or modified nucleotide residues and internucleotide linkages. These phrases also refer to DNA or RNA of natural (e.g., genomic) or synthetic origin which may be single-stranded, double-stranded, triple-stranded or tetra-stranded and may represent the sense or the antisense strand, or to any DNA-like or RNA-like material. An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all or most occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of 2'-deoxyribose. Additional alternative nucleic acid backbones suitable for the methods and

compositions provided herein include but are not limited to phosphorothioate, phosphoroselenoate, alkyl phosphotriester, aryl phosphotriester, alkyl phosphonate, aryl phosphonate, phosphoboronate, morpholino nucleic acid (MNA), locked nucleic acids (LNA), peptide nucleic acids (PNA).

[0050] As used herein, the term “primer” or “oligonucleotide primer” refers to a ribo- or deoxyribo- or chimeric ribo/deoxyribo-oligonucleotide, single stranded, may be naturally occurring or synthetic, and usually include a sequence of between about 2 to about 10 nucleotides, about 3 to about 8 nucleotides or about 3 to about 5 nucleotides. Oligonucleotide primers may contain one or more modification groups. Oligonucleotide primers may include RNA, DNA, and/or other modified nucleosides. The skilled artisan is capable of designing and preparing oligonucleotide primers that are appropriate for transcription of DNA template sequence.

[0051] As used herein, the terms “initiating capped oligonucleotide analogs” or “initiating capped oligonucleotide primers” refer to an initiating oligonucleotide primer containing Cap 0, Cap 1, Cap 2 or TMG-Cap structure on 5'-end of the primer. The capped primer has an unmodified or open 3'-OH group and it may be extended by RNA polymerase through the incorporation of an NTP onto the 3'-end of the primer. It is able to initiate *in vitro* transcription under the control of a promoter in a transcription system containing necessary components: DNA template (e.g. DNA plasmid), RNA polymerase, nucleoside 5'-triphosphates and appropriate buffer. Also used herein, “initiating primer” or “initiating oligonucleotide primer” refers to an oligonucleotide, carrying a terminal 3'-OH group that is a valid substrate for RNA polymerase. In certain embodiments, the initiating oligonucleotide primer is a substrate for RNA polymerase and may be elongated by incorporation of NTP onto the 3'-end of the primer. The initiating oligonucleotide primer is complementary to the DNA template at the initiation site.

[0052] As used herein, the term “unsubstituted” or “unmodified” in the context of the initiating capped oligonucleotide primer and NTPs refers to an initiating capped oligonucleotide primer and NTPs that have not been modified.

[0053] As used herein, the term “modified initiating capped oligonucleotide primer” refers to an initiating capped oligonucleotide primer that contains one or more additional modification groups.

[0054] As used herein, the term “modification group” refers to any chemical moiety that may be attached to the initiating primer at locations, which include, but are not limited to, the sugar, nucleoside base, triphosphate bridge, and/or internucleotide phosphate (*e.g.*, U.S. Patent Application No. 20070281308). The modification group of an initiating capped oligonucleotide primer may be a group of any nature compatible with the process of transcription.

[0055] As used herein, the term “internucleotide linkage” refers to the bond or bonds that connect two nucleosides of an oligonucleotide primer or nucleic acid and may be a natural phosphodiester linkage or modified linkage.

[0056] As used herein, the term “label” or “detectable label” refers to any compound or combination of compounds that may be attached or otherwise associated with a molecule so that the molecule can be detected directly or indirectly by detecting the label. A detectable label can be a radioisotope (*e.g.*, carbon, phosphorus, iodine, indium, sulfur, tritium *etc.*), a mass isotope (*e.g.*, H^2 , C^{13} or N^{15}), a dye or fluorophore (*e.g.*, cyanine, fluorescein or coumarin), a hapten (*e.g.*, biotin) or any other agent that can be detected directly or indirectly.

[0057] As used herein, the term “hybridize” or “specifically hybridize” refers to a process where initiating capped oligonucleotide primer anneals to a DNA template under appropriately stringent conditions during a transcription reaction. Hybridizations to DNA are conducted with an initiating capped oligonucleotide primer which, in certain embodiments, is 3-10 nucleotides in length including the 5'-5' inverted cap structure. Nucleic acid hybridization techniques are well known in the art (*e.g.*, Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Press, Plainview, N.Y.(1989); Ausubel, F.M., *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, Secaucus, N.J. (1994)).

[0058] As used herein, the term “complement,” “complementary,” or “complementarity” in the context of a complex of an initiating capped oligonucleotide primer and a DNA template refers to standard Watson/Crick base pairing rules. For example, the sequence “5'-A-G-T-C-3'” is complementary to the sequence “3'-T-C-A-G-5'.” Certain non-natural or synthetic nucleotides may be included in the nucleic acids described herein; these include but not limited to, base and sugar modified nucleosides, nucleotides, and nucleic acids, such as inosine, 7-deazaguanosine, 2'-O-methylguanosine, 2'-fluoro-2'-deoxycytidine, pseudouridine, Locked Nucleic Acids (LNA), and Peptide Nucleic Acids (PNA).

Complementarity does not need to be perfect; duplexes may contain mismatched base pairs, degenerative, or unmatched nucleotides. Those skilled in the art can determine duplex stability empirically considering a number of variables including, for example, the length of the oligonucleotide, base composition and sequence of the oligonucleotide, incidence of mismatched base pairs, ionic strength, components of the hybridization buffer and reaction conditions.

[0059] Complementarity may be “complete” or “total” where all of the nucleotide bases of two nucleic acid strands are matched according to recognized base pairing rules, it may be “partial” in which only some of the nucleotide bases of an initiating capped oligonucleotide primer and a DNA target are matched according to recognized base pairing rules or it may be “absent” where none of the nucleotide bases of two nucleic acid strands are matched according to recognized base pairing rules. The degree of complementarity between of an initiating capped oligonucleotide primer and a DNA template may have a significant effect on the strength of hybridization between the initiating capped oligonucleotide and the DNA template and correspondingly the efficiency of the reaction. The term complementarity may also be used in reference to individual nucleotides. For example, a particular nucleotide within an oligonucleotide may be noted for its complementarity, or lack thereof, to a nucleotide within another strand, in contrast or comparison to the complementarity between the rest of an initiating capped oligonucleotide primer and DNA strand.

[0060] As used herein the term “complete”, “total” or “perfectly” complementary means that each of the nucleotide bases of an initiating capped oligonucleotide primer and a DNA target are matched exactly according to recognized base pairing rules.

[0061] As used herein, the term “substantially complementary” refers to two sequences that hybridize under stringent hybridization conditions. Those skilled in the art will understand that substantially complementary sequences need not hybridize along their entire length. In particular, substantially complementary sequences may comprise a contiguous sequence of bases that do not hybridize to a target sequence and may be positioned 3’ or 5’ to a contiguous sequence of bases that hybridize under stringent hybridization conditions to the target sequence.

[0062] As used herein, the term “specific” when used in reference to an initiating capped oligonucleotide primer sequence and its ability to hybridize to a DNA template is a sequence that has at least 50% sequence identity with a portion of the DNA template when the

initiating capped oligonucleotide primer and DNA strand are aligned. Higher levels of sequence identity that may be preferred include at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, and most preferable 100% sequence identity.

[0063] As used herein, the term “nucleoside” includes all naturally occurring nucleosides, including all forms of nucleoside bases and furanosides found in nature. Base rings most commonly found in naturally occurring nucleosides are purine and pyrimidine rings. Naturally occurring purine rings include, for example, adenine, guanine, and N⁶-methyladenine. Naturally occurring pyrimidine rings include, for example, cytosine, thymine, 5-methylcytosine, pseudouracyl. Naturally occurring nucleosides for example include, but are not limited to, ribo, 2'-O-methyl or 2'-deoxyribo derivatives of adenosine, guanosine, cytidine, thymidine, uridine, inosine, 7-methylguanosine or pseudouridine.

[0064] As used herein, the terms “nucleoside analogs,” “modified nucleosides,” or “nucleoside derivatives” include synthetic nucleosides as described herein. Nucleoside derivatives also include nucleosides having modified base or/and sugar moieties, with or without protecting groups and include, for example, 2'-deoxy-2'-fluorouridine, 5-fluorouridine and the like. The compounds and methods provided herein include such base rings and synthetic analogs thereof, as well as unnatural heterocycle-substituted base sugars, and acyclic substituted base sugars. Other nucleoside derivatives that may be utilized with the present invention include, for example, LNA nucleosides, halogen-substituted purines (e.g., 6-fluoropurine), halogen-substituted pyrimidines, N⁶-ethyladenine, N⁴-(alkyl)-cytosines, 5-ethylcytosine, and the like (U.S. Patent No. 6,762,298).

[0065] As used herein, the terms “universal base,” “degenerate base,” “universal base analog” and “degenerate base analog” include, for example, a nucleoside analog with an artificial base which is, in certain embodiments, recognizable by RNA polymerase as a substitute for one of the natural NTPs (e.g., ATP, UTP, CTP and GTP) or other specific NTP. Universal bases or degenerate bases are disclosed in Loakes, D., *Nucleic Acids Res.*, 29:2437-2447 (2001); Crey-Desbiolles, C., *et. al.*, *Nucleic Acids Res.*, 33:1532-1543 (2005); Kincaid, K., *et. al.*, *Nucleic Acids Res.*, 33:2620-2628 (2005); Preparata, FP, Oliver, JS, *J. Comput. Biol.* 753-765 (2004); and Hill, F., *et. al.*, *Proc Natl Acad. Sci. U S A*, 95:4258-4263 (1998).

[0066] As used herein, the term “modified NTP” refers to a nucleoside 5'-triphosphate having a chemical moiety group bound at any position, including the sugar, base, triphosphate chain, or any combination of these three locations. Examples of such NTPs can

be found, for example in "Nucleoside Triphosphates and Their Analogs: Chemistry, Biotechnology and Biological Applications," Vaghefi, M., ed., Taylor and Francis, Boca Raton (2005).

[0067] As used herein, the term "modified oligonucleotide" includes, for example, an oligonucleotide containing a modified nucleoside, a modified internucleotide linkage, or having any combination of modified nucleosides and internucleotide linkages. Examples of oligonucleotide internucleotide linkage modifications including phosphorothioate, phosphotriester and methylphosphonate derivatives (Stec, W.J., *et al.*, *Chem. Int. Ed. Engl.*, 33:709-722 (1994); Lebedev, A.V., *et al.*, *E. Perspect. Drug Discov. Des.*, 4:17-40 (1996); and Zon, *et al.*, U.S. Patent Application No. 20070281308). Other examples of internucleotide linkage modifications may be found in Waldner, *et al.*, *Bioorg. Med. Chem. Letters* 6:2363-2366 (1996).

[0068] The term "promoter" as used herein refers to a region of dsDNA template that directs and controls the initiation of transcription of a particular DNA sequence (e.g. gene). Promoters are located on the same strand and upstream on the DNA (towards the 5' region of the sense strand). Promoters are typically immediately adjacent to (or partially overlap with) the DNA sequence to be transcribed. Nucleotide positions in the promoter are designated relative to the transcriptional start site, where transcription of DNA begins (position +1). The initiating oligonucleotide primer is complementary to initiation site of promoter sequence (which, in certain embodiments, is at positions +1 and +2 and, in the case of initiating tetramers, at positions +1, +2 and +3).

[0069] As used herein, the terms "transcription" or "transcription reaction" refers to methods known in the art for enzymatically making RNA that is complementary to DNA template, thereby producing the number of RNA copies of a DNA sequence. The RNA molecule synthesized in transcription reaction called "RNA transcript", "primary transcript" or "transcript". Transcription reaction involving the compositions and methods provided herein employs "initiating capped oligonucleotide primers". Transcription of DNA template may be exponential, nonlinear or linear. A DNA template may be a double stranded linear DNA, a partially double stranded linear DNA, circular double stranded DNA, DNA plasmid, PCR amplicon, a modified nucleic acid template which is compatible with RNA polymerase.

[0070] As used herein, the term "acyl" denotes the group $-C(O)R^a$, where R^a is hydrogen, lower alkyl, cycloalkyl, heterocyclyl, aryl, heteroaryl, and the like.

[0071] As used herein, the term “substituted acyl” denotes the group $-C(O)R^a$, where R^a is substituted lower alkyl, substituted cycloalkyl, substituted heterocyclyl, substituted aryl, substituted heteroaryl, and the like.

[0072] As used herein, the term “acyloxy” denotes the group $-OC(O)R^b$, where R^b is hydrogen, lower alkyl, substituted lower alkyl, cycloalkyl, substituted cycloalkyl, heterocyclyl, substituted heterocyclyl, aryl, substituted aryl, heteroaryl, substituted heteroaryl, and the like.

[0073] As used herein, the term “alkyl” refers to a single bond chain of hydrocarbons ranging, in some embodiments, from 1-20 carbon atoms, and ranging in some embodiments, from 1-8 carbon atoms; examples include methyl, ethyl, propyl, isopropyl, n-butyl, sec-butyl, isobutyl, tert-butyl, pentyl, isoamyl, hexyl, octyl, dodecanyl, and the like.

[0074] As used herein, the term “lower alkyl” refers to a straight chain or a branched chain of hydrocarbons ranging, in some embodiments, from 1-6 carbon atoms, and ranging in some embodiments from 2-5 carbon atoms. Examples include ethyl, propyl, isopropyl, and the like.

[0075] As used herein, the term “alkenyl” refers to a straight-chain or branched-chain hydrocarbyl, which has one or more double bonds and, unless otherwise specified, contains from about 2 to about 20 carbon atoms, and ranging in some embodiments from about 2 to about 10 carbon atoms, and ranging in some embodiments from about 2 to about 8 carbon atoms, and ranging in some embodiments from about 2 to about 6 carbon atoms. Examples of alkenyl radicals include vinyl, allyl, 1,4-butadienyl, isopropenyl, and the like.

[0076] As used herein, the term “alkenylaryl” refers to alkenyl-substituted aryl groups and “substituted alkenylaryl” refers to alkenylaryl groups further bearing one or more substituents as set forth herein.

[0077] As used herein, the term “alkenylene” refers to divalent straight-chain or branched-chain hydrocarbyl groups having at least one carbon-carbon double bond, and typically containing 2-20 carbon atoms, and ranging in some embodiments from 2-12 carbon atoms, and ranging in some embodiments from 2-8 carbon atoms, and “substituted alkenylene” refers to alkenylene groups further bearing one or more substituents as set forth herein.

[0078] As used herein, the term “alkylene” refers to divalent hydrocarbyl group

containing 1-20 carbon atoms, and ranging in some embodiments from 1-15 carbon atoms, straight-chain or branched-chain, from which two hydrogen atoms are taken from the same carbon atom or from different carbon atoms. Examples of alkylene include, but are not limited to, methylene ($-\text{CH}_2-$), ethylene ($-\text{CH}_2\text{CH}_2-$), and the like.

[0079] As used herein, the term “alkynyl” refers to a straight-chain or branched-chain hydrocarbyl, which has one or more triple bonds and contains from about 2-20 carbon atoms, and ranging in some embodiments from about 2-10 carbon atoms, and ranging in some embodiments from about 2- 8 carbon atoms, and ranging in some embodiments from about 2-6 carbon atoms. Examples of alkynyl radicals include ethynyl, propynyl (propargyl), butynyl, and the like.

[0080] As used herein, the term “alkynylaryl” refers to alkynyl-substituted aryl groups and “substituted alkynylaryl” refers to alkynylaryl groups further bearing one or more substituents as set forth herein.

[0081] As used herein, the term “alkoxy” denotes the group $-\text{OR}^c$, where R^c is lower alkyl, substituted lower alkyl, aryl, substituted aryl, aralkyl, substituted aralkyl, heteroalkyl, heteroarylalkyl, cycloalkyl, substituted cycloalkyl, cycloheteroalkyl, or substituted cycloheteroalkyl as defined.

[0082] As used herein, the term “lower alkoxy” denotes the group $-\text{OR}^d$, where R^d is lower alkyl.

[0083] As used herein, the term “alkylaryl” refers to alkyl-substituted aryl groups and “substituted alkylaryl” refers to alkylaryl groups further bearing one or more substituents as set forth herein.

[0084] As used herein, the term “alkylthio” refers to the group $-\text{S-R}^h$, where R^h is alkyl.

[0085] As used herein, the term “substituted alkylthio” refers to the group $-\text{S-R}^i$, where R^i is substituted alkyl.

[0086] As used herein, the term “alkynylene” refers to divalent straight-chain or branched-chain hydrocarbyl groups having at least one carbon-carbon triple bond, and typically having in the range of about 2-12 carbon atoms, and ranging in some embodiments from about 2-8 carbon atoms, and “substituted alkynylene” refers to alkynylene groups further bearing one or more substituents as set forth herein.

[0087] As used herein, the term “amido” denotes the group $-C(O)NR^jR^j$, where R^j and R^j may independently be hydrogen, lower alkyl, substituted lower alkyl, alkyl, substituted alkyl, aryl, substituted aryl, heteroaryl, or substituted heteroaryl.

[0088] As used herein, the term “substituted amido” denotes the group $-C(O)NR^kR^{k'}$, where R^k and $R^{k'}$ are independently hydrogen, lower alkyl, substituted lower alkyl, aryl, substituted aryl, heteroaryl, or substituted heteroaryl, provided, however, that at least one of R^k and $R^{k'}$ is not hydrogen. $R^kR^{k'}$ in combination with the nitrogen may form an optionally substituted heterocyclic or heteroaryl ring.

[0089] As used herein, the term “amino” or “amine” denotes the group $-NR^nR^n$, where R^n and R^n may independently be hydrogen, lower alkyl, substituted lower alkyl, alkyl, substituted alkyl, aryl, substituted aryl, heteroaryl, or substituted heteroaryl as defined herein. A “divalent amine” denotes the group $-NH-$. A “substituted divalent amine” denotes the group $-NR-$ where R is lower alkyl, substituted lower alkyl, alkyl, substituted alkyl, aryl, substituted aryl, heteroaryl, or substituted heteroaryl.

[0090] As used herein, the term “substituted amino” or “substituted amine” denotes the group $-NR^pR^p$, where R^p and R^p are independently hydrogen, lower alkyl, substituted lower alkyl, alkyl, substituted alkyl, aryl, substituted aryl, heteroaryl, substituted heteroaryl, provided, however, that at least one of R^p and R^p is not hydrogen. R^pR^p in combination with the nitrogen may form an optionally substituted heterocyclic, or heteroaryl ring.

[0091] As used herein, the term “aroyl” refers to aryl-carbonyl species such as benzoyl and “substituted aroyl” refers to aroyl groups further bearing one or more substituents as set forth herein.

[0092] As used herein, the term “aryl” alone or in combination refers to phenyl, naphthyl or fused aromatic heterocyclic optionally with a cycloalkyl of 5-10 ring members, and in some embodiments 5-6, ring members and/or optionally substituted with 1 to 3 groups or substituents of halo, hydroxy, alkoxy, alkylthio, alkylsulfinyl, alkylsulfonyl, acyloxy, aryloxy, heteroaryloxy, amino optionally mono- or di-substituted with alkyl, aryl or heteroaryl groups, amidino, urea optionally substituted with alkyl, aryl, heteroaryl or heterocyclyl groups, aminosulfonyl optionally N-mono- or N,N-di-substituted with alkyl, aryl or heteroaryl groups, alkylsulfonylamino, arylsulfonylamino, heteroarylsulfonylamino, alkylcarbonylamino, arylcarbonylamino, heteroarylcarbonylamino, or the like.

[0093] As used herein, the term “aryloxy” denotes the group -OAr, where Ar is an aryl, or substituted aryl group.

[0094] As used herein, the term “carbocycle” refers to a saturated, unsaturated, or aromatic group having a single ring or multiple condensed rings composed of linked carbon atoms. The ring(s) can optionally be unsubstituted or substituted with, *e.g.*, halogen, lower alkyl, alkoxy, alkylthio, acetylene ($-C\equiv CH$), amino, amido, azido, carboxyl, hydroxyl, aryl, aryloxy, heterocycle, heteroaryl, substituted heteroaryl, nitro ($-NO_2$), cyano ($-CN$), thiol ($-SH$), sulfamido ($-S(O)_2NH_2$), and the like.

[0095] As used herein, the term “guanidinyll” denotes the group $-N=C(NH_2)_2$ and “substituted guanidinyll” denotes the group $-N=C(NR_2)_2$, where each R is independently H, alkyl, substituted alkyl, aryl, or substituted aryl as set forth herein.

[0096] As used herein, the term “halo” or “halogen” refers to all halogens, *i.e.*, chloro (Cl), fluoro (F), bromo (Br), and iodo (I).

[0097] As used herein, the term “heteroaryl” refers to a monocyclic aromatic ring structure containing 5 or 6 ring atoms, or a bicyclic aromatic group having 8-10 atoms, containing one or more, and in some embodiments 1-4, and in some embodiments 1-3, and in some embodiments 1-2 heteroatoms independently selected from the group O, S, and N, and optionally substituted with 1-3 groups or substituents of halo, hydroxy, alkoxy, alkylthio, alkylsulfinyl, alkylsulfonyl, acyloxy, aryloxy, heteroaryloxy, amino optionally mono- or di-substituted with alkyl, aryl or heteroaryl groups, amidino, urea optionally substituted with alkyl, aryl, heteroaryl, or heterocyclyl groups, aminosulfonyl optionally N-mono- or N,N-di-substituted with alkyl, aryl or heteroaryl groups, alkylsulfonylamino, arylsulfonylamino, heteroarylsulfonylamino, alkylcarbonylamino, arylcarbonylamino, heteroarylcarbonylamino, or the like. Heteroaryl is also intended to include oxidized S or N, such as sulfinyl, sulfonyl, and N-oxide of a tertiary ring nitrogen. A carbon or nitrogen atom is the point of attachment of the heteroaryl ring structure such that a stable aromatic ring is retained. Examples of heteroaryl groups are phthalimide, pyridinyl, pyridazinyl, pyrazinyl, quinazolinyl, purinyl, indolyl, quinolinyl, pyrimidinyl, pyrrolyl, oxazolyl, thiazolyl, thienyl, isoxazolyl, oxathiadiazolyl, isothiazolyl, tetrazolyl, imidazolyl, triazinyl, furanyl, benzofuryl, indolyl, and the like. A substituted heteroaryl contains a substituent attached at an available carbon or nitrogen to produce a stable compound.

[0098] As used herein, the term “substituted heteroaryl” refers to a heterocycle optionally mono or poly substituted with one or more functional groups, *e.g.*, halogen, lower alkyl, lower alkoxy, alkylthio, acetylene, amino, amido, carboxyl, hydroxyl, aryl, aryloxy, heterocycle, substituted heterocycle, hetaryl, substituted hetaryl, nitro, cyano, thiol, sulfamido, and the like.

[0099] As used herein, the term “heterocycle” refers to a saturated, unsaturated, or aromatic group having a single ring (*e.g.*, morpholino, pyridyl or furyl) or multiple condensed rings (*e.g.*, naphthpyridyl, quinoxalyl, quinoliny, indoliziny or benzo[b]thienyl) and having carbon atoms and at least one hetero atom, such as N, O or S, within the ring, which can optionally be unsubstituted or substituted with, *e.g.*, halogen, lower alkyl, lower alkoxy, alkylthio, acetylene, amino, amido, carboxyl, hydroxyl, aryl, aryloxy, heterocycle, hetaryl, substituted hetaryl, nitro, cyano, thiol, sulfamido, and the like.

[00100] As used herein, the term “substituted heterocycle” refers to a heterocycle substituted with 1 or more, *e.g.*, 1, 2, or 3, substituents selected from the group consisting of optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, halo, hydroxy, alkoxy, alkylthio, alkylsulfinyl, alkylsulfonyl, acyloxy, aryl, substituted aryl, aryloxy, heteroaryloxy, amino, amido, amidino, urea optionally substituted with alkyl, aryl, heteroaryl or heterocyclyl groups, aminosulfonyl optionally N-mono- or N,N-di-substituted with alkyl, aryl or heteroaryl groups, alkylsulfonylamino, arylsulfonylamino, heteroarylsulfonylamino, alkylcarbonylamino, arylcarbonylamino, heteroarylcarbonylamino, acyl, carboxyl, heterocycle, substituted heterocycle, hetaryl, substituted hetaryl, nitro, cyano, thiol, sulfonamido, and oxo, attached at any available point to produce a stable compound.

[00101] As used herein, the term “hydrocarbyl” refers to any organic radical where the backbone thereof comprises carbon and hydrogen only. Thus, hydrocarbyl embraces alkyl, cycloalkyl, alkenyl, cycloalkenyl, alkynyl, aryl, alkylaryl, arylalkyl, arylalkenyl, alkenylaryl, arylalkynyl, alkynylaryl, and the like.

[00102] As used herein, the term “substituted hydrocarbyl” refers to any of the above-referenced hydrocarbyl groups further bearing one or more substituents selected from hydroxy, hydrocarbyloxy, substituted hydrocarbyloxy, alkylthio, substituted alkylthio, arylthio, substituted arylthio, amino, alkylamino, substituted alkylamino, carboxy, -C(S)SR, -C(O)SR, -C(S)NR₂, where each R is independently hydrogen, alkyl or substituted alkyl, nitro, cyano, halo, -SO₃M or -OSO₃M, where M is H, Na, K, Zn, Ca, or meglumine, guanidiny, and the like.

substituted guanidiny, hydrocarbyl, substituted hydrocarbyl, hydrocarbylcarbonyl, substituted hydrocarbylcarbonyl, hydrocarbyloxycarbonyl, substituted hydrocarbyloxycarbonyl, hydrocarbylcarbonyloxy, substituted hydrocarbylcarbonyloxy, acyl, acyloxy, heterocyclic, substituted heterocyclic, heteroaryl, substituted heteroaryl, heteroarylcarbonyl, substituted heteroarylcarbonyl, carbamoyl, monoalkylcarbamoyl, dialkylcarbamoyl, arylcarbamoyl, a carbamate group, a dithiocarbamate group, aroyl, substituted aroyl, organosulfonyl, substituted organosulfonyl, organosulfinyl, substituted alkylsulfinyl, alkylsulfonylamino, substituted alkylsulfonylamino, arylsulfonylamino, substituted arylsulfonylamino, a sulfonamide group, sulfuryl, and the like, including two or more of the above-described groups attached to the hydrocarbyl moiety by such linker/spacer moieties as -O-, -S-, -NR-, where R is hydrogen, alkyl or substituted alkyl, -C(O)-, -C(S)-, -C(=NR')-, -C(=CR')-, where R' is alkyl or substituted alkyl, -O-C(O)-, -O-C(O)-O-, -O-C(O)-NR- (or -NR-C(O)-O-), -NR-C(O)-, -NR-C(O)-NR-, -S-C(O)-, -S-C(O)-O-, -S-C(O)-NR-, -O-S(O)₂-, -O-S(O)₂-O-, -O-S(O)₂-NR-, -O-S(O)-, -O-S(O)-O-, -O-S(O)-NR-, -O-NR-C(O)-, -O-NR-C(O)-O-, -O-NR-C(O)-NR-, -NR-O-C(O)-, -NR-O-C(O)-O-, -NR-O-C(O)-NR-, -O-NR-C(S)-, -O-NR-C(S)-O-, -O-NR-C(S)-NR-, -NR-O-C(S)-, -NR-O-C(S)-O-, -NR-O-C(S)-NR-, -O-C(S)-, -O-C(S)-O-, -O-C(S)-NR- (or -NR-C(S)-O-), -NR-C(S)-, -NR-C(S)-NR-, -S-S(O)₂-, -S-S(O)₂-O-, -S-S(O)₂-NR-, -NR-O-S(O)-, -NR-O-S(O)-O-, -NR-O-S(O)-NR-, -NR-O-S(O)₂-, -NR-O-S(O)₂-O-, -NR-O-S(O)₂-NR-, -O-NR-S(O)-, -O-NR-S(O)-O-, -O-NR-S(O)-NR-, -O-NR-S(O)₂-O-, -O-NR-S(O)₂-NR-, -O-NR-S(O)₂-, -O-P(O)R₂-, -S-P(O)R₂-, or -NR-P(O)R₂-, where each R is independently hydrogen, alkyl or substituted alkyl, and the like.

[00103] As used herein, the term "hydroxyl" or "hydroxy" refers to the group -OH.

[00104] As used herein, the term "oxo" refers to an oxygen substituent double bonded to the attached carbon.

[00105] As used herein, the term "sulfinyl" denotes the group -S(O)-.

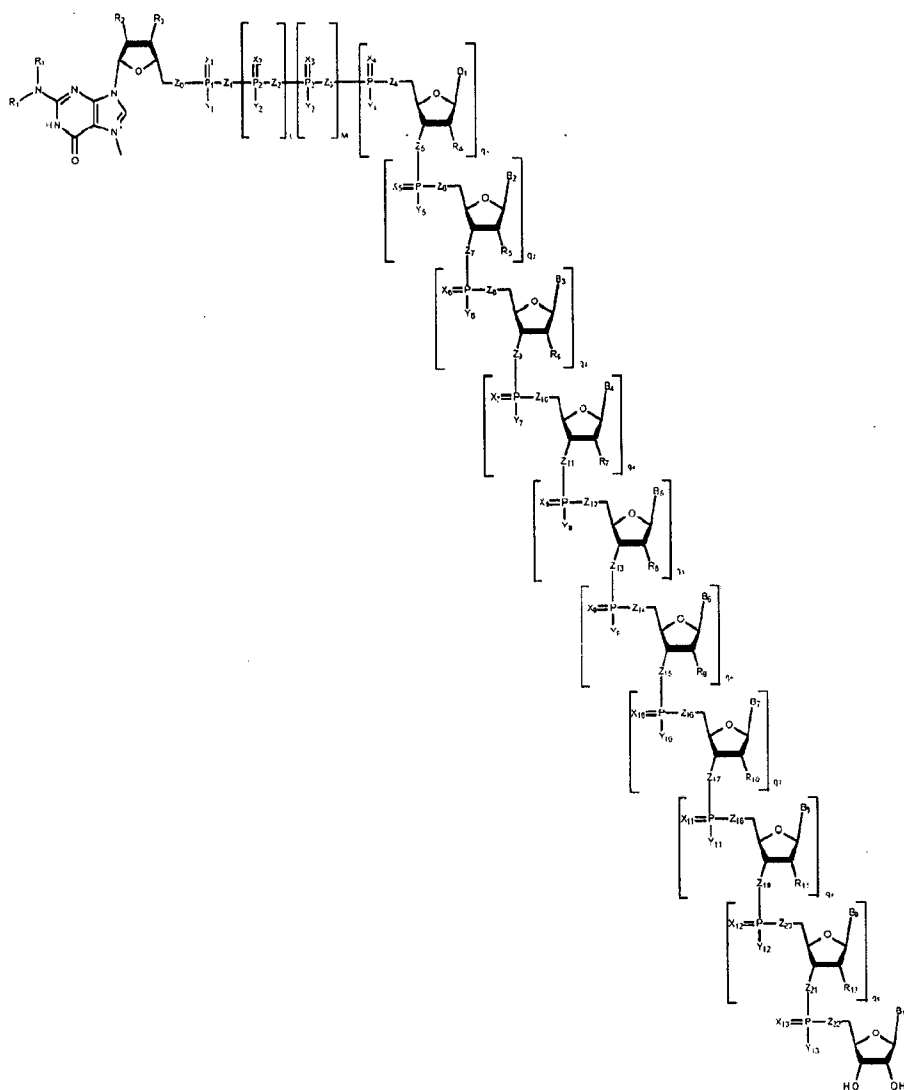
[00106] As used herein, the term "substituted sulfinyl" denotes the group -S(O)R¹, where R¹ is lower alkyl, substituted lower alkyl, cycloalkyl, substituted cycloalkyl, cycloalkylalkyl, substituted cycloalkylalkyl, heterocyclyl, substituted heterocyclyl, heterocyclylalkyl, substituted heterocyclylalkyl, aryl, substituted aryl, heteroaryl, substituted heteroaryl, heteroaralkyl, substituted heteroaralkyl, aralkyl, or substituted aralkyl.

[00107] As used herein, the term “sulfonyl” denotes the group $-S(O)_2-$.

[00108] As used herein, the term “substituted sulfonyl” denotes the group $-S(O)_2R^1$, where R^1 is lower alkyl, substituted lower alkyl, cycloalkyl, substituted cycloalkyl, cycloalkylalkyl, substituted cycloalkylalkyl, heterocyclyl, substituted heterocyclyl, heterocyclalkyl, substituted heterocyclalkyl, aryl, substituted aryl, heteroaryl, substituted heteroaryl, heteroaralkyl, substituted heteroaralkyl, aralkyl, or substituted aralkyl.

[00109] As used herein, the term “sulfuryl” denotes the group $-S(O)_2-$.

[00110] The present invention provides methods and compositions for synthesizing 5' Capped RNAs wherein the initiating capped oligonucleotide primers have the general form ${}^{m7}\text{Gppp}[\text{N}_{2'}\text{Omc}]_n[\text{N}]_m$ wherein ${}^{m7}\text{G}$ is N7-methylated guanosine or any guanosine analog, N is any natural, modified or unnatural nucleoside “n” can be any integer from 1 to 4 and “m” can be an integer from 1 to 9. In one aspect of the invention, the initiating capped oligonucleotide primers have the structure of Formula I:



Formula I

wherein:

each of B₁ through B₁₀ is independently a natural, modified or unnatural nucleoside

base;

M is 0 or 1;

L is 0 or 1;

q₁ is 1;

each of q₂ through q₉ is independently 0 or 1;

R_1 is H or methyl;

R_2 and R_3 are independently H, OH, alkyl, O-alkyl, amine, azide, halogen, a linker or a detectable marker;

each of X_1 through X_{13} is independently O or S;

each of Y_1 through Y_{13} is independently OH, SH, BH_3 , aryl, alkyl, O-alkyl or O-aryl;

each of Z_1 through Z_{27} is independently O, S, NH, CH_2 , $C(halogen)_2$ or $CH(halogen)$

and

each of R_4 through R_{12} are independently, H, OH, OMe or a detectable marker.

[00111] In certain embodiments the initiating capped oligonucleotide primer is a trimer ($q^2-q^9=0$), tetramer ($q^3-q^9=0$), pentamer ($q^4-q^9=0$), hexamer ($q^5-q^9=0$), heptamer ($q^6-q^9=0$), octamer ($q^7-q^9=0$), nanomer ($q^8-q^9=0$), decamer ($q^9=0$) or an undecamer. A number of examples of the initiating capped oligonucleotide trimer primers are presented in Table I below:

Table I

	Sequence of Initiating Capped Oligonucleotide Primer
	$m^7GpppApA$
	$m^7GpppApC$
	$m^7GpppApG$
	$m^7GpppApU$
	$m^7GpppCpA$
	$m^7GpppCpC$
	$m^7GpppCpG$
	$m^7GpppCpU$
	$m^7GpppGpA$
	$m^7GpppGpC$
	$m^7GpppGpG$
	$m^7GpppGpU$
	$m^7GpppUpA$
	$m^7GpppUpC$
	$m^7GpppUpG$
	$m^7GpppUpU$
	$m^7G_{3'OMe}pppApA$
	$m^7G_{3'OMe}pppApC$
	$m^7G_{3'OMe}pppApG$
	$m^7G_{3'OMe}pppApU$
	$m^7G_{3'OMe}pppCpA$
	$m^7G_{3'OMe}pppCpC$
	$m^7G_{3'OMe}pppCpG$
	$m^7G_{3'OMe}pppCpU$
	$m^7G_{3'OMe}pppGpA$
	$m^7G_{3'OMe}pppGpC$

	m ⁷ G _{3'} Ome pppGpG
	m ⁷ G _{3'} Ome pppGpU
	m ⁷ G _{3'} Ome pppUpA
	m ⁷ G _{3'} Ome pppUpC
	m ⁷ G _{3'} Ome pppUpG
	m ⁷ G _{3'} Ome pppUpU
	m ⁷ G _{3'} Ome pppA _{2'} Ome pA
	m ⁷ G _{3'} Ome pppA _{2'} Ome pC
	m ⁷ G _{3'} Ome pppA _{2'} Ome pG
	m ⁷ G _{3'} Ome pppA _{2'} Ome pU
	m ⁷ G _{3'} Ome pppC _{2'} Ome pA
	m ⁷ G _{3'} Ome pppC _{2'} Ome pC
	m ⁷ G _{3'} Ome pppC _{2'} Ome pG
	m ⁷ G _{3'} Ome pppC _{2'} Ome pU
	m ⁷ G _{3'} Ome pppG _{2'} Ome pA
	m ⁷ G _{3'} Ome pppG _{2'} Ome pC
	m ⁷ G _{3'} Ome pppG _{2'} Ome pG
	m ⁷ G _{3'} Ome pppG _{2'} Ome pU
	m ⁷ G _{3'} Ome pppU _{2'} Ome pA
	m ⁷ G _{3'} Ome pppU _{2'} Ome pC
	m ⁷ G _{3'} Ome pppU _{2'} Ome pG
	m ⁷ G _{3'} Ome pppU _{2'} Ome pU
	m ⁷ GpppA _{2'} Ome pA
	m ⁷ GpppA _{2'} Ome pC
	m ⁷ GpppA _{2'} Ome pG
	m ⁷ GpppA _{2'} Ome pU
	m ⁷ GpppC _{2'} Ome pA
	m ⁷ GpppC _{2'} Ome pC
	m ⁷ GpppC _{2'} Ome pG
	m ⁷ GpppC _{2'} Ome pU
	m ⁷ GpppG _{2'} Ome pA
	m ⁷ GpppG _{2'} Ome pC
	m ⁷ GpppG _{2'} Ome pG
	m ⁷ GpppG _{2'} Ome pU
	m ⁷ GpppU _{2'} Ome pA
	m ⁷ GpppU _{2'} Ome pC
	m ⁷ GpppU _{2'} Ome pG
	m ⁷ GpppU _{2'} Ome pU

[00112] Other initiating capped oligonucleotide primers encompassed by this invention include those primers having known or novel base analogues. Methods for synthesizing - initiating capped oligonucleotide primers are exemplified in the examples below.

Transcription

[00113] In Eukaryotes, transcription of messenger RNAs (mRNAs) is done by RNA polymerase II. This is a complicated multi-subunit enzyme with complex regulation. To carry out large scale transcription *in vitro*, researches commonly use single subunit phage

polymerases derived from T7, T3, SP6, K1-5, K1E, K1F or K11 bacteriophages. This family of polymerases has simple, minimal promoter sequences of ~17 nucleotides which require no accessory proteins and have minimal constraints of the initiating nucleotide sequence. While this application focuses on T7 RNA Polymerase (T7 RNAP), one skilled in the art would understand that this invention could be practiced with other RNA polymerases.

[00114] T7 RNAP exists in at least two protein states. The first is referred to as the “abortive complex” and is associated with transcriptional initiation. The second is a very processive conformation called the “elongation complex”. *In vitro* transcription can be broken into six steps: 1) binding of the RNA polymerase to the promoter sequence, 2) initiation of transcription, 3) non-processive elongation termed abortive transcription during which the polymerase frequently releases the DNA template and short abortive transcripts 4) conversion of the open complex to the closed complex, 5) processive elongation and 6) transcriptional termination. A significant amount of RNA produced during transcription consists of short abortive fragments of ~2-8 nucleotides in length (Biochemistry 19:3245-3253 (1980); Nucleic Acids Res. 9:31-45 (1981); Nucleic Acids Res. 15:8783-8798 (1987); Biochemistry 27:3966-3974 (1988)). After synthesis of about 10–14 bases, RNA polymerases escape from abortive cycling, at the same time losing sequence-specific contacts with the promoter DNA, and forming a processive elongation complex, in which the RNA chain is extended in a sequence-independent manner (*J. Mol. Biol.* 183:165-177(1985); *Proc. Natl. Acad. Sci. U. S. A.* 83:3614-3618(1986); *Mol. Cell Biol.* 7:3371-3379 (1987)).

[00115] The consensus sequence for the most active Class III T7 promoters encompasses 17 bp of sequence upstream, and 6 bp downstream, of the transcription start site (*Cell* 16:815-25. (1979)). The position of the first transcribed nucleotide is commonly referred to as the +1 transcript nucleotide of the RNA, the second transcribed nucleotide as +2 transcript nucleotide and so on (Table 2). During transcription, the two strands are melted to form a transcription bubble and the bottom strand of the duplex (shown 3' to 5' in Table 2) is the template for transcription. For transcript nucleotides +3 and beyond, the template strand defines the identity of the transcribed nucleotides primarily through Watson-Crick base pairing interactions. Here the nucleotide encoding the first RNA transcript nucleotide is defined as the +1 nucleotide of the template. In the example shown in Table 2, the +1 transcript nucleotide is G and the +1 template nucleotide is C. Likewise the +4 transcript nucleotide is A and the +4 template nucleotide is T.

Table 2

Position in transcript	+1+2+3+4+5+6
Transcript sequence	pppGGGAGA
Promoter top strand	5'-TAATACGACTCACTATAGGGAGA...-3' SEQ ID NO. 1
Promoter bottom strand	3'-ATTATGCTGAGTGATATCCCTCT...-5' SEQ ID NO. 2
Position in template	+1+2+3+4+5+6

[00116] Unlike DNA polymerases, T7 RNAP initiates RNA synthesis in the absence of a primer. The first step in initiation is called *de novo* RNA synthesis, in which RNA polymerase recognizes a specific sequence on the DNA template, selects the first pair of nucleotide triphosphates complementary to template residues at positions +1 and +2, and catalyzes the formation of a phosphodiester bond to form a dinucleotide. The initiating nucleotides have lower affinities for the polymerase than those used during elongation. The K_d value is 2 mM for the first initiating NTP and 80 μ M for the second, whereas the K_d is approximately 5 μ M for NTPs during elongation (*J. Mol. Biol.* (2007) 370, 256–268). It has been found that *de novo* synthesis is the rate-limiting step during transcription. T7 RNAP exhibits a strong bias for GTP as the initiating nucleotide (*J. Biol. Chem.* 248: 2235-2244 (1973)). Among the 17 T7 promoters in the genome, 15 initiate with GTP (and 13 with pppGpG), whereas there is no obvious NTP preference during transcription elongation (*J. Mol. Biol.* 370:256-268 (2007)). T7 RNA polymerase initiates poorly on promoters encoding A at position +1; transcription instead initiates predominantly with an encoded G at position +2 (*J. Biol. Chem.* 278:2819-2823 (2003)).

[00117] During *de novo* RNA synthesis, binding of the initiating nucleotides is achieved primarily by the free energy created from base stacking, specific interactions between the polymerase residues, the guanine moieties of the initiating nucleotides and base complementarity interactions (*J. Mol. Biol.* 370:256-268 (2007)).

[00118] It is known that T7 RNAP can also initiate with short oligonucleotide primers. For example, it is known that 13 promoters in the T7 genome initiate with pppGpG (*J. Mol. Biol.* 370:256-268 (2007)). Several groups showed that T7 RNAP can initiate from dinucleotide primers (*Biochemistry* 24:5716-5723 (1985)). Axelrod *et al.* showed that an uncapped GpA dinucleotide could initiate from +1 and +2 template nucleotides that were 2'-deoxycytidine and 2'-deoxythymidine, respectively ("CT" template). Their reaction conditions were 200 micromolar (μ M) dimer and 100 μ M ATP, CTP, GTP and UTP. Their

reaction mixture also contained 100 μ M 3' dATP, 3' dCTP 3' dUTP or 50 μ M 3' dGTP. They observe only GpA initiated RNAs and not a mixture of GpA initiated RNAs and 5' triphosphate RNAs from GTP initiation. This is likely due to the reaction conditions employed. 100 μ M GTP is well below the 2 mM K_d of T7 polymerase for the first initiating guanosine (J. Mol. Biol. (2007) 370, 256–268). Since GTP competes for initiation with the initiating oligonucleotide, using a low GTP concentration favors GpA initiation but results in low transcription yield (maximum calculated yield estimated to be <150 μ g / mL of reaction). When initiating transcription on "CT" template with ApG, CpG, UpG or GpG, they observed formation of RNA transcripts with an additional untemplated 5' nucleotide (A, C, U or G, respectively).

[00119] Axelrod *et al.* also used uncapped GpG dinucleotide to initiate RNA synthesis on a promoter where template nucleotides +1 and +2 were 2'-deoxycytidines ("CC" template). They observed low fidelity of initiation and observed three different transcription products. They state, "An examination of the autoradiograph indicates that one member of each triplet resulted from initiation with GpG at the normal (+1) position, the second member of each triplet resulted from initiation with GpG at the abnormal (-1) position, and the third member of each triplet resulted from initiation with guanosine triphosphate at the normal position. Thus, GpG is a relatively weak initiator with the ϕ 10 promoter ("CC" template), as well as with the ϕ 1.1A promoter ("CT" template), and is unable to prevent normal initiation with guanosine triphosphate at the concentration that was used." They did not observe initiation with a GpA dinucleotide with ϕ 10 promoter ("CC" template). CpA, ApC and ApA did not serve as initiators on either of the above "TC" and "CC" templates, presumably because these cannot hybridize to the template nucleotides at positions +1 and +2. The method described in Axelrod *et al.* was designed for production of very small amounts of radioactive transcripts for sequencing and is not suitable for large scale production of useful pharmaceutical amounts of RNA. If larger concentrations ($\sim$ 5 mM) of initiating dimer and NTPs, including GTP, were used to increase the yield of RNA, the expected result is a low proportion of RNA that starts with initiating dimer since GTP competes efficiently with dimer for initiation from the +1 nucleotides at NTP concentrations closer to the K_d (2 mM), producing large proportion of RNA that starts with pppG.

[00120] Pitulle *et al.* showed that transcription with T7 RNAP could be initiated with uncapped oligonucleotides (2-mer to 6-mer) (Gene, 112:101-105 (1992)). These

oligonucleotides had either a 5'-OH or a 5' monophosphate. They also initiated transcription with an oligonucleotide of the structure Biotin-ApG. All of the oligonucleotides used in this study contained a 3' terminal G. Pitulle *et al.* also showed that 2'-O-methyl residues and deoxy residues could be included within the primer sequence to produce RNA transcripts with 2'-O-methylated or 2'-deoxy residues at or near 5'-end of RNA. It is clear in this publication, that the 3' terminal guanosine residue of the any initiating oligonucleotide paired with the +1 template nucleotide. This results in untemplated nucleotides appended to the 5' end of the transcribed RNA. Specifically, the authors state, "Also sequence variations in this segment are easily possible, since no base-pairing with the template DNA is required apart from the Y-terminal G." Thus none of the initiating oligonucleotide primers is completely complementary only to the template nucleotide "C" at position +1 and not to any nucleotides at the following positions (+2, +3 etc.). This is confirmed in subsequent methods papers by this group (Methods Mol Biol. 74: 99-110 (1997), Methods Mol. Biol. 252:9-17, (2004)). Their methods differ from the method described herein where all the nucleotides of the initiating capped oligonucleotide primer completely complementary to template nucleotides at positions +1 and following positions. Kleinedam *et al.* created 5' modified tRNA transcripts by initiation with dimers or trimers that were modified with 2'-deoxy or 2'-O-methyl sugars (*Nucleic Acids Research* 21:1097-1101 (1993). Again the authors state that the 3'-terminal guanosine of the primer initiates at the +1 template nucleotide "C".

[00121] Another study by Ishikawa *et al.* showed that capped initiating oligonucleotide trimers of the structure ${}^m{}^7\text{GpppApG}$, ${}^m{}^7\text{Gppp}{}^m{}^6\text{ApG}$, ${}^m{}^7\text{GpppA}_2'\text{OmepG}$ or ${}^m{}^7\text{Gppp}{}^m{}^6\text{A}_2'\text{OmepG}$ could initiate transcription on template with 2'-deoxycytidine residues at template positions +1 and +2 ("CC" template: *Nucleic Acids Symposium Series No. 53: 129* (2009)). The authors state, "The different result from the case of using ${}^m{}^7\text{G5'pppG}$ may be caused from base pairing between additional adenosine (N1) in ${}^m{}^7\text{G5'pppN1pG}$ and 2'-deoxythymidine in T7 promoter at -1 position." This method clearly differs from the method described in the present invention where the +1 and +2 nucleotides of the initiating capped oligonucleotide trimer pair with the +1 and +2 of the template nucleotides. Ishikawa *et al.* used 6 mM initiating oligonucleotide trimer, 0.9 mM GTP and 7.5 mM each of ATP, CTP and UTP. The authors used a greater than 6 fold excess of the capped initiating oligonucleotide primer, the most expensive nucleotide component of transcription reaction, over competing GTP to drive the transcription reaction toward capped RNA over pppRNA which increases the total cost of

synthesized RNA. On the other hand, a low concentration of GTP (0.9 mM) limits the total yield of the RNA in transcription reaction (theoretically to less than 1.4 mg/mL). On the contrary the method described herein does not require restricting the concentration of any NTP to achieve both an efficient RNA capping and a higher yield of RNA (2 to 6 mg/mL) and thus allowing a production of high quality mRNA at a commercially useful cost.

[00122] None of the publications discussed above directly measured RNA capping efficiency so the extent of capping in those studies is unknown

[00123] Importantly, in all studies described above, the +1 template nucleotide is 2'-deoxycytidine (*Biochemistry* 24:5716-5723 (1985), *Gene*, 112:101-105 (1992), *Methods Mol. Biol.* 74: 99-110 (1997), *Methods Mol. Biol.* 252:9-17, (2004), *Nucleic Acids Research* 21:1097-1101 (1993), *Nucleic Acids Symposium Series No. 53: 129 (2009)*).

[00124] In more than 20 years since the publication of the transcription initiation studies with oligonucleotide primers, there have been no published examples of transcription initiation with initiating oligonucleotide primers containing 5'- to 5' inverted cap structures until a publication of short report in *Nucleic Acids Symposium Series No. 53: 129 (2009)*.

[00125] The methods and compositions provided herein for preparation of 5'-capped RNA include, but are not limited to, mRNA, small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), small cajal body-specific RNA (scaRNA). These methods involve the use of a Cap containing oligonucleotide primers, nucleoside 5'-triphosphates (NTPs) and RNA polymerase for DNA-templated and promoter controlled synthesis of RNA. In certain aspects, the methods use an initiating capped oligonucleotide primer that provides utility in RNA synthesis, in particular synthesis of capped mRNAs. The initiating oligonucleotide primer has a structure that resembles Cap 0, Cap 1, Cap 2 or TMG-Cap of natural RNA molecules, which include 2'-O-methylated nucleoside units at, penultimate Cap 1 and next to penultimate Cap 2 5'-positions of RNA. The natural Cap 0 structure does not have 2'-O-methylated nucleoside units.

[00126] The methods and compositions for preparation of RNA including, but not limited to, mRNA, snRNA, snoRNA, scaRNA, transfer RNA (tRNA), ribosomal RNA (rRNA), and transfer-messenger RNA (tmRNA) that carry modifications at or near 5'-end of the molecule. These methods involve the use of initiating oligonucleotide primers with or without Cap, nucleoside 5'-triphosphates (NTPs) and RNA polymerase for DNA-templated and promoter

controlled synthesis of RNA. In certain aspects, the methods use a modified initiating oligonucleotide primer carrying structural modifications that provide utility in RNA synthesis; in particular synthesis of 5'-modified RNAs.

[00127] The initiating capped oligonucleotide primer has an open 3'-OH group that allows for initiation of RNA polymerase mediated synthesis of RNA on a DNA template by adding nucleotide units to the 3'-end of the primer. The initiating capped oligonucleotide primer is substantially complementary to template DNA sequence at the transcription initiation site (*i.e.*, the initiation site is located closer to 3'-terminus of a promoter sequence and may overlap with promoter sequence). In certain embodiments, the initiating capped oligonucleotide primer directs synthesis of RNA predominantly in one direction ("forward") starting from the 3'-end of the primer. In certain aspects and embodiments, the initiating capped oligonucleotide primer outcompetes any nucleoside 5'-triphosphate for initiation of RNA synthesis, thereby maximizing the production of the RNA that starts with initiating capped oligonucleotide primer and minimizing a production of RNA that starts with 5'-triphosphate-nucleoside (typically GTP).

[00128] The manufacture of mRNA by *in vitro* transcription utilizes highly active phage RNA polymerases (T3, T7, SP6 and others). RNA polymerase works under control of specific promoter which is incorporated in DNA plasmid construct in front of a template nucleotide sequence. Transcription process usually starts with purine nucleoside 5'-triphosphate (typically GTP) and continues until the RNA polymerase encounters a terminating sequence or completes the DNA template.

[00129] As discussed above, mCAP dinucleotide analogs, $^m\text{G}(5')\text{ppp}(5')\text{N}$, that contain Cap 0 have been used for initiation *in vitro* transcription (*e.g.*, RNA 1: 957-967 (1995)). The capped RNA molecules produced using these dinucleotide analogs contain Cap 0. However only about 50% of synthesized capped RNA molecules have the correct "forward" orientation of Cap 0. To convert RNA with Cap 0 to RNA with Cap 1 an additional enzymatic reaction must be preformed using (nucleoside-2'-O) methyltransferase. However this conversion may be not quantitative; it is not easy to control and it is difficult to separate the remaining Cap 0 RNAs from Cap 1 RNAs. In addition, the competition from NTPs (specifically GTP) for initiation transcription further reduces the quantity of active capped RNA molecules produced.

[00130] In addition, modified dinucleotide analogs, such as $^m\text{G}_{3'\text{Ome}}(5')\text{ppp}(5')\text{N}$ and

other related ARCA analogs, that carry modified 7^mG residue with blocked $3'$ and/or $2'$ position on ribose, have been used for initiation of *in vitro* transcription (e.g., RNA 7:1486-1495 (2001)). These ARCA cap analogs direct RNA synthesis only in the "forward" orientation and therefore produce a RNA molecule with (natural) Cap 0 on the $5'$ -terminus (having $2'$ and/or $3'$ modifications to 7^mG residue). Such RNAs are more active in translation systems compared to RNAs prepared using standard dinucleotide analogs, $7^m\text{G}(5')\text{ppp}(5')\text{N}$. To convert RNA with ARCA Cap 0 to RNA with ARCA Cap 1 an additional enzymatic reaction must be performed with (nucleoside- $2'$ -O) methyltransferase similar to that required for the dinucleotide analogs previously discussed. This method has the same disadvantages as that elaborated for the mCAP dinucleotide analogs; the conversion of RNA with Cap 0 to RNA with ARCA Cap 1 may be not quantitative; the reaction is not easy to control, it is difficult to separate remaining Cap 0 RNAs from Cap 1 RNAs and competition from NTPs (specifically GTP) for initiation of transcription further reduces the quantity of active capped RNA molecules produced.

[00131] Short oligonucleotide primers (2 to 6-mer) with $3'$ -terminal guanosine residue have been used for initiation of *in vitro* transcription (Pitulle, C. *et al.*, *Gene*, 112:101-105(1992)). These oligonucleotide primers contained modified and unmodified ribonucleoside residues (e.g., modified ribonucleoside residues included $2'$ -O-methylated nucleoside residues and $2'$ -deoxyribonucleoside residues). The shorter oligonucleotide primers (dimers to tetramer) substantially out-compete GTP for initiation of transcription while longer primers (pentamer to hexamer) are much less efficient in initiation of transcription compared to GTP. It may be because these longer primers (as they are designed) have a low percent of complementarity with DNA template at initiation site. In contrast, dimer, AG (as designed), was complementary to the DNA template at initiation site. The RNA molecules produced using oligonucleotide primers, discussed in this section, had internal $2'$ -O-methylated nucleoside but did not contain $5'$ -Cap 0, Cap 1, Cap 2 or TMG-cap. To convert RNA without cap structure to RNA with Cap 1, Cap 2 or TMG-cap structure an additional enzymatic reactions using capping enzymes would have to be performed. However such conversion has the same disadvantages as those stated above.

[00132] Other short RNA oligonucleotides containing cap structures and internal $2'$ -O-methylated nucleoside residues have been chemically prepared (Ohkubo *et al.*, *Org. Letters* 15:4386-4389 (2013)). These short capped oligonucleotides were ligated with a

“decapitated” (without 5'-cap structure) fragment of long RNA using T4 DNA ligase and a complementary DNA splinter oligonucleotide. The final RNA synthesized using this chemical-enzymatic method had both internal 2'-O-methylated nucleoside residues and 5'-TMG-cap structure. However only short capped RNAs (<200-mer) were prepared using this ligation approach. Moreover, the yields were low (15-30%). It is not easy to control and optimize the T4 DNA ligation reaction and it requires a laborious separation process using PolyAcrylamide Gel Electrophoresis and isolation of capped RNAs from remaining uncapped RNAs. Separation of long (500 -10000 bases) capped mRNAs from remaining uncapped mRNAs by PAGE method is not feasible.

[00133] Finally, 5'-modified nucleoside or 5'-modified mononucleotide or 5'-modified dinucleotide, typically a derivative of guanosine, have been used for initiating *in vitro* transcription of RNA (*Gene*, 112:101-105 (1992) and *Bioconjug. Chem.*, 10371-378 (1999)). These initiator nucleosides and nucleotides may carry labels or affinity groups (e.g. biotin) and, when incorporated on the 5'-end of RNA, would allow for easy detection, isolation and purification of synthesized RNA. This 5'-labeled or tagged RNAs may be necessary for some applications. However this strategy was not used for the preparation of mRNA with Cap 0, Cap 1, Cap 2 or TMG-cap structures.

[00134] In certain aspects of the present invention, compositions of the initiating capped oligonucleotide primers of Formula 1 are provided. In related aspects, are methods in which RNA is synthesized using the initiating capped oligonucleotide primers of Formulas 1.

Initiating Capped Oligonucleotide Primer

[00135] The initiating capped oligonucleotide primers of the present invention have a hybridization sequence which may be complementary to a sequence on DNA template at initiation site. The length of the hybridization sequence of the primers for use in the methods and compositions provided herein depends on several factors including the identity of template nucleotide sequence and the temperature at which this primer is hybridized to DNA template or used during *in vitro* transcription. Determination of the desired length of a specific nucleotide sequence of an initiating capped oligonucleotide primer for use in transcription can be easily determined by a person of ordinary skill in the art or by routine experimentation. For example, the length of a nucleic acid or oligonucleotide may be determined based on a desired hybridization specificity or selectivity.

[00136] In some embodiments, the nucleotide length of initiating capped oligonucleotide primer (including the inverted 5'-5' Cap nucleotide) is between 3 to about 9, in some embodiments the nucleotide length of initiating capped oligonucleotide primer (including Cap) is between 3 to about 7, in some embodiments the nucleotide length of initiating capped oligonucleotide primer (including Cap) is between 3 to about 5, and in some embodiments the nucleotide length of initiating capped oligonucleotide primer (including Cap) is about 3. The length of hybridization sequence within the initiating capped oligonucleotide primer may be equal to or shorter than the total length of initiating capped oligonucleotide primer.

[00137] The presence of hybridization sequence forces an initiating capped oligonucleotide primer to predominantly align with complementary sequence of the DNA template at the initiation site in only the desired orientation (*i.e.*, the "forward" orientation). In the forward orientation, the RNA transcript begins with the inverted guanosine residue (*i.e.*, $^7mG(5')ppp(5')N...$). The dominance of the forward orientation of the primer alignment on DNA template (Figure 1) over incorrect "reverse" orientation is maintained by the thermodynamics of the hybridization complex. The latter is determined by the length of the hybridization sequence of initiating capped oligonucleotide primer and the identity of bases involved in hybridization with DNA template. Hybridization in the desired forward orientation may also depend on the temperature and reaction conditions at which DNA template and initiating capped oligonucleotide primer are hybridized or used during *in vitro* transcription.

[00138] The initiating capped oligonucleotide primer of the present invention enhances efficacy of initiation of transcription compared to efficacy of initiation with standard GTP, ATP, CTP or UTP. In some embodiments, initiation of transcription is considered enhanced when synthesis of RNA starts predominantly from initiating capped oligonucleotide primer and not from any NTP in transcription mixture. The enhanced efficiency of initiation of transcription results in a higher yield of RNA transcript. The enhanced efficiency of initiation of transcription may be increased to about 10%, about 20%, about 40%, about 60%, about 80%, about 90%, about 100%, about 150%, about 200% or about 500% over synthesis of RNA with conventional methods without initiating capped primer. In certain embodiments "initiating capped oligonucleotide primers" out-compete any NTP (including GTP) for initiation of transcription. One of ordinary skill in the art is able to readily determine the level of substrate activity and efficacy of initiating capped oligonucleotide primers. One example

of a method of determining substrate efficacy is illustrated in Example 13). In certain embodiments, initiation takes place from the capped oligonucleotide primer rather than an NTP, which results in a higher level of capping of the transcribed mRNA.

[00139] In some aspects, methods are provided in which RNA is synthesized utilizing an initiating capped oligonucleotide primer that has substitutions or modifications. In some aspects, the substitutions and modifications of the initiating capped oligonucleotide primer do not substantially impair the synthesis of RNA. Routine test syntheses can be preformed to determine if desirable synthesis results can be obtained with the modified initiating capped oligonucleotide primers. Those skilled in the art can perform such routine experimentation to determine if desirable results can be obtained. The substitution or modification of initiating capped oligonucleotide primer include for example, one or more modified nucleoside bases, one or more modified sugars, one or more modified internucleotide linkage and/or one or more modified triphosphate bridges.

[00140] The modified initiating capped oligonucleotide primer, which may include one or more modification groups of the methods and compositions provided herein, can be elongated by RNA polymerase on DNA template by incorporation of NTP onto open 3'-OH group. The initiating capped oligonucleotide primer may include natural RNA and DNA nucleosides, modified nucleosides or nucleoside analogs. The initiating capped oligonucleotide primer may contain natural internucleotide phosphodiester linkages or modifications thereof, or combination thereof.

[00141] In one embodiment the modification group may be a thermally labile group which dissociates from a modified initiating capped oligonucleotide primer at an increasing rate as the temperature of the enzyme reaction medium is raised. Examples of thermally labile groups for oligonucleotides and NTPs are described in *Nucleic Acids Res.*, 36:e131 (2008), *Collect. Symp. Ser.*, 10:259-263 (2008) and *Analytical Chemistry*, 81:4955-4962 (2009).

[00142] In some aspects, methods are provided in which RNA is synthesized where at least one or more NTP is added to a transcription reaction may have a modification as disclosed herein. In some aspects, the modification of the at least one NTP does not substantially impair RNA polymerase mediated synthesis of RNA. The modification of NTP may include for example, one or more modified nucleoside bases, one or more modified sugars, one or more modified 5'-triphosphate. The modified NTP may incorporate onto the 3'-end of the initiating capped oligonucleotide primer and it does not block transcription and

supports further elongation of the primer.

[00143] In another embodiment, the modification group of an initiating capped oligonucleotide primer may be a detectable label or detectable marker. Thus, following transcription, the target RNA, containing the detectable label or marker, can be identified by size, mass, color and/or affinity capture. In some embodiments, the detectable label or marker is a fluorescent dye; and the affinity capture label is biotin. In certain embodiments, one or more components of a transcription reaction (initiating capped oligonucleotide primer and/or NTPs) may be labeled with a detectable label or marker. Thus, following transcription, the RNA molecule can be identified, for example, by size, mass, affinity capture or color. In some embodiments, the detectable label is a fluorescent dye; and the affinity capture label is biotin.

[00144] Standard chemical and enzymatic synthesis methods may be utilized to synthesize the “initiating capped oligonucleotide primers” of the present invention and are disclosed herein in the Examples section.

Kits

[00145] Kits including, the “initiating capped oligonucleotide primer” for performing transcription are also contemplated. For example, kits may contain all transcription reagents for synthesis of common RNAs (*e.g.*, FLuc mRNA). More specifically, a kit may contain: an “initiating capped oligonucleotide primer”; a container marked for transcription; instructions for performing RNA synthesis; and one or more reagents selected from the group consisting of one or more modified or unmodified initiating capped oligonucleotide primers, one or more unmodified NTPs, one or more modified NTPs (*e.g.*, pseudouridine 5'-triphosphate), an RNA polymerase, other enzymes, a reaction buffer, magnesium and a DNA template.

[00146] The initiating capped oligonucleotide primers of the present invention have a significant advantage over current methods and compositions involving use of various initiating nucleosides, nucleotides and oligonucleotides or use of polyphosphate dinucleotide derivatives containing Cap 0 structure, such as mCAP and ARCA. The initiating capped oligonucleotide primers are compatible with existing transcription systems and reagents and no additional enzymes or reagents are required. In addition, the use of initiating capped oligonucleotide primers makes several non-enzymatic and enzymatic steps (such as capping

and 2'-O-methylation) unnecessary thus reducing complexity of the process and a cost of RNA synthesis.

[00147] While the exemplary methods described herein relate to T7 RNA polymerase mediated transcription reaction, a number of other RNA polymerases known in the art for use in transcription reactions may be utilized with the compositions and methods of the present invention. Other enzymes, including natural or mutated variants that may be utilized include, for example, SP6 and T3 RNA polymerases and RNA polymerases from other sources including thermostable RNA polymerases.

[00148] Some nucleic acid replication and amplification methods may include transcription as a part of the process. Among these methods are: transcription mediated amplification (TMA) and nucleic acid sequence-based amplification (NASBA), DNA and RNA sequencing and other nucleic acid extension reactions known in the art. The skilled artisan will understand that other methods may be used either in place of, or together with, transcription methods, including variants of transcription reactions developed in the future.

Therapeutic Uses

[00149] The present invention also contemplates the production of mRNAs containing the initiation capped oligonucleotide primer for use as therapeutic agents in a pharmaceutical composition, the introduction of RNAs containing the initiating capped oligonucleotide primer into cells to treat a medical condition of the cells or the introduction of RNAs containing the initiating capped oligonucleotide primer into cells that utilize those RNAs to produce proteins that may have a therapeutic affect on the host cells.

[00150] One method for treating a condition utilizing an RNA containing an initiating capped oligonucleotide primer comprises the step of administering the RNA containing the initiating capped oligonucleotide primer of formula I or a composition comprising such RNA to a subject having, or suspected of having a condition whose symptoms/symptomologies may be reduced in severity or eliminated.

[00151] An RNA containing the initiating capped oligonucleotide primer of formula I "compound", when formulated in a pharmaceutically acceptable carrier and/or pharmaceutically acceptable salt at a concentration of 4 mg/ml or less, is effective to produce a reduction of the symptoms and/or symptomologies by at least 20%, 30%, 40%, 50%, 60%,

70%, 80%, 90% or greater than an untreated individual with the pharmaceutically acceptable carrier alone.

[00152] Pharmaceutical compositions may be formulated for administration by injection or other appropriate routes known to those skilled in the art for treating a particular condition. An injectable composition for parenteral administration typically contains the active compound in a suitable solution and/or pharmaceutical carrier such as sterile physiological saline. The composition may also be formulated as a suspension in a lipid or phospholipid, in a liposomal suspension, or in an aqueous emulsion.

[00153] Methods for preparing a variety of compositions and/or formulations are known to those skilled in the art *see* Remington's Pharmaceutical Sciences (19th Ed., Williams & Wilkins, 1995). The composition to be administered will contain a quantity of the selected compound in a pharmaceutically safe and effective amount for increasing expression of the desired protein in the target cells or tissue.

[00154] In some embodiments, the pharmaceutical composition contains at least 0.1% (w/v) of the compound, as described above, in some embodiments, the pharmaceutical composition contains greater than 0.1%, in some embodiments, the pharmaceutical composition contains up to about 10%, in some embodiments, the pharmaceutical composition contains up to about 5%, and in some embodiments, the pharmaceutical composition contains up to about 1% (w/v) of the compound. Choice of a suitable concentration depends on factors such as the desired dose, frequency and method of delivery of the active agent.

[00155] For treatment of a subject, such as a mammal or a human, dosages are determined based on factors such as the weight and overall health of the subject, the condition treated, severity of symptoms, *etc.* Dosages and concentrations are determined to produce the desired benefit while avoiding any undesirable side effects. Typical dosages of the subject compounds are in the range of about 0.0005 to 500 mg/day for a human patient, and ranging in some embodiments between about 1-100 mg/day. For example, higher dose regimens include *e.g.* 50-100, 75-100, or 50-75 mg/day, and lower dose regimens include *e.g.* 1-50, 25-50, or 1-25 mg/day.

[00156] Various aspects of the present invention are illustrated by the following non-limiting examples. The examples are for illustrative purposes and are not a limitation on any

practice of the present invention. It will be understood that variations and modifications can be made without departing from the spirit and scope of the invention. One of ordinary skill in the art readily knows how to synthesize or commercially obtain the reagents and components described herein.

EXAMPLE 1

Preparation of 7-methylguanosine 5'-diphosphate (pp^{7m}G) from guanosine 5'-diphosphate (Figure 1)

[00157] To a stirring solution of guanosine 5'-diphosphate (2.5 mmol) in 40.0 mL of water, acetic acid is added to adjust the pH of the solution to 4.0. To this mixture dimethyl sulfate (4.0 mL) is added dropwise over a period of 30 minutes and the reaction mixture is stirred at room temperature for 4 hours while maintaining the pH of the reaction mixture at about 4.0 using 0.1M NaOH solution. After 4 hours, the reaction mixture is extracted with CH₂Cl₂ (3 x 50 mL) to remove unreacted dimethyl sulfate. The aqueous layer is diluted with water to 500 mL, adjusted to pH 6.5 with 1M TEAB and loaded on a DEAE Sephadex column (3 x 50 cm). The product is eluted using a linear gradient of 0–1 M TEAB, pH 7.5 (3 L). Fractions containing pure pp^{7m}G (triethylammonium salt) are pooled, evaporated, and dried under high vacuum to give a fine white powder (yield: 80%). A similar procedure is disclosed in *Bioorgan. Med. Chem. Letters* 17:5295-5299 (2007).

EXAMPLE 2

Preparation of 7-methylguanosine 5'-diphosphate imidazolide (Im-pp^{7m}G) from pp^{7m}G (Figure 2)

[00158] The triethylammonium salt of pp^{7m}G (0.4 mmol) is reacted with imidazole (4 mmol), triphenylphosphine (2 mmol) and 2, 2'-dipyridyl disulfide (2 mmol) in dry DMF (20 mL) for 8 hrs. Crude Im-pp^{7m}G is precipitated by pouring the reaction mixture into 250 mL of a 0.2 M sodium perchlorate solution. The mixture is cooled to -20°C and the resulting precipitate is collected by centrifugation, washed with acetone (3 x 50 mL) and dried under high vacuum. The isolated yield of Im-pp^{7m}G is approximately 100%. A similar procedure is disclosed in *Nucleosides, Nucleotides, and Nucleic Acids* 24:1131-1134 (2005) and *J. Org. Chem.* 64:5836-5840 (1999).

EXAMPLE 3

Preparation of 3'-O-Methylguanosine 5'-phosphate (pG_(3'Ome)) from 3'-O-methylguanosine

(Figure 3)

[00159] 3'-O-methylguanosine (10 mmol) is dissolved in triethylphosphate (40 mL) at 60-70°C. The mixture is cooled to 0°C in an ice-water bath, phosphorus oxychloride (30 mmol) is added and the mixture is stirred under argon for 3 hrs at room temperature. The reaction is quenched by slow addition of 1M TEAB (100 mL; pH 8.5) with stirring. The mixture is stirred for 8 hrs and diluted with 1L of water. The resulting solution is loaded onto a DEAE Sephadex column (3 x 40 cm) and the product is eluted with a linear gradient 0.05 to 1.0M (3L) of TEAB (pH 7.5). Fractions containing pure product are combined, evaporated to a solid residue and co-evaporated with methanol (4 x 50 mL) to give pG_(3'Ome) (triethylammonium salt) as a white solid (yield 60%). A similar procedure is disclosed in U.S. Patent Application serial no. 2012/0156751.

EXAMPLE 4

Preparation of 3'-O-Methylguanosine 5'-phosphorimidazolid (Im-pG_(3'Ome)) from pG_(3'Ome)

(Figure 4)

[00160] The triethylammonium salt of pG_(3'Ome) (0.5 mmol) is reacted with imidazole (5 mmol), triphenylphosphine (2.5 mmol) and 2,2'-dipyridyl disulfide (2.5 mmol) in dry DMF (25 mL) for 5 hrs. Crude Im-pG_(3'Ome) is precipitated by pouring the reaction mixture into 400 mL of 0.2 M sodium perchlorate in acetone solution. The mixture is cooled to -20°C and the resulting precipitate is collected by centrifugation, washed with acetone (3 x 60 mL) and dried under high vacuum (yield: 100%). A similar procedure is disclosed in U.S. Patent Application serial no. 2012/0156751.

EXAMPLE 5

Preparation of 3'-O-Methylguanosine 5'-diphosphate (ppG_(3'Ome)) from Im-pG_(3'Ome)

(Figure 5)

[00161] Solid zinc chloride (14.0 mmol) is added with small portions to a solution of Im-pG_(3'Ome) (7.0 mmol) in dry DMF (40 mL). The mixture is stirred for 15 minutes under

argon until all solids are dissolved. A solution of 1M tributylammonium phosphate in DMF (40 mL) is added and the mixture is stirred at room temperature. After 5 hours the mixture is diluted with 200 mL of water and extracted with dichloromethane (2 x 200 mL). The aqueous layer is diluted with water (1L), loaded onto a DEAE Sephadex column (5 x 40 cm) and eluted with linear gradient of 0.05 to 1.0M (6L) TEAB (pH 7.5). Fractions containing pure product are combined, evaporated and co-evaporated with methanol (4 x 50 mL) to give ppG(3'-Ome) (triethylammonium salt) as a white solid (yield: 60%). A similar procedure is disclosed in U.S. Patent Application serial no. 2012/0156751.

EXAMPLE 6

Preparation of 7-methyl-3'-O-methylguanosine 5-diphosphate (pp^{7m}G(3'-Ome)) from ppG(3'-Ome)

(Figure 6)

[00162] A solution of ppG(3'-Ome) (triethylammonium salt; 3.0 mmol) in 50 mL of water is prepared and glacial acetic acid is added to adjust the pH of the solution to 4.0. Dimethyl sulfate (10.0 mL) is added dropwise to this mixture over a period of 30 minutes and the reaction mixture is stirred at room temperature for 4 hours while maintaining a pH of 4.0±0.5 using 0.1M NaOH solution. After 4 hours, the reaction mixture is extracted with CH₂Cl₂ (3 x 150 mL) to remove unreacted dimethyl sulfate. The aqueous layer is adjusted to pH 5.5, diluted with water (500 mL) and loaded onto a DEAE Sephadex column (3 x 50 cm). The product is eluted using a linear gradient of 0–1.0M TEAB, pH 7.5 (3 L). Fractions containing pure pp^{7m}G(3'-Ome) (triethylammonium salt) are pooled, evaporated, and dried under high vacuum to give a fine white powder (yield: 80%). A similar procedure is disclosed in *RNA* 9:1108-1122(2003); *Nucleoside Nucleotides & Nucleic acids* 25:337-340 (2006); and U.S. Patent Application Serial no.2012/0156751.

EXAMPLE 7

Preparation of 7-methyl-3'-O-methylguanosine 5-diphosphate imidazolide (Im-pp^{7m}G(3'-Ome)) from pp^{7m}G(3'-Ome)

(Figure 7)

[00163] The protocol described in Example 1 is utilized for preparation of Im-pp^{7m}G(3'-Ome). A similar procedure is disclosed in *RNA* 14:1119-1131 (2008).

EXAMPLE 8

General procedure for preparation of pN(2'-OR₁)pN dinucleotides (R₁ = H or Me) (Figure 8)

[00164] Phosphoramidite monomer (i) (1.0 mmol) and 2',3', N-protected nucleoside (ii) (1.0 mmol) are reacted in 10 mL of acetonitrile containing 2.5 molar equivalents of activator (tetrazole). After 60 minutes of stirring at room temperature the intermediate product is oxidized from the P(III) to P(V) state with iodine and extracted with dichloromethane (200 mL) and brine (200 mL). The organic layer is dried with sodium sulfate and is evaporated to solid foam (intermediate (iii)).

[00165] To remove the DMT-protecting group, intermediate (iii) is dissolved in 10 mL of 80% acetic acid and, after reaction is completed (about 1-2 hrs), the mixture is evaporated and co-evaporated with methanol (5 x 30 mL) to remove acetic acid. The crude 5'-OH dimer (iv) is isolated and purified by silica gel chromatography using 5% methanol in dichloromethane as an eluent.

[00166] The 5'-OH dimer (iv) (1.0 mmol) is phosphitylated with 2 equivalents of bis-cyanoethyl-N,N-diisopropyl-phosphoramidite and 2 equivalents of activator (tetrazole) in 10 mL of acetonitrile. After 30 minutes of stirring at room temperature the 5'-phosphitylated dimer is oxidized from the P(III) to P(V) state with iodine and extracted with dichloromethane (150 mL) and brine (150 mL). The organic layer is evaporated to an oily residue, co-evaporated with methanol (2 x 30 mL), dissolved in 12 mL of methanol and concentrated ammonia (12 mL) was added. The mixture is kept at room temperature for over 48 hours until deprotection of the pN₂'OR₁pN dimer (v) is complete. The mixture is evaporated and co-evaporated with methanol (2 x 30 mL).

[00167] When R = methyl the crude dimer (v) is directly purified by anion exchange and reverse phase chromatography (Step 5). Fractions are evaporated to give a final pN₂'OR₁pN dimer (v) (v; R₁ = methyl; triethylammonium salt) as a white solid (35% overall yield).

[00168] When R = TBDMS the crude dimer (v) is treated (Step 4B) with HF-3TEA mixture to remove 2'-OTBDMS protecting group (*Org. Biomol. Chem.* 3:3851-3868 (2005))

and *Nucl. Acids Res.* 22:2430-2431 (1994)). When the reaction is complete, the mixture is diluted with 0.05M TEAB and purified by anion exchange and reverse phase chromatography (Step 5). Fractions are evaporated to give a final $pN_2'OR_1pN$ dimer (v) ($R_1 = H$; triethylammonium salt) as a white solid (30% overall yield).

EXAMPLE 9

General procedure for the synthesis of initiating oligonucleotides with Cap 0, Cap 1 or Cap 2 structures (5'-phosphorylated dinucleotide is used in example)

(Figure 9)

A. Approach 1:

[00169] To a suspension of Im-pp^{7m}G(3'-Ome) or Im-pp^{7m}G (2 mmol; sodium salt form) and 5'-phosphorylated dinucleotide (1 mmol; triethylammonium salt) in DMF (50 mL), anhydrous ZnCl₂ (1 g) is added slowly while the mixture is stirred at 35°C. After 24 hours the reaction is stopped by addition of a 25 mM solution of EDTA in water (500 mL) and neutralized by 1M solution of sodium bicarbonate. The mixture is diluted to 1L with water and loaded on a DEAE Sephadex column (3 x 50 cm). The product is eluted using a linear gradient of 0–1 M ammonium bicarbonate, pH 7.2 (2 L). Fractions containing pure product are pooled, evaporated, and dried under high vacuum to give a fine white powder (yield: 60%). A similar approach is disclosed in U.S. patent application serial no. 2012/0156751; *Bioorg. Med. Chem. Lett.* 17:5295-5299 (2007) and *RNA* 14:1119-1131 (2008).

B. Approach 2:

[00170] Im-pp^{7m}G(3'-Ome) or Im-pp^{7m}G (2 mmol) is dissolved in N-methyl morpholine buffer (0.2 M, pH 7.0, 10 mL) containing MnCl₂ (2 mmol) and added to solid 5'-phosphorylated dinucleotide (1 mmol; triethylammonium salt). The reaction is stirred at room temperature. After 24-40 hours the reaction is stopped with 10 mL of 0.25M solution of EDTA. The mixture is loaded onto a DEAE Sephadex column (3 x 50 cm). The product is eluted using a linear gradient of 0–1.0M ammonium bicarbonate, pH 7.2 (2 L). Fractions containing pure product are pooled, evaporated, and dried under high vacuum to give a fine white powder (yield 50%). A similar approach is disclosed in *Bioorganic & Medicinal Chemistry* 21:7921-7928 (2013), *Nucleic Acids Research* 37:1925-1935 (2009); *J. Org. Chem.* 64:5836-5840 (1999). (Note: 1. A 5'-phosphorylated trimer, tetramer, pentamer, hexamer, heptamer, octamer, nanomer or decamer oligonucleotide can be used instead of 5'-phosphorylated dinucleotide as illustrated in Example 9. 2. In order to prepare the initiating

capped oligonucleotide with Cap 0, the 5'-phosphorylated oligonucleotide does not have 2'-O-methyl groups on the first 5'-nucleoside residues, e.g. pApG. 3. In order to prepare the initiating capped oligonucleotide with Cap 1, the 5'-phosphorylated oligonucleotide carries one 2'-O-methyl group on the first 5'-nucleoside residue, e.g. pA(2'-Ome)pG. 4. In order to prepare the initiating capped oligonucleotide with Cap 2, the 5'-phosphorylated oligonucleotide carries two 2'-O-methyl groups on the first and second 5'- nucleoside residues, e.g. pA(2'-Ome)pG(2'-Ome)pG).

EXAMPLE 10

Structures of Initiating Capped Oligonucleotide Primers According to Formula I

Figure 10 shows the structures of the initiating capped oligonucleotide primers used in the examples according to Structure I. A) $m^7G_{3'Ome}pppG$, B) $m^7GpppG_{2'Ome}pG$, C) $m^7G_{3'Ome}pppG_{2'Ome}pG$, D) $m^7GpppA_{2'Ome}pG$, E) $m^7G_{3'Ome}pppA_{2'Ome}pG$, F) $m^7GpppC_{2'Ome}pG$, G) $m^7G_{3'Ome}pppC_{2'Ome}pG$, H) $m^7GpppA_{2'Ome}pG_{2'Ome}pG$.

EXAMPLE 11

In vitro Transcription with ARCA Primer

[00171] A double stranded DNA transcription template encoding firefly luciferase was generated by polymerase chain reaction. The +1 template nucleotide was 2'-deoxycytidine. Transcription reactions were assembled with 25 ug/mL transcription template, 40 mM Tris-HCl (pH 8.0), 27 mM MgCl₂, 2 mM spermidine, 10 mM DTT, 0.002% Triton X-100, 1000 unit/mL murine RNase inhibitor (New England Biolabs catalog #M0314), 2 unit/mL inorganic pyrophosphatase (New England Biolabs catalog # M2403), 4000 units/mL T7 RNA polymerase (New England Biolabs catalog #M0251), 6 mM ARCA ($m^7G_{3'Ome}pppGG$), 1.5 mM GTP, 7.5 mM ATP, 7.5 mM CTP and 7.5 mM UTP. Transcriptions were incubated at 37°C for 2 hours. Reactions were supplemented with 10 mM Tris-HCl (pH 7.6), 2.5 mM MgCl₂, 0.5 mM CaCl₂ and 100 units/mL DNase I (New England Biolabs catalog #M0303) and incubated for 1 hour at 37°C. The resulting mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs were eluted in water and dephosphorylated by adjusting the solution to 50 mM Bis-Tris-Propane HCl (pH 6.0), 1 mM MgCl₂, 0.1 mM ZnCl₂ and 250 units/mg Antarctic phosphatase (New England Biolabs catalog #M0289). The reaction were incubated at 37 °C for 1 hour. The

resulting mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs were eluted in water.

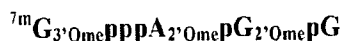
EXAMPLE 12

In vitro transcription with Trimeric Initiating Capped Oligonucleotide Primers

[00172] Some initiating capped oligonucleotide primers that are utilized in transcriptions are indicated in Table 1. A double stranded firefly luciferase DNA transcription template specific for each trimer is generated by polymerase chain reaction. These templates differed in template nucleotides +1 and +2. For a given trimer, template nucleotide +1 was complementary to B₁ (the nucleotide destined to become transcript nucleotide +1). Likewise template nucleotide +2 was complementary to B₁₀ (the nucleotide destined to become transcript nucleotide +2; see Example 10)). Transcription reactions are assembled with 25 ug/mL transcription template, 40 mM Tris-HCl (pH 8.0), 27 mM MgCl₂, 2 mM spermidine, 10 mM DTT, 0.002% Triton X-100, 1000 unit/mL murine RNase inhibitor (New England Biolabs catalog #M0314), 2 unit/mL inorganic pyrophosphatase (New England Biolabs catalog # M2403), 4000 units/mL T7 RNA polymerase (New England Biolabs catalog #M0251) and 6 mM of initiating capped oligonucleotide primer, 1.5 mM GTP and 7.5 mM each of ATP, CTP and UTP. In subsequent examples, this primer/NTP formulation will be referred to as Primer/NTP Formulation 1. It is clear to one skilled in the art that other polymerases such as T7, T3 or SP6 RNA polymerases could be used instead of T7 RNA polymerase to perform the same function by using their respective promoters. Transcription reaction mixtures are incubated at 37°C for 2 hours. Reactions are supplemented with 10 mM Tris-HCl (pH 7.6), 2.5 mM MgCl₂, 0.5 mM CaCl₂ and 100 units/mL DNase I (New England Biolabs catalog #M0303) and incubated for 1 hour at 37 °C. The resulting mRNAs are purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs are eluted in water and dephosphorylated by adjusting the solution to 50 mM Bis-Tris-Propane HCl (pH 6.0), 1 mM MgCl₂, 0.1 mM ZnCl₂ and 250 units/mg Antarctic phosphatase (New England Biolabs catalog #M0289). The reaction is incubated at 37 °C for 1 hour. The resulting mRNAs are purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs are eluted in water.

EXAMPLE 13

In vitro transcription with Tetrameric Initiating Capped Oligonucleotide Primer



[00173] Initiating capped oligonucleotide primer $^7\text{mG}_{3'}\text{OMepppA}_{2'}\text{OMepG}_{2'}\text{OMepG}$ were used in transcription. Transcription were carried out as in Example 11 with the following modifications. A double stranded DNA transcription template specific for tetramer was generated by polymerase chain reaction. The template nucleotide +1 (2'-deoxythymidine) was complementary to Adenosine (the first nucleotide of the primer destined to become a transcript nucleotide +1). Likewise template nucleotide +2 (2'-deoxycytidine) was complementary to guanosine (the second nucleotide of the primer destined to become transcript nucleotide +2). Likewise template nucleotide +3 (2'-deoxycytidine) was complementary to guanosine (the third nucleotide of the primer destined to become transcript nucleotide +3). Transcription reactions were assembled with 25 ug/mL transcription template, 40 mM Tris-HCl (pH 8.0), 27 mM MgCl₂, 2 mM spermidine, 10 mM DTT, 0.002% Triton X-100, 1000 unit/mL murine RNase inhibitor (New England Biolabs catalog #M0314), 2 unit/mL inorganic pyrophosphatase (New England Biolabs catalog # M2403), 4000 units/mL T7 RNA polymerase (New England Biolabs catalog #M0251), 6 mM of initiating capped oligonucleotide primer $^7\text{mG}_{3'}\text{OMepppA}_{2'}\text{OMepG}_{2'}\text{OMepG}$, 1.5 mM of GTP, 7.5 mM of ATP, 7.5 mM of UTP and 7.5 mM of CTP. It is clear to one skilled in the art that other polymerases such as T7, T3 or SP6 RNA polymerases could be used instead of T7 RNA polymerase to perform the same function by using their respective promoters. Transcription reaction mixtures were incubated at 37°C for 2 hours. Reactions were supplemented with 10 mM Tris-HCl (pH 7.6), 2.5 mM MgCl₂, 0.5 mM CaCl₂ and 100 units/mL DNase I (New England Biolabs catalog #M0303) and incubated for 1 hour at 37 °C. The resulting mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs were eluted in water and dephosphorylated by adjusting the solution to 50 mM Bis-Tris-Propane HCl (pH 6.0), 1 mM MgCl₂, 0.1 mM ZnCl₂ and 250 units/mg Antarctic phosphatase (New England Biolabs catalog #M0289). The reaction was incubated at 37 °C for 1 hour. The resulting mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs were eluted in water.

EXAMPLE 14***In vitro* Transcription with ^{m7}GpppA_{2'}OmeP Initiating Capped Oligonucleotide Primers on a Transcription Template With 2'-deoxythymidine and 2'-deoxycytidine Residues at Template Positions +1 and +2, respectively.**

[00174] A double stranded firefly luciferase DNA transcription template was used with 2'-deoxythymidine and 2'-deoxycytidine residues at template positions +1 and +2, respectively. Two transcription reactions were assembled with 25 ug/mL transcription template, 40 mM Tris-HCl (pH 8.0), 27 mM MgCl₂, 2 mM spermidine, 10 mM DTT, 0.002% Triton X-100, 1000 unit/mL murine RNase inhibitor (New England Biolabs catalog #M0314), 2 unit/mL inorganic pyrophosphatase (New England Biolabs catalog # M2403), 4000 units/mL T7 RNA polymerase (New England Biolabs catalog #M0251). The two transcriptions differed in the amount of initiating capped oligonucleotide, amount of NTPs and the identity of NTPs. The first transcription was designed to mimic the transcription conditions of Ishicawa *et al.* (Nucleic Acids Symposium Series No. 53:129 (2009)). This transcription reaction contained 6 mM of ^{m7}GpppA_{2'}OmeP initiating capped oligonucleotide primer, 0.9 mM GTP and 7.5 mM of ATP, CTP and UTP. In subsequent examples this primer/NTP formulation will be referred to as Primer/NTP Formulation 2. In this formulation, the initiating oligonucleotide primer was in greater than 6 fold excess over GTP. The second transcription used 5 mM of ^{m7}GpppA_{2'}OmeP initiating capped oligonucleotide primer, 5 mM GTP, ATP, CTP and pseudouridine triphosphate (ΨTP). In subsequent examples this primer/NTP formulation will be referred to as Primer/NTP Formulation 3. In Primer/NTP Formulation 3, the GTP concentration was increased in order to produce commercially useful amounts of RNA. It is clear to one skilled in the art that other polymerases such as T7, T3 or SP6 RNA polymerases could be used instead of T7 RNA polymerase to perform the same function by using their respective promoters. Transcription reaction mixtures are incubated at 37°C for 2 hours. Reactions were supplemented with 10 mM Tris-HCl (pH 7.6), 2.5 mM MgCl₂, 0.5 mM CaCl₂ and 100 units/mL DNase I (New England Biolabs catalog #M0303) and incubated for 1 hour at 37°C. The resulting mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs were eluted in water and dephosphorylated by adjusting the solution to 50 mM Bis-Tris-Propane HCl (pH 6.0), 1 mM MgCl₂, 0.1 mM ZnCl₂ and 250 units/mg Antarctic phosphatase (New England Biolabs catalog #M0289). The reaction was incubated at 37 °C for 1 hour for Primer/NTP

Formulation 2 and 3 hours for Primer/NTP Formulation 3 hour. The resulting mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs were eluted in water. The purified transcription yields with Primer/NTP Formulation 1 and 2 were 0.7 milligram/milliliter transcription reaction (mg/mL) and 3.9 g/mL transcription reaction, respectively. The yield with Primer/NTP Formulation 2 was greatly superior to that obtained with Primer/NTP Formulation 1.

EXAMPLE 15

***In vitro* Transcription with ^{m7}GpppA_{2'}OmpG Initiating Capped Oligonucleotide Primers on a Transcription Template With Cytidine Residues at Template Positions +1 and +2**

[00175] A double stranded firefly luciferase DNA transcription template was used in which the +1 and +2 template nucleotides were cytidines and thus not completely complementary to the ^{m7}GpppA_{2'}OmpG initiating capped oligonucleotide primer. Two transcription reactions were assembled with 25 ug/mL transcription template, 40 mM Tris-HCl (pH 8.0), 27 mM MgCl₂, 2 mM spermidine, 10 mM DTT, 0.002% Triton X-100, 1000 unit/mL murine RNase inhibitor (New England Biolabs catalog #M0314), 2 unit/mL inorganic pyrophosphatase (New England Biolabs catalog # M2403), 4000 units/mL T7 RNA polymerase (New England Biolabs catalog #M0251). The two transcriptions differed in the amount of initiating capped oligonucleotide and NTPs. The first transcription was designed to mimic the transcription conditions of Ishicawa *et al.* (Nucleic Acids Symposium Series No. 53: 129 (2009)). This transcription reaction contained Primer/NTP Formulation 2. The second transcription used Primer/NTP Formulation 3. It is clear to one skilled in the art that other polymerases such as T7, T3 or SP6 RNA polymerases could be used instead of T7 RNA polymerase to perform the same function by using their respective promoters. Transcription reaction mixtures were incubated at 37°C for 2 hours. Reactions were supplemented with 10 mM Tris-HCl (pH 7.6), 2.5 mM MgCl₂, 0.5 mM CaCl₂ and 100 units/mL DNase I (New England Biolabs catalog #M0303) and incubated for 1 hour at 37°C. The resulting mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions or by reverse phase high performance liquid chromatography. mRNAs were dephosphorylated by adjusting the solution to 50 mM Bis-Tris-Propane HCl (pH 6.0), 1 mM MgCl₂, 0.1 mM ZnCl₂ and 250 units/mg Antarctic phosphatase (New England Biolabs catalog #M0289). The reaction were incubated at 37 °C for 1 hour for Primer/NTP Formulation 2 and 3 hours for Primer/NTP Formulation 3. The resulting

mRNAs were purified using an RNeasy Maxi kit (Qiagen catalog # 75162) according to manufacturer's instructions. mRNAs were eluted in water. The purified transcription yields with Primer/NTP Formulation 1 and 2 were 0.6 milligram/milliliter transcription reaction (mg/mL) and 3.9 g/mL transcription reaction, respectively. The yield with Primer/NTP Formulation 2 was greatly superior to that obtained with Primer/NTP Formulation 1.

EXAMPLE 16

Translation of mRNAs in Huh-7 Cells

(Figure 11)

[00176] Translational activity of luciferase mRNAs generated with initiating capped oligonucleotide primers was assessed in cultured hepatocytes in triplicate. Huh-7 cells were cultured in DMEM supplemented with 10% FBS, L-glutamine, non-essential amino acids and penicillin/streptomycin at 37°C under an atmosphere of 5% CO₂. Cells were transfected with 400 ng of mRNAs. For comparison, cells were also transfected with a Cap0 luciferase mRNA generated by initiation with ARCA. At 20 hours, cells were harvested and luciferase activity was measured using a ONE-Glo Luciferase Assay System kit (Promega catalog #E6120) according to manufacturer's recommendations. Luminescence was measured using a GloMax-Multi+ Detection System instrument according to manufacturer's recommendations. Luciferase activity was detected for all mRNAs tested (Figure 11). The fact that mRNAs generated with initiating capped oligonucleotide primers or 3'-O-methyl initiating capped oligonucleotide primers translate with similar efficiencies to ARCA capped RNAs indicates that they are efficiently capped co-transcriptionally.

EXAMPLE 17

Assay to determine capping efficiency of mRNA generated by co-transcriptional capping with ARCA or Initiating Capped Oligonucleotide Primer Using Primer/NTP Formulation 1.

(Figures 12A-12H)

[00177] For each initiating capped oligonucleotide primer tested, sufficient quantities of mRNA to be detected by liquid chromatography mass spectroscopy (LC-MS) was subjected to a capping assay. In this assay, a small fragment was cleaved from the 5' end of a full length mRNA and analyzed by LC-MS. Prior to cleavage of the mRNA, it was treated with Antarctic phosphatase (New England Biolabs catalog # M0289) to convert uncapped

monophosphates, diphosphates and triphosphates to a 5' OH to facilitate analysis. The phosphatase treated mRNA was then cleaved and purified. The purified RNA was subjected to LC-MS analysis. Figure 12 shows the LC traces. LC peaks corresponding to uncapped (5' OH after phosphatase treatment) and Cap 1 are indicated with observed masses. The inset schematic shows the alignment of the initiating oligonucleotide primer on the transcription template. Note that “[]” indicates a base pair of the capped initiating nucleotide with the template nucleotide in the schematic. Subscript “m” indicates a 2'-O-methyl group and superscript “m7” indicates a base methylation. For comparison, an mRNA co-transcriptionally capped with ARCA was subjected to the capping assay. An estimate of the capping efficiency was made using the following formula (intensity of capped peaks)/[(intensity of capped peak) + (intensity of the 5' OH peak)]. The observed % capping in Figure 12 was A) $m^7G_{3'Ome}pppG$ = 79%, B) $m^7GpppG_{2'Ome}pG$ = 89%, C) $m^7G_{3'Ome}pppG_{2'Ome}pG$ = 87%, D) $m^7GpppA_{2'Ome}pG$ = 99%, E) $m^7G_{3'Ome}pppA_{2'Ome}pG$ = 99%, F) $m^7GpppC_{2'Ome}pG$ = 98%, G) $m^7G_{3'Ome}pppC_{2'Ome}pG$ = 97%, H) $m^7GpppA_{2'Ome}pG_{2'Ome}pG$ = 50%. In each case, the capping efficiency of transcripts co-transcriptionally capped with initiating trimeric capped oligonucleotide primers was greater than that observed with $m^7G_{3'Ome}pppG$ (ARCA).

Example 18

Comparison of Capping with $m^7GpppA_{2'Ome}pG$ Initiating Capped Oligonucleotide On a Transcription Template With 2'-deoxythymidine and 2'-deoxycytidine Residues at Template Positions +1 and +2 vs. a Transcription Template With Cytidine Residues at Template Positions +1 and +2

(Figures 13A-13D)

[00178] mRNAs made in Examples 14 and 15 were subjected to a capping assay to determine relative efficiency and specificity of capping of transcripts with 2'-deoxythymidine and 2'-deoxycytidine residues at template positions +1 and +2 vs. a transcription template with cytidine residues at template positions +1 and +2. Sufficient quantities of mRNA to be detected by liquid chromatography mass spectroscopy (LC-MS) were subjected to a capping assay. In this assay, a small fragment was cleaved from the 5' end of a full length mRNA and analyzed by LC-MS. Prior to cleavage of the mRNA, it was treated with Antarctic phosphatase (New England Biolabs catalog # M0289) to convert uncapped monophosphates, diphosphates and triphosphates to a 5' OH to facilitate analysis. The phosphatase treated

mRNA was then cleaved and purified. The purified RNA was subjected to LC-MS analysis. Figure 13 shows the LC traces. LC peaks corresponding to uncapped and Cap 1 are indicated with observed masses. The inset schematic shows the alignment of the initiating oligonucleotide primer on the transcription template. Subscript "m" indicates a 2'-O-methyl group and superscript "m7" indicates a base methylation. Figure 13A and 13B transcripts were transcribed from templates with 2'-deoxythymidine and 2'-deoxycytidine at template nucleotides +1 and +2, respectively, using Primer/NTP Formulations 2 and 3, respectively. Figures 13C and 13D were transcripts transcribed from templates with 2'-deoxycytidines at template nucleotides +1 and +2 using Primer/NTP Formulations 2 and 3, respectively. When the ^{m7}GpppA_{2'}OmeP_G initiating oligonucleotide was completely complementary to the +1 and +2 nucleotides, the major product observed was the desired templated Cap 1 transcript initiated with ^{m7}GpppA_{2'}OmeP_G... (Figures 12A and 12B). For Figures 13A and 13B, an estimate of the capping efficiency was made using the following formula (intensity of capped peaks)/[(intensity of capped peaks) + (intensity of the 5' OH peak)] and capping efficiency with Primer/NTP Formulation 2 and Primer/NTP Formulation 3 was 99 % and 96%, respectively. Only minor aberrant initiation products were detected. In contrast, when template nucleotides +1 and +2 are cytidine, there is not perfect complementarity between the ^{m7}GpppA_{2'}OmeP_G initiating oligonucleotide and these template nucleotides. Figures 12C and 12D show that the initiating capped oligonucleotide initiated in two registers. In the first register, the 3' guanosine initiating capped oligonucleotide residue paired with the +1 template cytidine to produce a transcript with an additional untemplated 5' adenosine (Cap 1 + A). Note that "[]" indicates a base pair of the capped initiating nucleotide with the template nucleotide in the schematic. Designated template nucleotide position is indicated. In the second register, the 3' guanosine initiating capped oligonucleotide residue paired with the +2 template cytidine and the +1 initiating capped oligonucleotide adenosine does not form a complete hybrid with the +1 template nucleotide. This produces a transcript where the templated 5' guanosine has been replaced with an untemplated adenosine (Cap 1 GtoA). An estimate of the capping efficiency was made using the following formula (intensity of capped peaks "Cap1 + A" + "Cap1 GtoA")/[(intensity of capped peaks "Cap 1 + A" + "Cap 1 GtoA") + (intensity of the 5' OH peak)]. The calculated capping efficiency in Figure 13C and 13D were 97 and 77 %. We note the surprising finding that capping efficiency and templated transcription initiation fidelity were much higher when the initiating capped oligonucleotide primer was completely complementary to the corresponding template +1 and

+2 nucleotides. In addition, with our method, efficient capping can be achieved without reducing the concentration of an NTP to drive capping, allowing much greater transcription yields. Thus the capping method we describe is distinct from and superior to that of Ishikawa *et al.*

Example 19

(Figure 14A-14B)

Comparison of Translation in Differentiated THP-1 Cells of mRNA Made with ^{m7}GpppA_{2'}Ome pG Initiating Capped Oligonucleotide On a Transcription Template With 2'-deoxythymidine and 2'-deoxycytidine Residues at Template Positions +1 and +2 vs. a Transcription Template With Cytidine Residues at Template Positions +1 and +2

[00179] In order to assess the expression of luciferase mRNAs generated in Examples 14 and 15, mRNAs were transfected into THP-1 cells (ATCC, Catalog# TIB-202) with six replicates. THP-1 cells were cultured in ATCC formulated RPMI-1640 (ATCC, Catalog# 30-2001) supplemented with 10% FBS, sodium pyruvate and penicillin/streptomycin at 37°C under an atmosphere of 5% CO₂. Cells were seeded at 2E+05 cells per well in a 24-well plate format in the presence of the phorbol ester 12-O-Tetradecanoylphorbol-13-acetate (TPA; Cell Signaling Technologies, Catalog# 4174) to induce differentiation. Cells were transfected with 100 ng of mRNA per well 72 hours after seeding. For comparison, cells were also transfected with a Cap0 luciferase mRNA generated by initiation with ARCA. At 20 hours post-transfection, cells were harvested and luciferase activity was measured using a ONE-Glo Luciferase Assay System kit (Promega Catalog# E6120) according to the manufacturer's recommendations. Luminescence was measured using a GloMax-Multi+ Detection System instrument according to the manufacturer's recommendations. Data were graphed as the mean of six replicates +/- the standard deviation from the mean. Data were analyzed using the Unpaired t-test to generate p values as a measure of significance. The p-value is between pairs is indicated. Translation was compared in THP-1 cells for transcripts were generated with ^{m7}GpppA_{2'}Ome pG initiating capped oligonucleotide and a template comprising 2'-deoxythymidine and 2'-deoxycytidine residues at template positions +1 and +2 ("TC" Template) vs. a transcription template With cytidine residues at template Positions +1 and +2 ("CC" Template). A) Transcription was conducted with Primer/NTP Formulations 2 or B) Primer/NTP Formulations 3. With both formulations, translation in cultured THP-1 cells was significantly superior when the ^{m7}GpppA_{2'}Ome pG initiating capped oligonucleotide primer

was completely complementary to template nucleotides +1 and +2 ("TC" template) as described in the current invention. In Figure 14B, we also assessed the activity of an mRNA that was made with the "TC" Template and capped with ARCA (Cap 0). This is the current industry standard for generating co-transcriptionally capped mRNAs. The ARCA capped RNA had significantly less activity than transcripts made where the ^{m7}GpppA_{2'}Omc pG initiating capped oligonucleotide primer was completely complementary to template nucleotides +1 and +2 ("TC" template).

[00180] In sum, the Examples shown here demonstrate that, relative to previously published methods, the methods described here generate RNAs with a combination of 1) high yield, 2) high extent of capping, 3) high fidelity of templated transcription and 4) superior activity in cells.

[00181] The inventions illustratively described herein may suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms "a" and "an" and "the" and similar referents in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms "comprising", "having," "including," "containing", etc. shall be read expansively and without limitation (*e.g.*, meaning "including, but not limited to,"). Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (*e.g.*, "such as") provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention

has been specifically disclosed with reference to certain embodiments and optional features, modification and variation of the inventions embodied therein herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention. Thus, it should be understood that although the present invention has been specifically disclosed by reference to certain embodiments and optional features, modification, improvement and variation of the inventions embodied therein herein disclosed may be resorted to by those skilled in the art, and that such modifications, improvements and variations are considered to be within the scope of this invention. The materials, methods, and examples provided here are representative of certain embodiments, are exemplary, and are not intended as limitations on the scope of the invention.

[00182] The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein.

[00183] In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[00184] All publications, patent applications, patents, and other references mentioned herein are expressly incorporated by reference in their entirety, to the same extent as if each were incorporated by reference individually. In case of conflict, the present specification, including definitions, will control.

[00185] Applicants reserve the right to physically incorporate into this application any and all materials and information from any such articles, patents, patent applications, or other physical and electronic documents.

[00186] Other embodiments are set forth within the following claims.

CLAIMS

What is claimed is:

1. An initiating capped oligonucleotide primer comprising the following structure:



wherein:

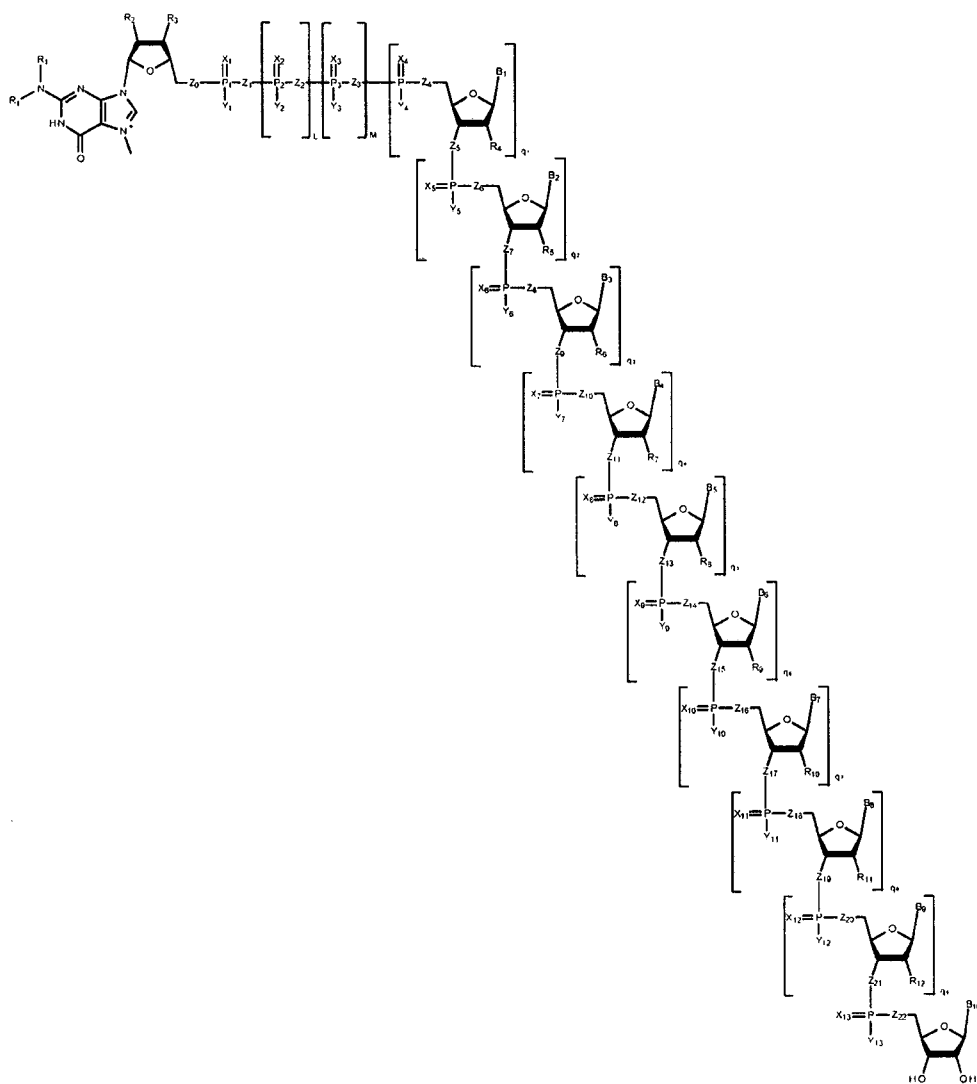
- each of B₁ through B₁₀ is independently a natural, modified or unnatural nucleoside base;
- M is 0 or 1;
- L is 0 or 1;
- q₁ is 1 and each of q₂ through q₉ is independently 0 or 1;
- R₁ is H or methyl;

R_2 and R_3 are independently H, OH, alkyl, O-alkyl, halogen, a linker or a detectible marker;
 each of X_1 through X_{13} is independently O or S;
 each of Y_1 through Y_{13} is independently OH, SH, BH_3 , aryl, alkyl, O-alkyl or O-aryl;
 each of Z_1 through Z_{22} is independently O, S, NH, CH_2 , C(halogen)₂ or CH(halogen);
 and
 each of R_4 through R_{12} are independently H, OH, OMe or a detectable marker.

2. An initiating capped oligonucleotide primer according to claim 1 wherein:
 $q_1=1$, and
 each of $q_2-q_9 = 0$ or 1.
3. An initiating capped oligonucleotide primer according to claim 1 or 2, wherein:
 B_1 is completely complementary to nucleoside base on the nucleic acid template at transcription template position +1,
 B_2 through B_9 , if present, are completely complementary to respective nucleoside bases on nucleic acid template at transcription start site from position +2 and on, and
 B_{10} , the last nucleoside base of initiating capped oligonucleotide primer, is completely complementary to the last hybridized transcription template nucleotide.
4. An initiating capped oligonucleotide primer according to claim 1 wherein:
 $q_2-q_9 = 0$, and
 B_{10} hybridizes to template nucleotide 2.
5. An initiating capped oligonucleotide primer according to claim 1 wherein:
 $q_1-q_2=1$
 $q_3-9=0$, and
 B_2 and B_{10} hybridize to template nucleotide 2 and 3, respectively.
6. An initiating capped oligonucleotide primer according to claim 1 wherein:
 $q_1-q_3=1$,
 $q_4-9=0$, and
 B_2 , B_3 and B_{10} hybridize to template nucleotide 2, 3 and 4, respectively.
7. An initiating capped oligonucleotide primer according to claim 1 wherein:
 $q_1-q_4=1$,
 $q_5-9=0$, and
 B_2 , B_3 , B_4 and B_{10} hybridize to template nucleotide 2, 3, 4 and 5, respectively.
8. An initiating capped oligonucleotide primer according to claim 1 wherein:
 $q_1-q_5=1$,

- q6-9=0, and
B2, B3, B4, B5 and B10 hybridize to template nucleotide 2, 3, 4, 5 and 6,
respectively.
9. An initiating capped oligonucleotide primer according to claim 1 wherein:
q1-q6=1,
q7-9=0, and
B2, B3, B4, B5, B6 and B10 hybridize to template nucleotide 2, 3, 4, 5, 6 and 7,
respectively.
10. An initiating capped oligonucleotide primer according to claim 1 wherein:
q1-q7=1,
q8-9=0, and
B2, B3, B4, B5, B6, B7 and B10 hybridize to template nucleotide 2, 3, 4, 5, 6, 7 and
8, respectively.
11. An initiating capped oligonucleotide primer according to claim 1 wherein:
q1-q8=1,
q9=0, and
B2, B3, B4, B5, B6, B7, B8 and B10 hybridize to template nucleotide 2, 3, 4, 5, 6, 7,
8 and 9, respectively.
12. An initiating capped oligonucleotide primer according to claim 1 wherein:
q1-q9=1, and
B2, B3, B4, B5, B6, B7, B8, B9 and B10 hybridize to template nucleotide 2, 3, 4, 5, 6,
7, 8, 9 and 10, respectively.
13. An initiating capped oligonucleotide primer according to claim 1 wherein:
q1-q2=1,
q3-9=0, and
B2 and B10 hybridize to template nucleotide 2 and 3, respectively.
14. An RNA molecule comprising the initiating capped oligonucleotide primer according
to claim 1.
15. A cell containing an RNA molecule comprising the initiating capped oligonucleotide
primer according to claim 1.
16. A cell containing a protein or a peptide translated from an RNA molecule comprising
the initiating capped oligonucleotide primer according to claim 1.

17. A pharmaceutical composition comprising an RNA molecule comprising the initiating capped oligonucleotide primer according to claim 1 and a pharmaceutical acceptable carrier.
18. A method for synthesizing a fully templated RNA molecule comprising the steps of:
introducing the initiating capped oligonucleotide primer with the structure



wherein:

- each of B₁ through B₁₀ is independently a natural, modified or unnatural nucleoside base;
- M is 0 or 1;
- L is 0 or 1;

R_1 is H or methyl;

R_2 and R_3 are independently H, OH, alkyl, O-alkyl, halogen, a linker or a detectible marker;

each of X_1 through X_{13} is independently O or S;

each of Y_1 through Y_{13} is independently OH, SH, BH_3 , aryl, alkyl, O-alkyl or O-aryl;

each of Z_1 through Z_{22} is independently O, S, NH, CH_2 , $C(halogen)_2$ or $CH(halogen)$;

and

each of R_4 through R_{12} are independently H, OH, OMe or a detectable marker

into a mixture comprising an RNA polymerase under conditions conducive to transcription by the RNA polymerase of the polynucleotide template and

incubating said mixture for a time sufficient to allow for transcription of said template.

Figure 1

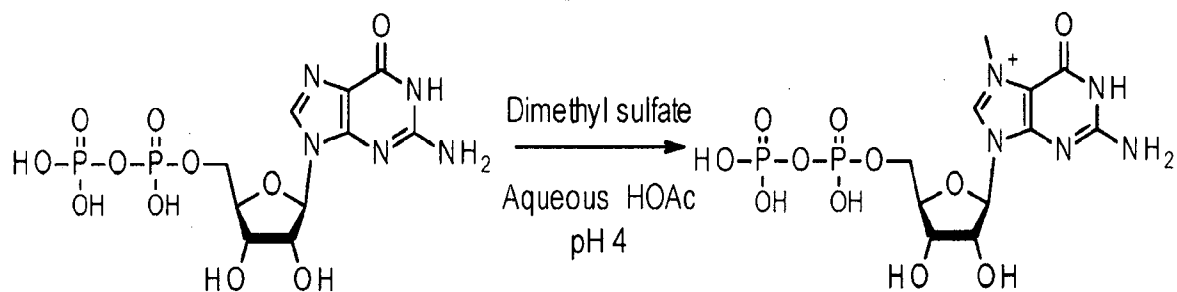


Figure 2

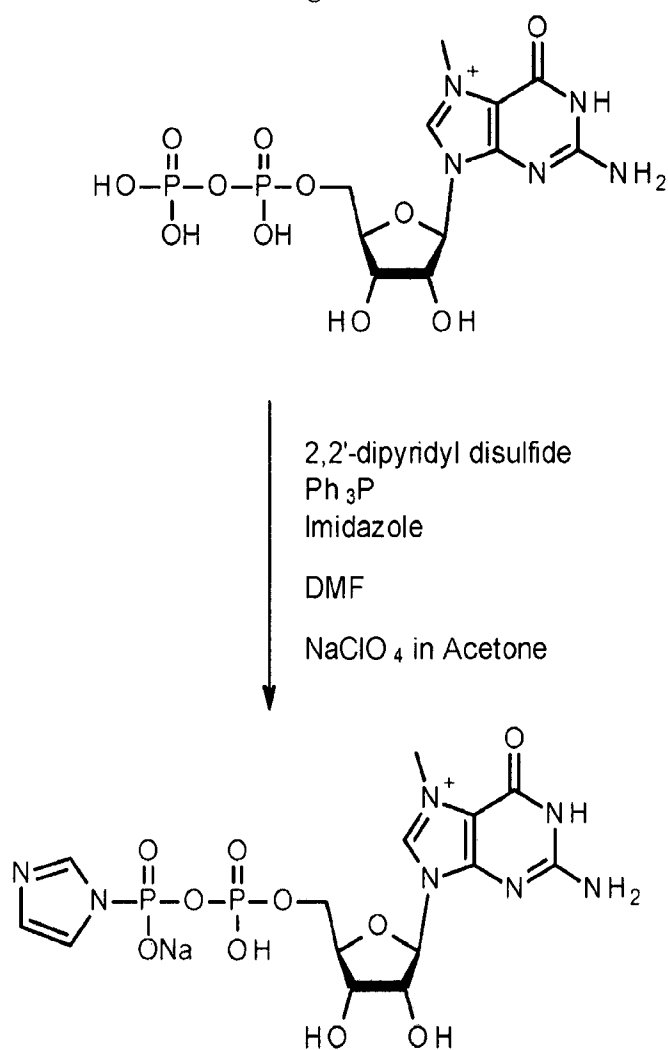


Figure 3

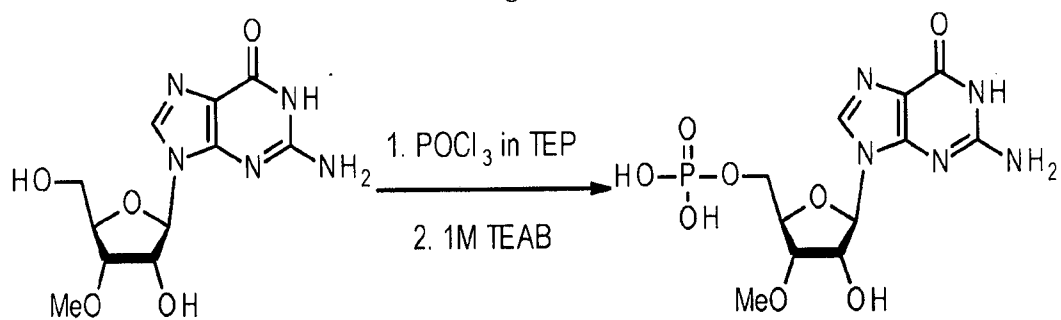


Figure 4

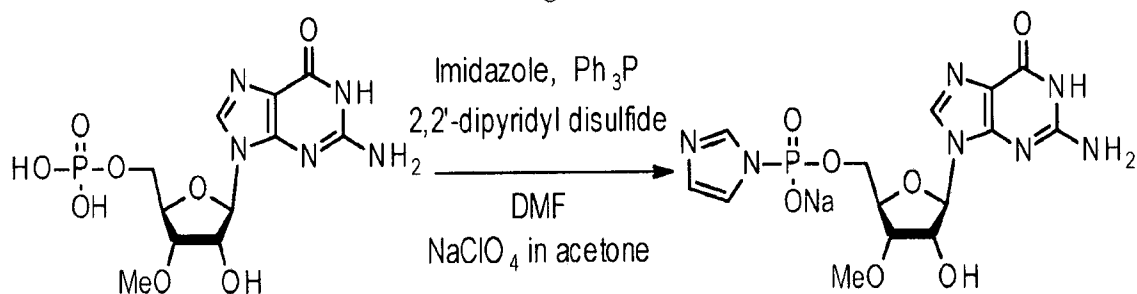


Figure 5

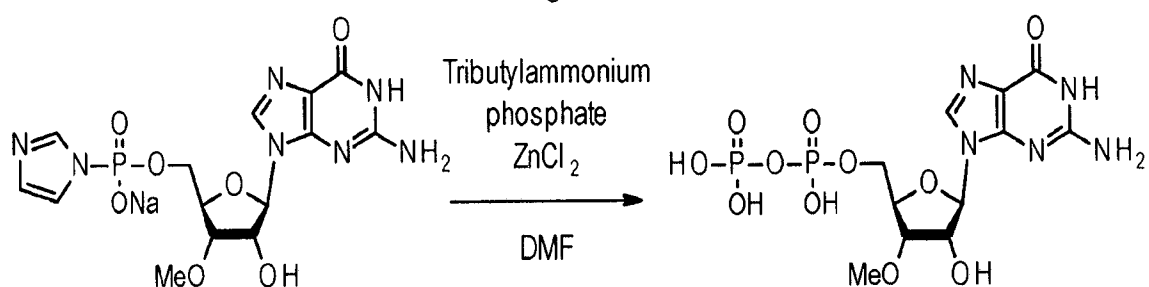


Figure 6

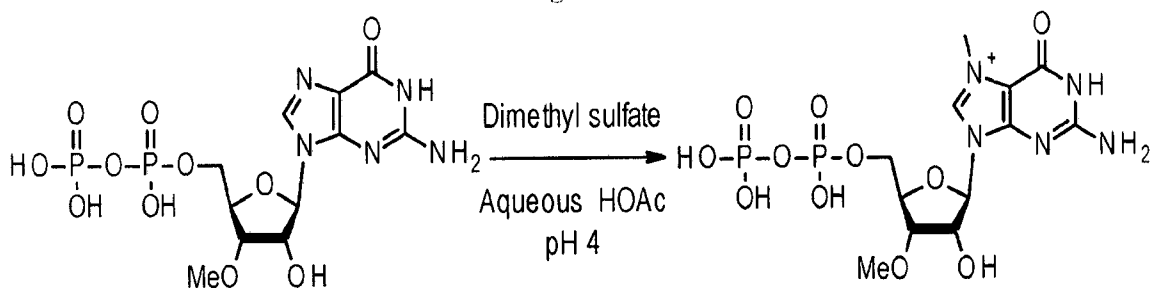


Figure 7

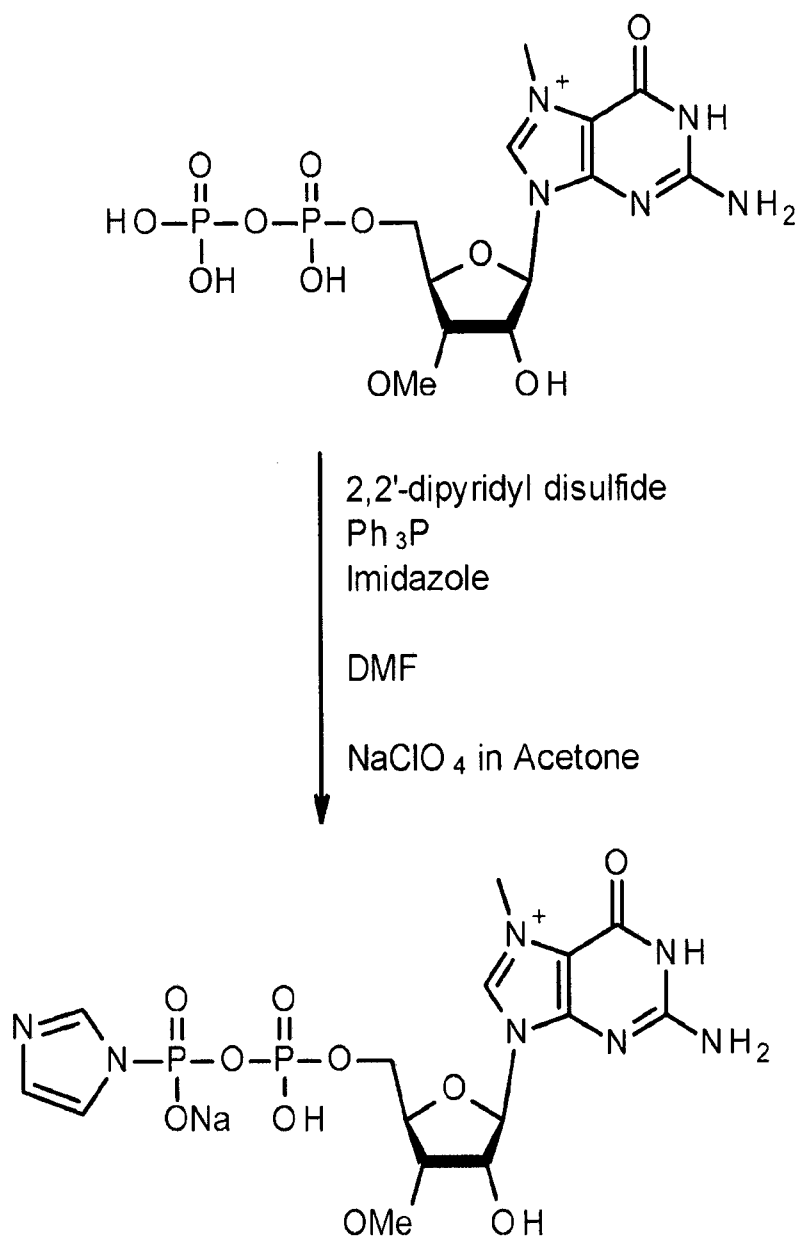
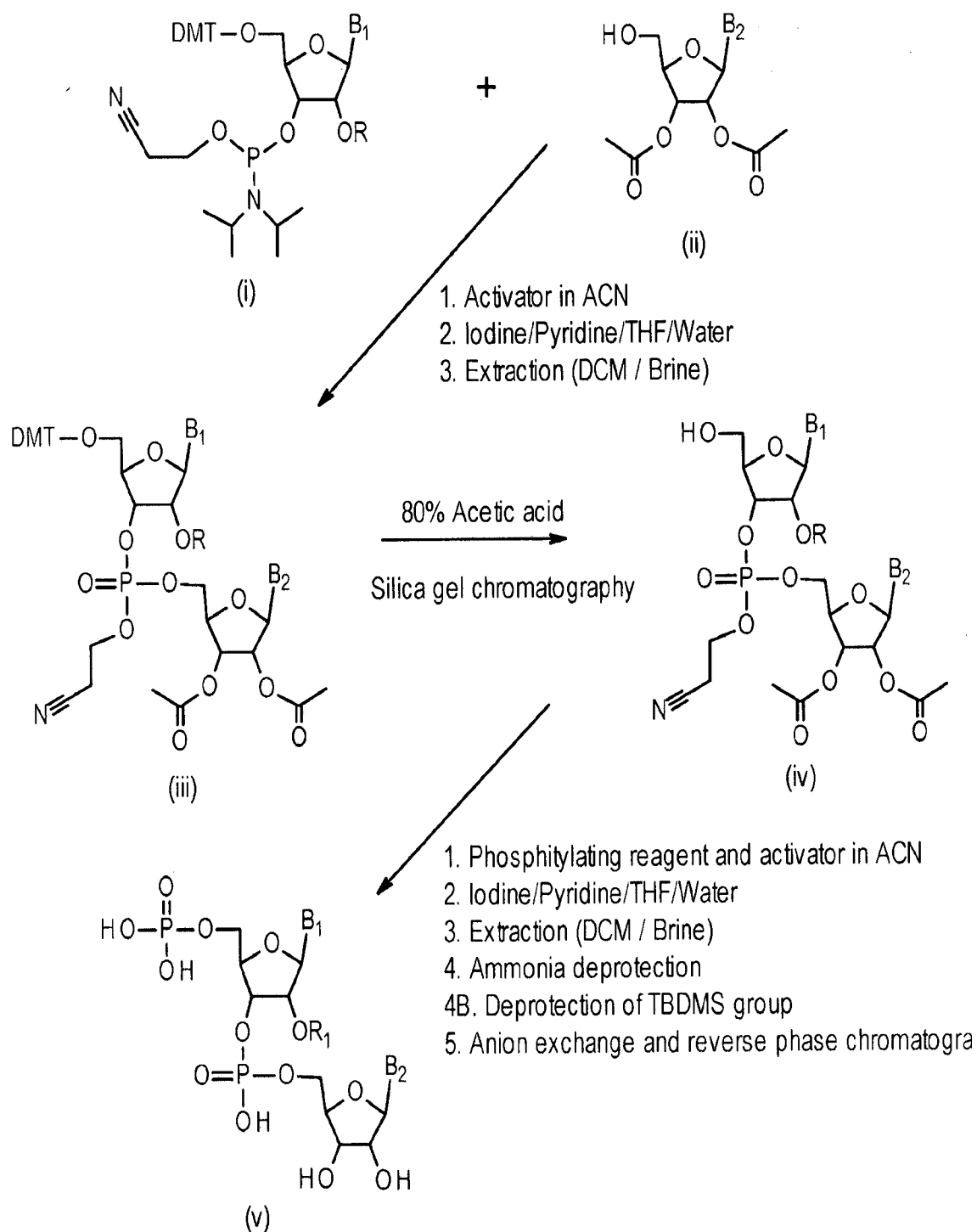


Figure 8



B1 and B2 are appropriately protected or unprotected nucleoside bases
 R = Me or TBDMS; R₁ = Me or H

Figure 9

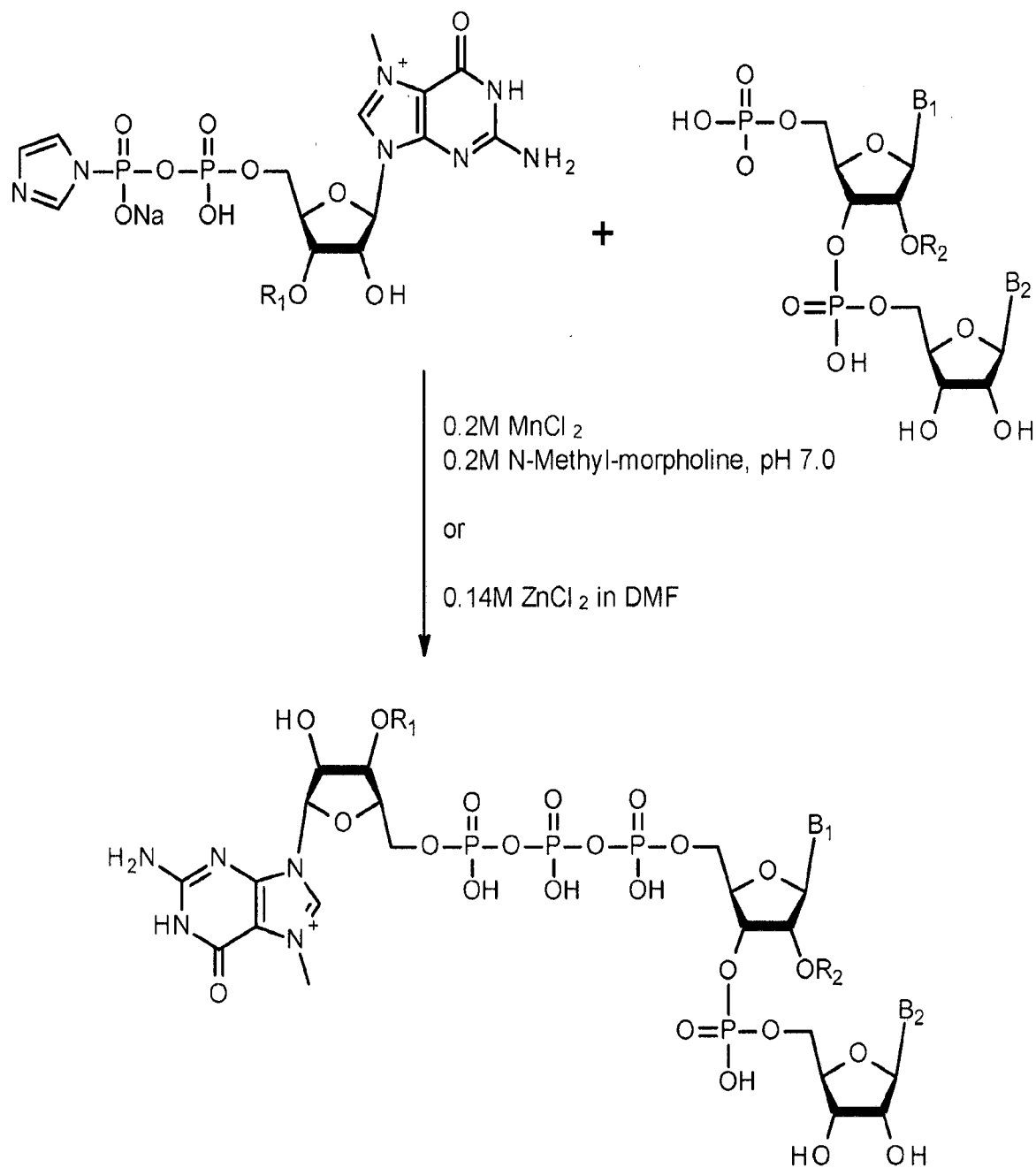


Figure 10A

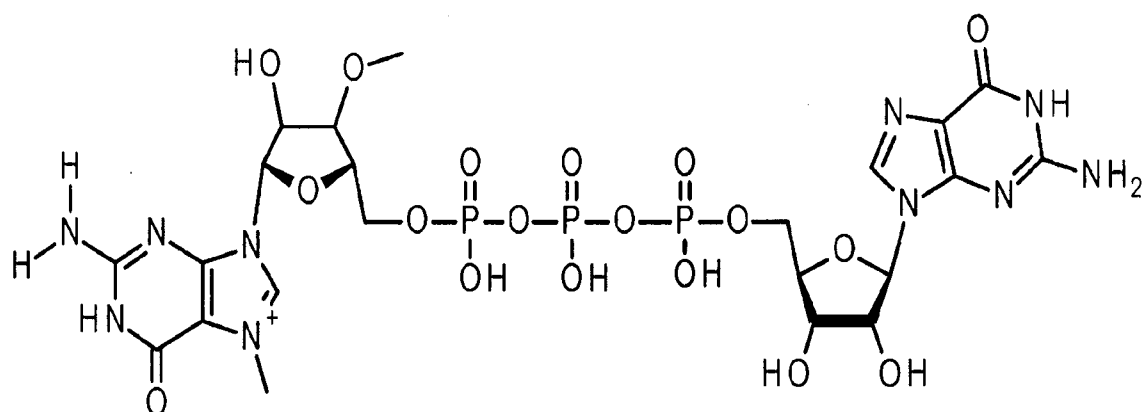


Figure 10B

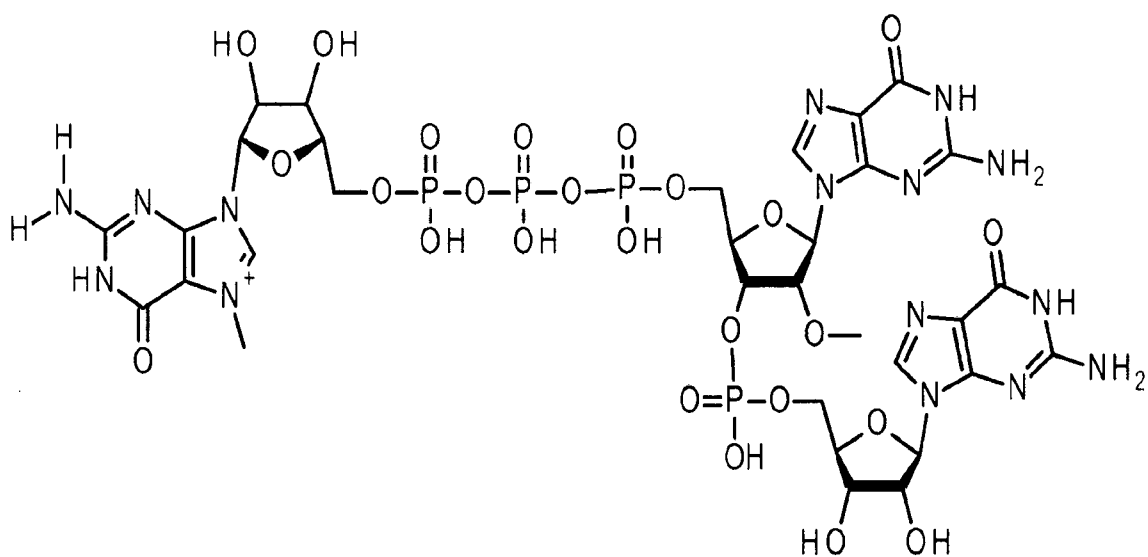


Figure 10C

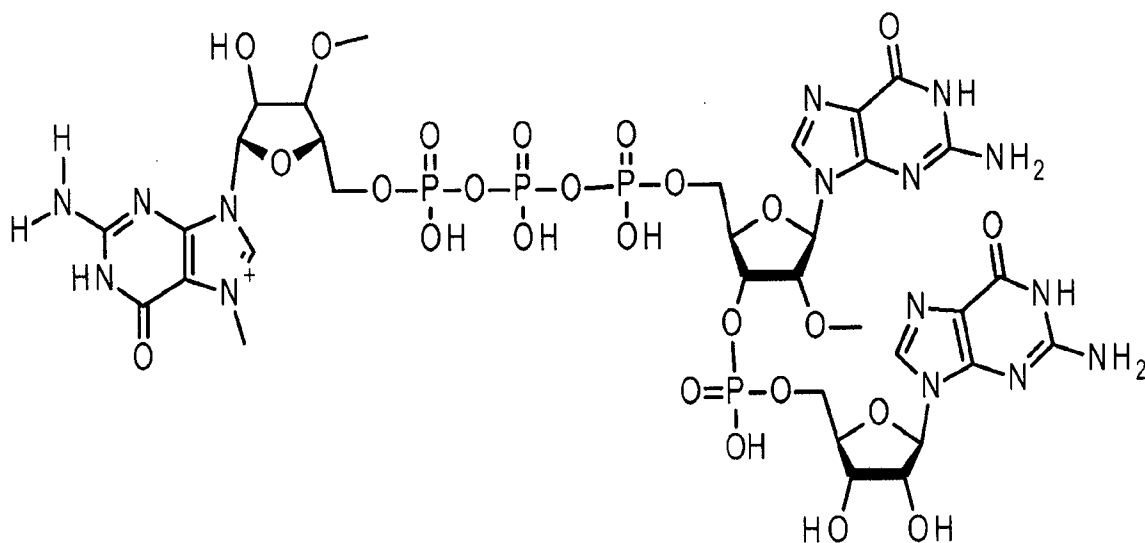


Figure 10D

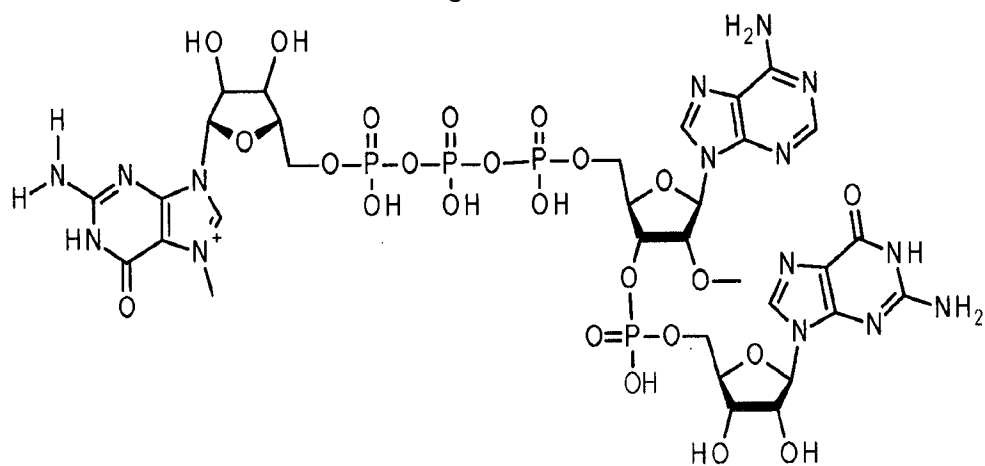


Figure 10E

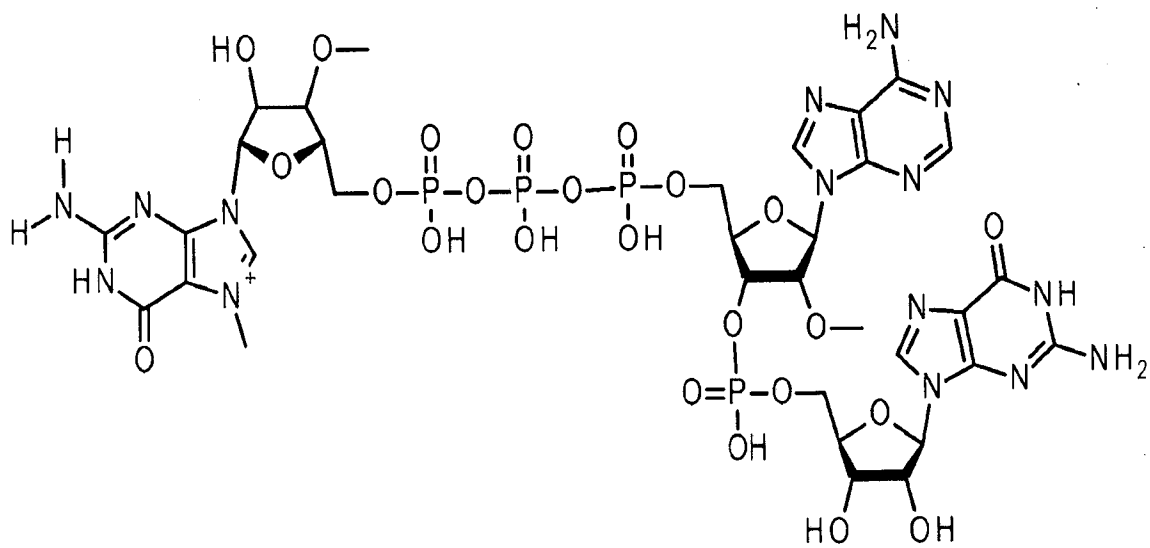


Figure 10F

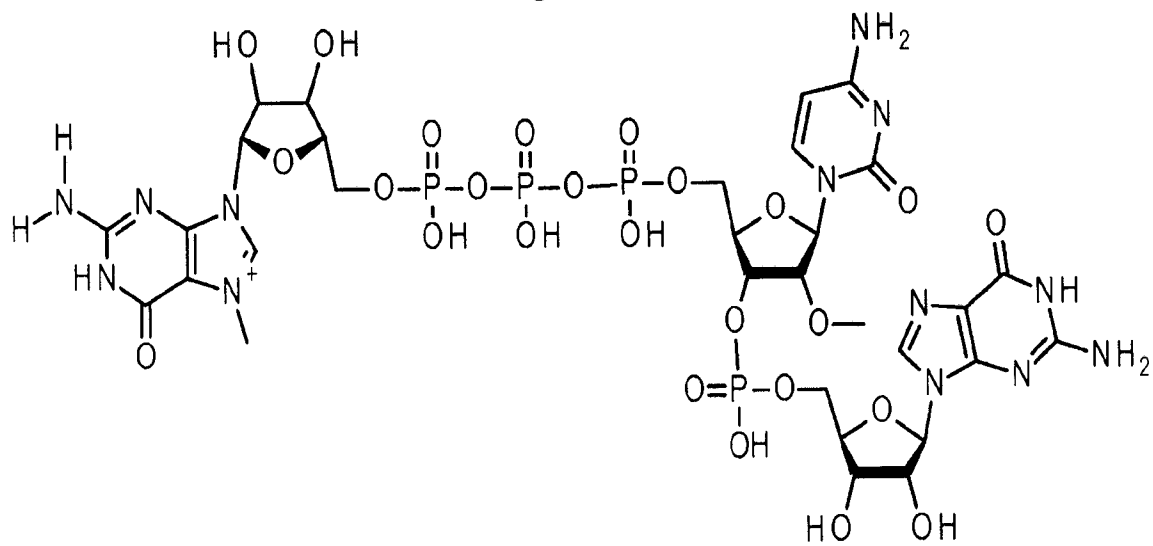


Figure 10G

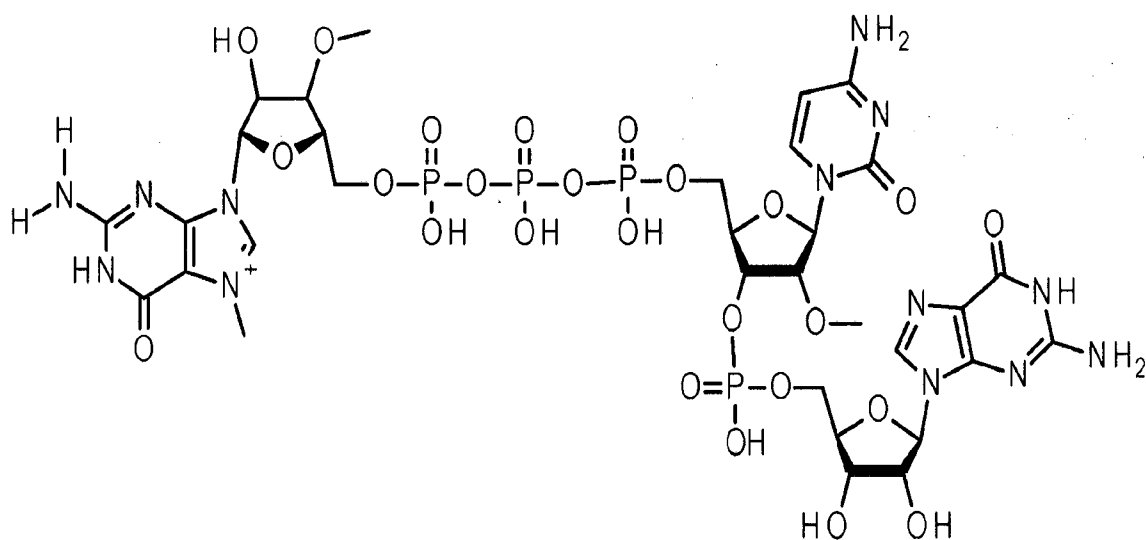


Figure 10H

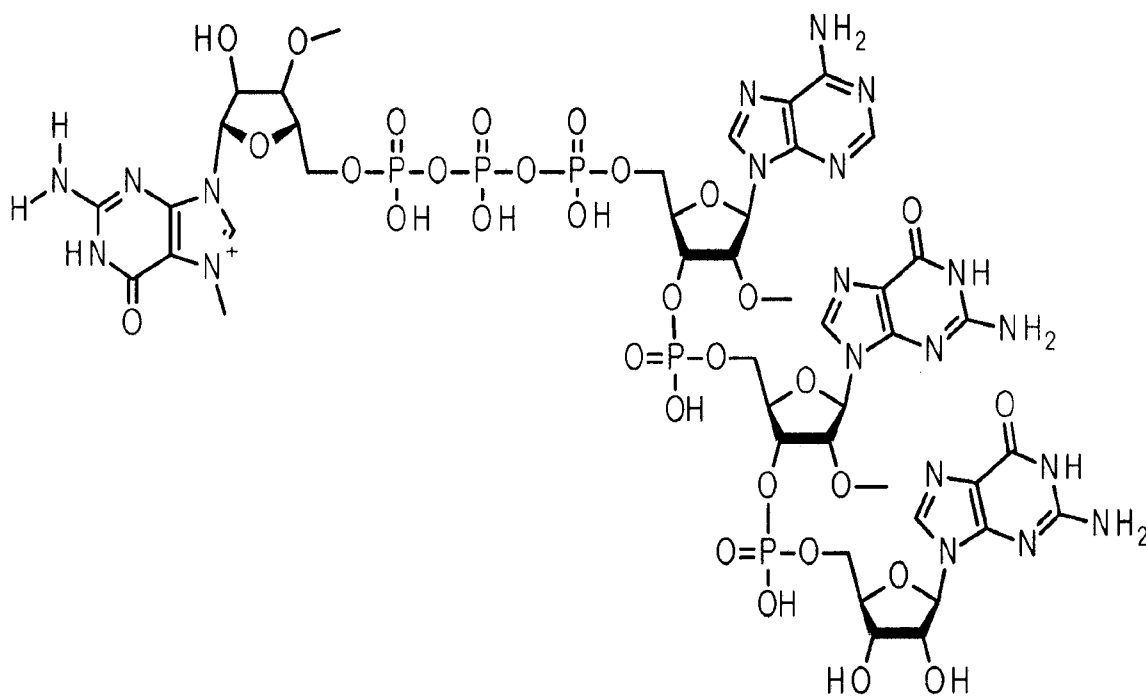


Figure 11

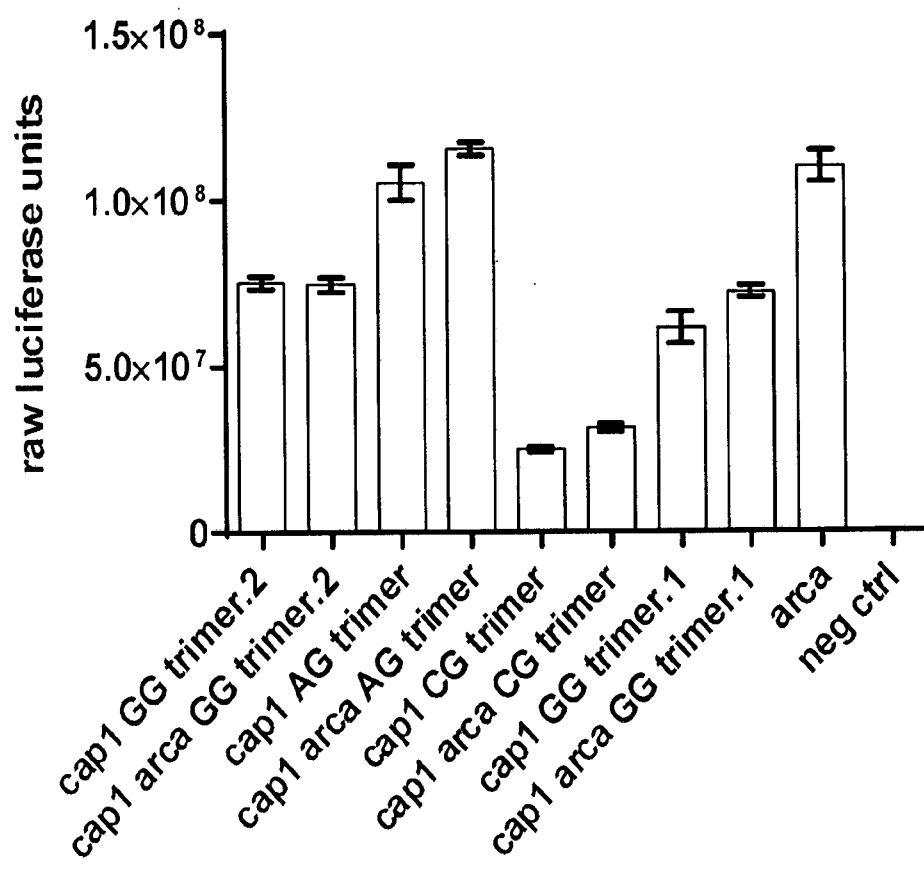


Figure 12A

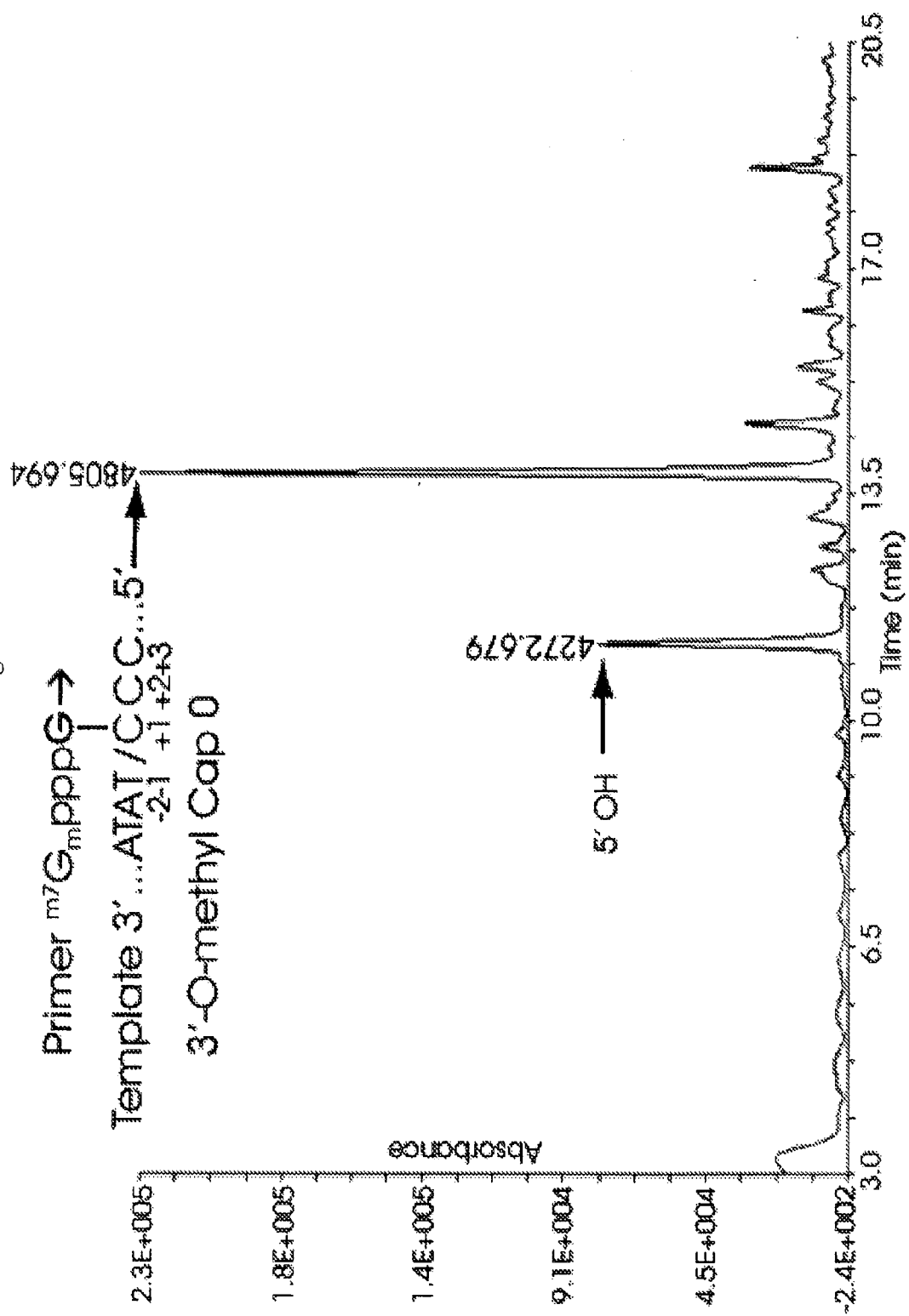


Figure 12B

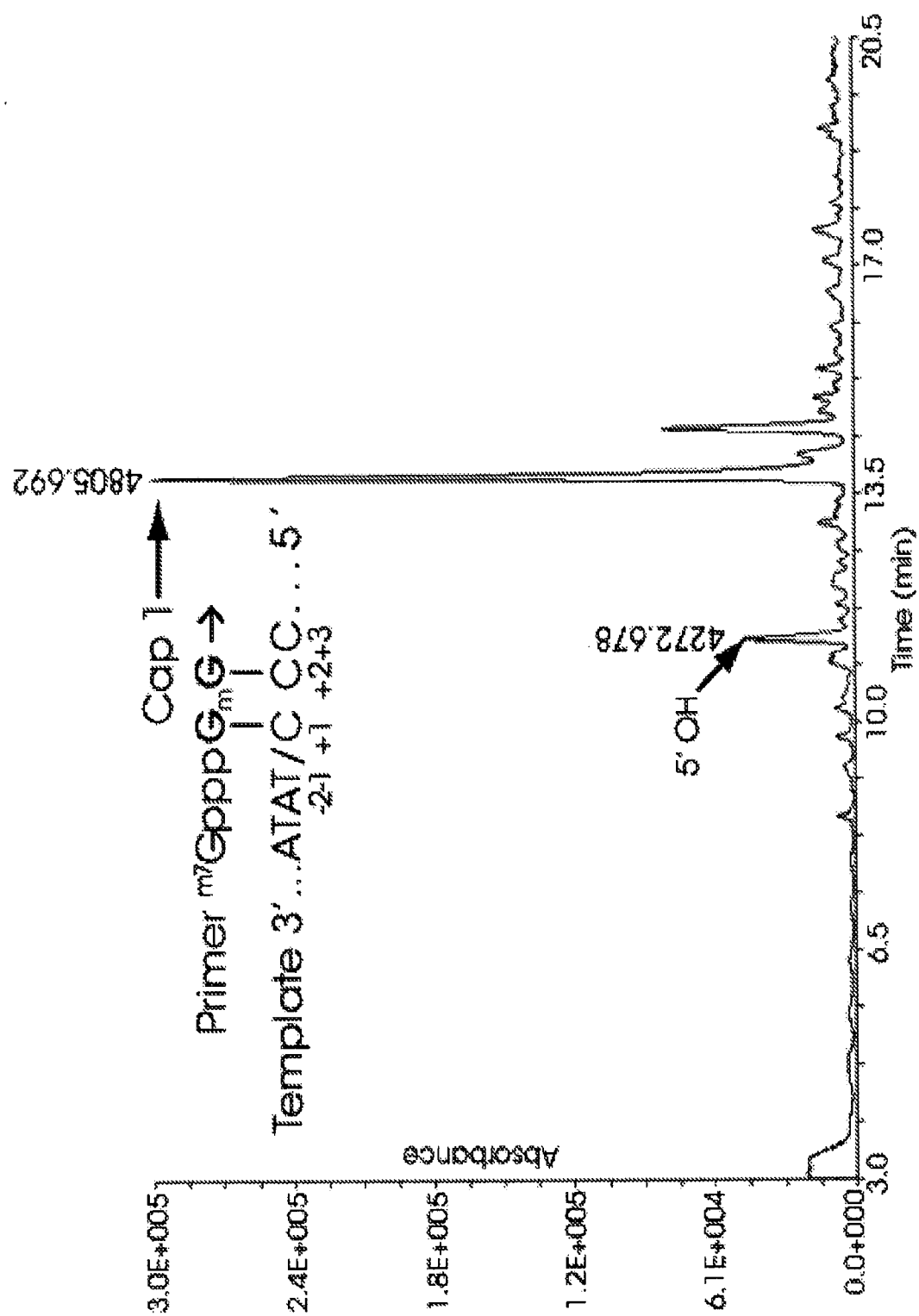


Figure 12C

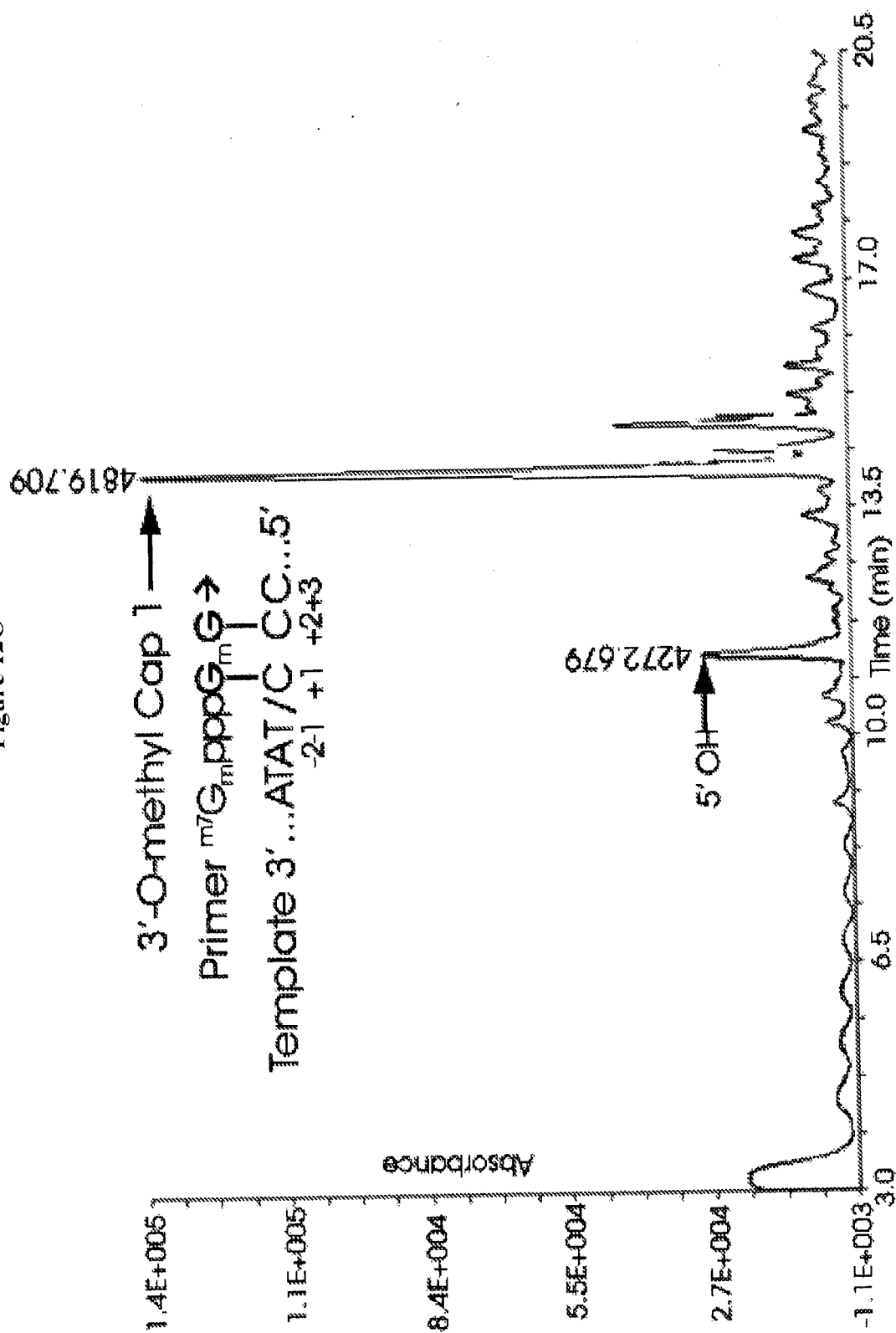


Figure 12D

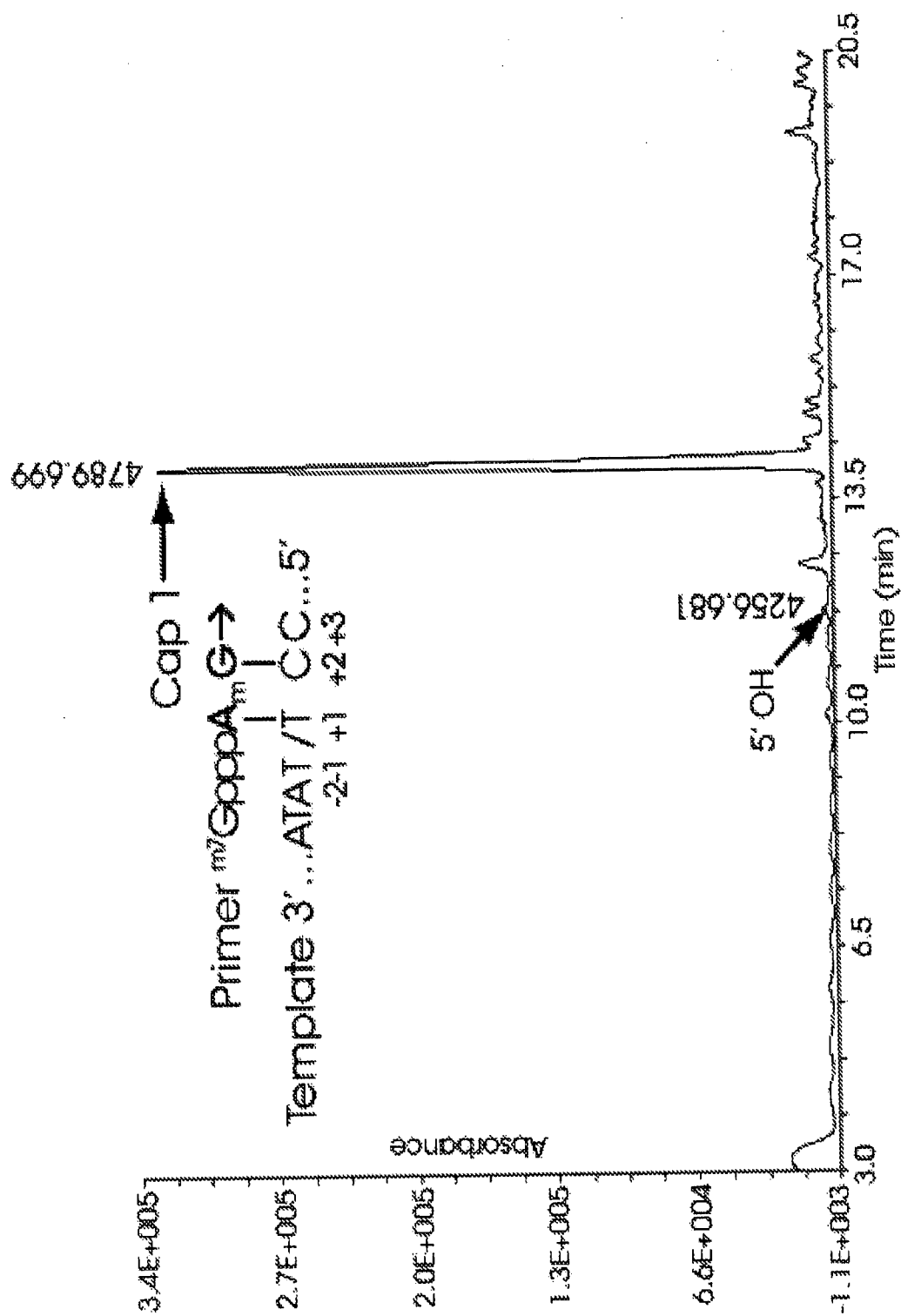


Figure 12E

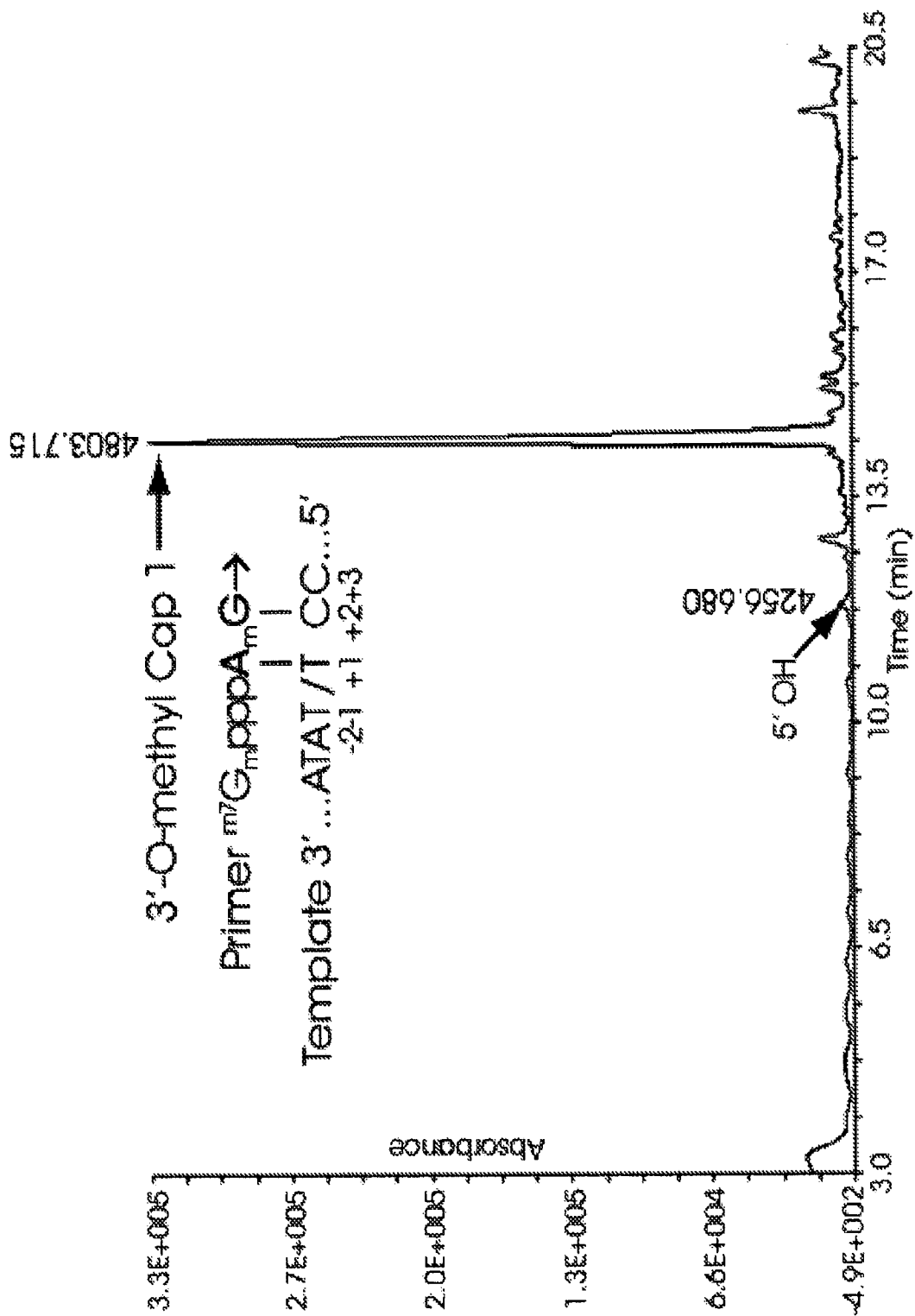


Figure 12F

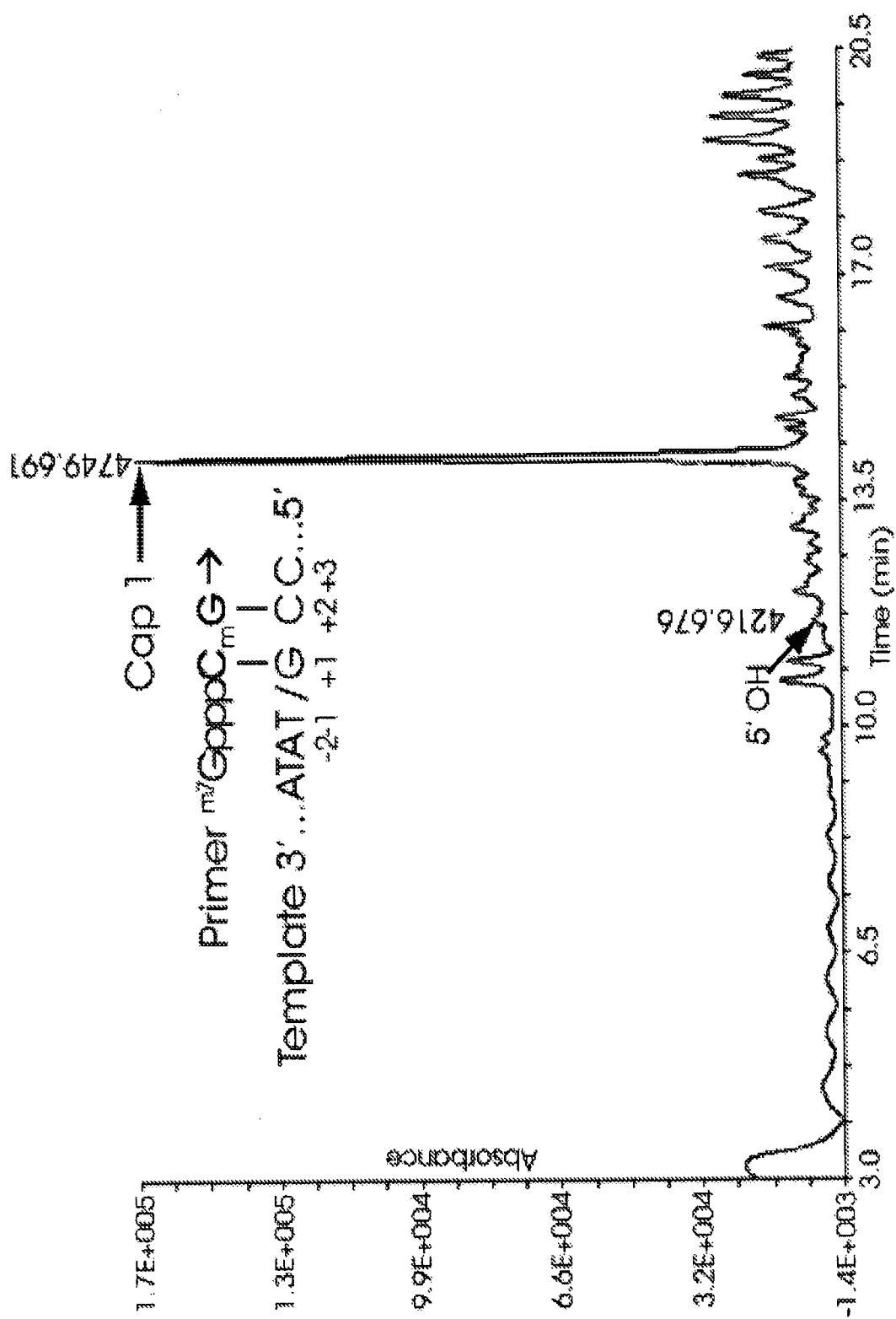


Figure 12G

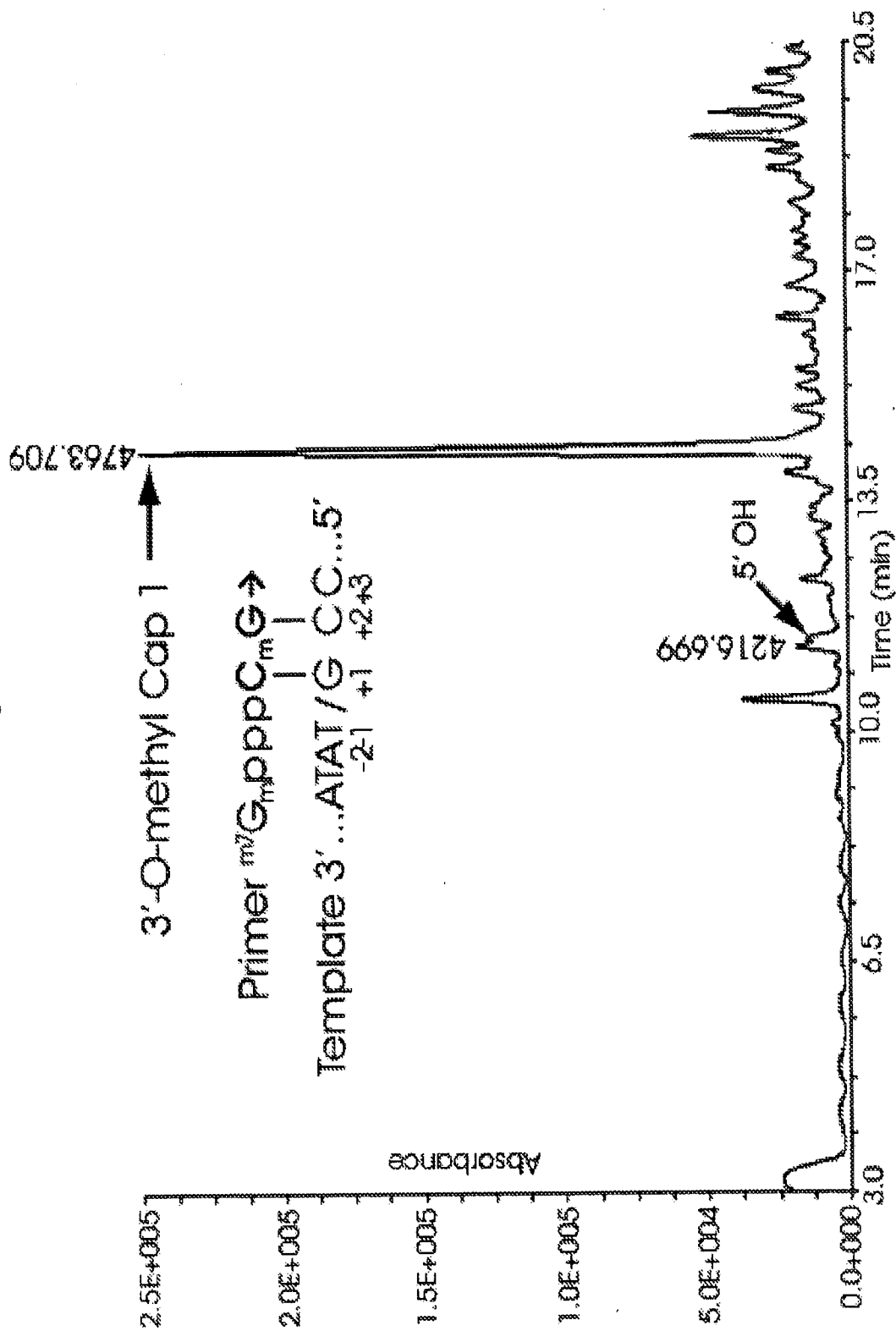


Figure 12H

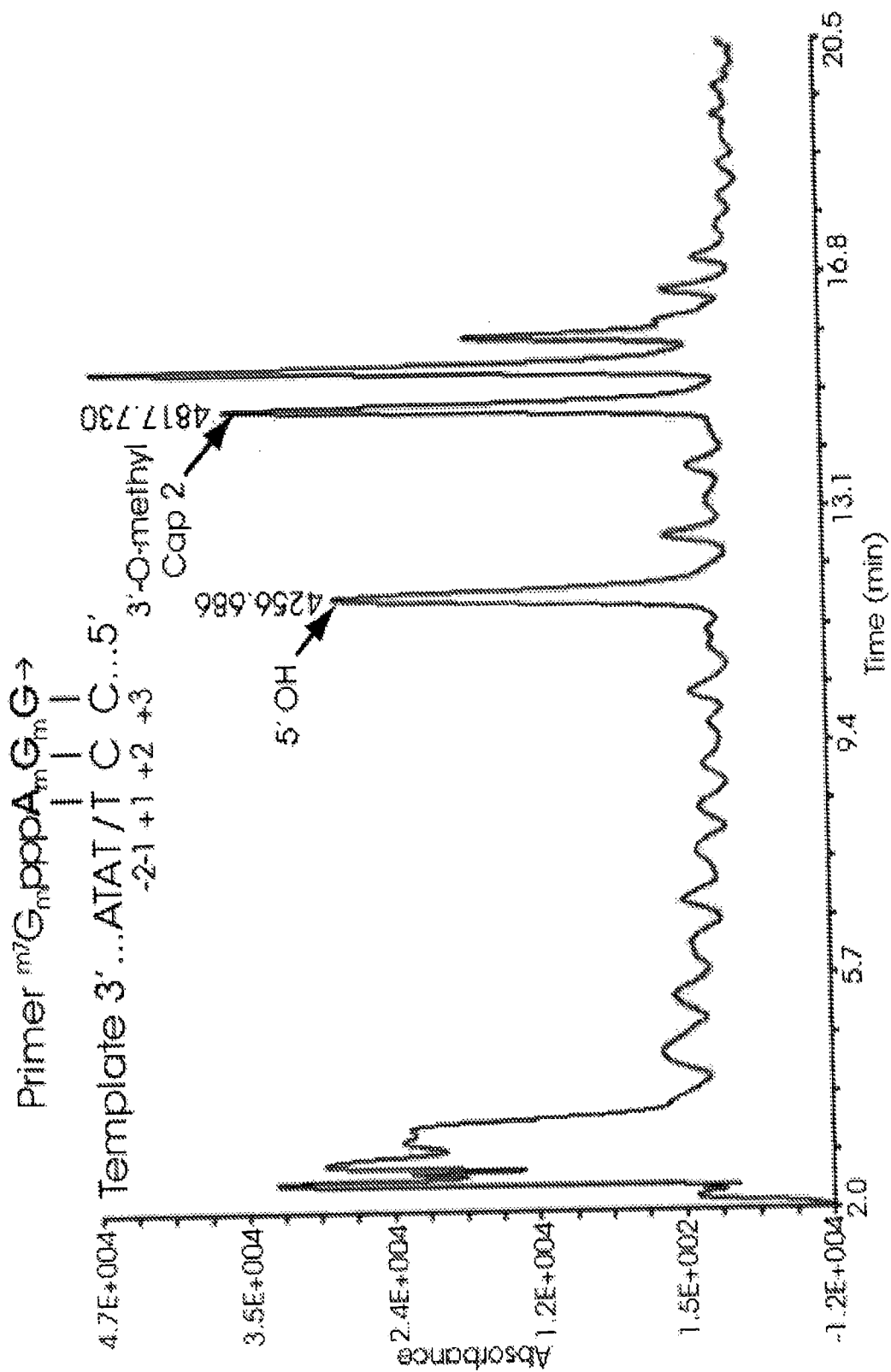


Figure 13A

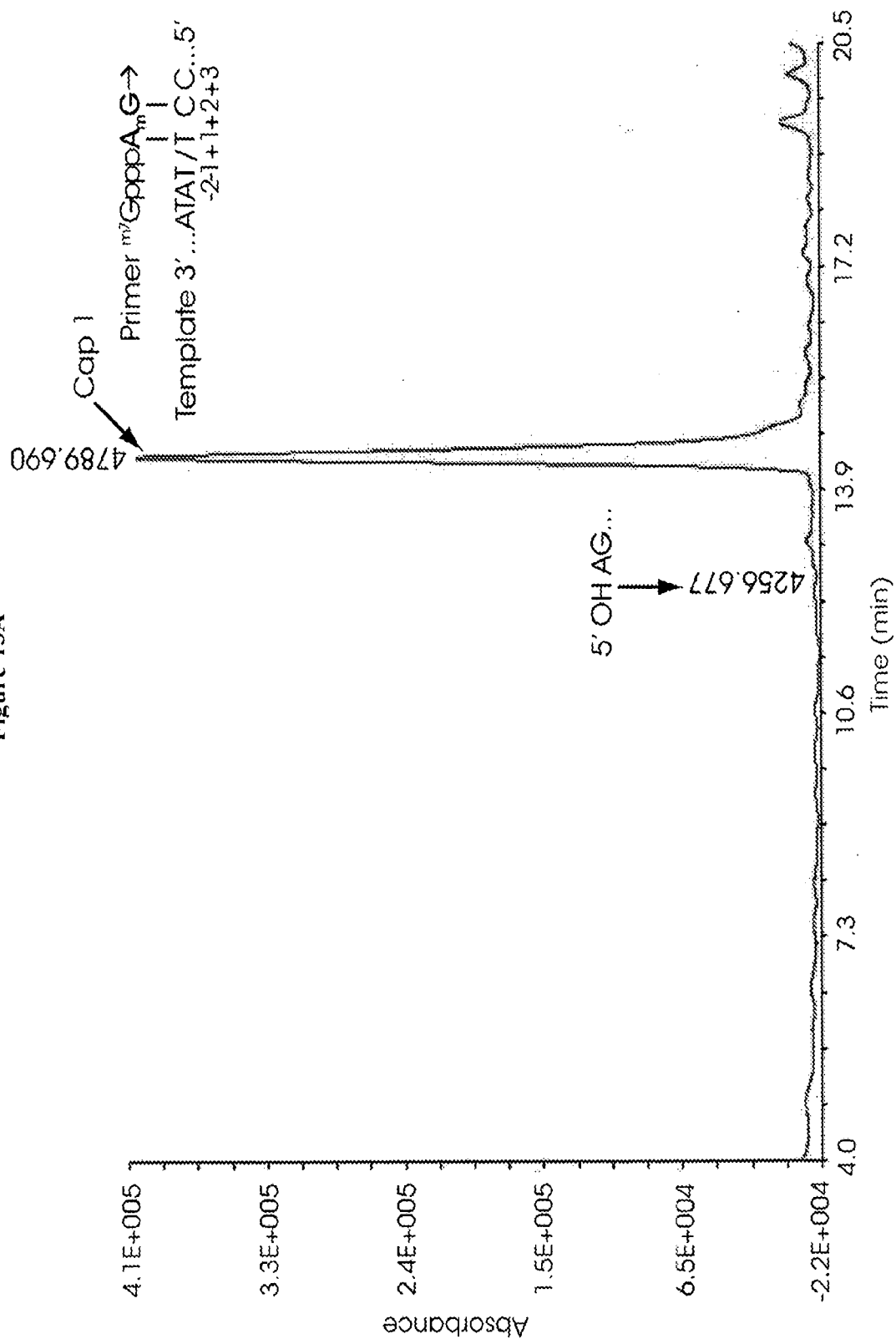


Figure 13B

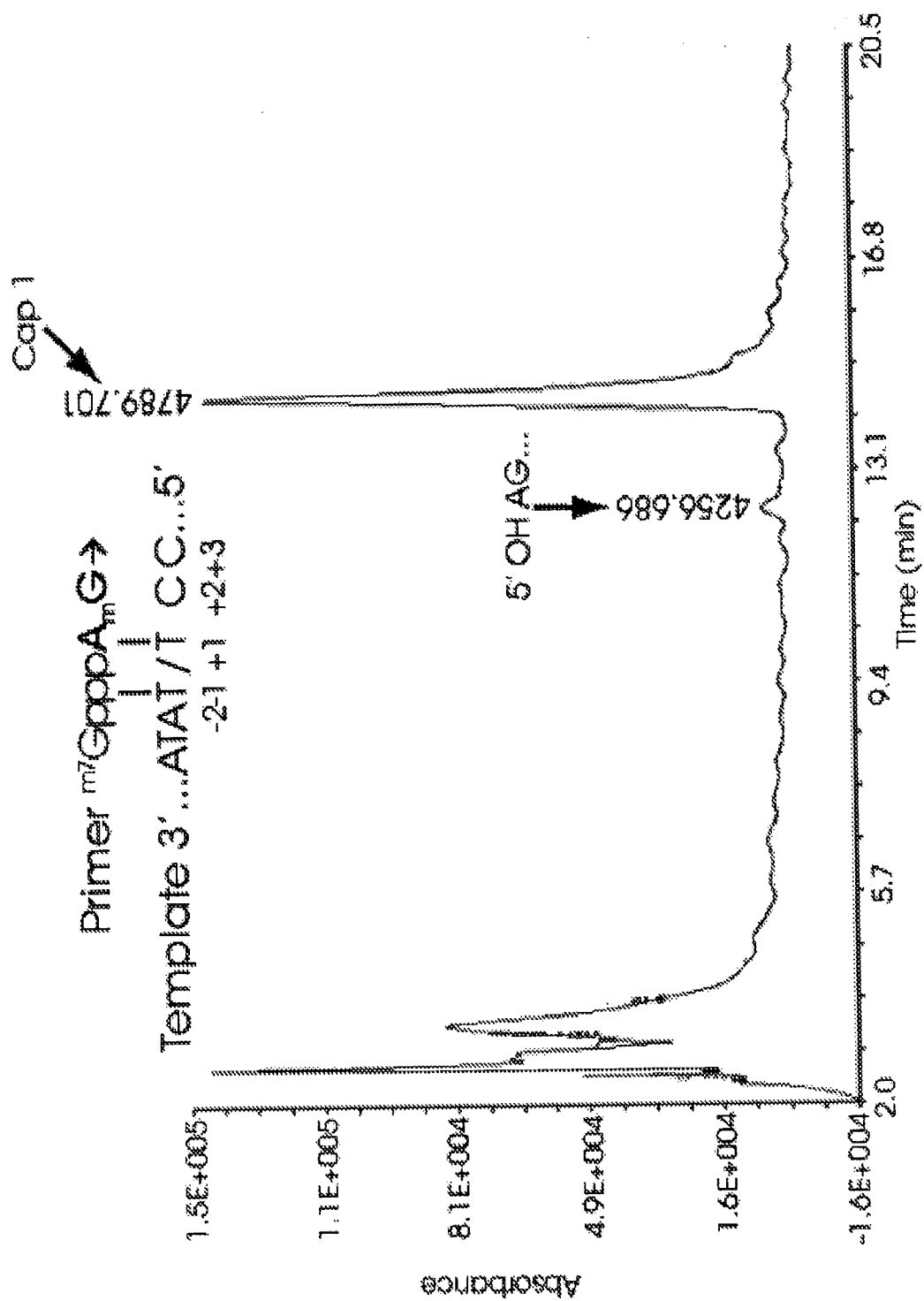


Figure 13C

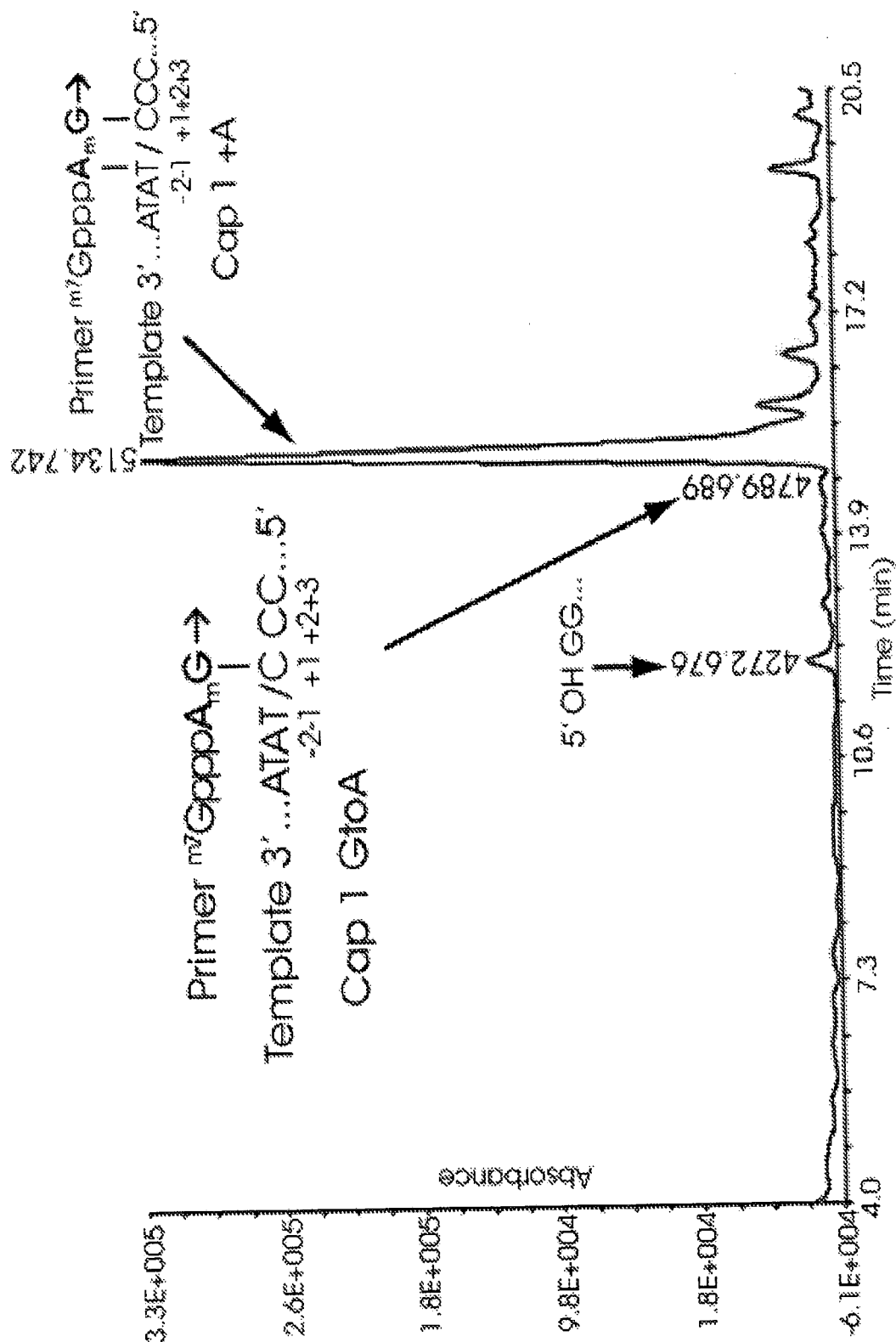


Figure 13D

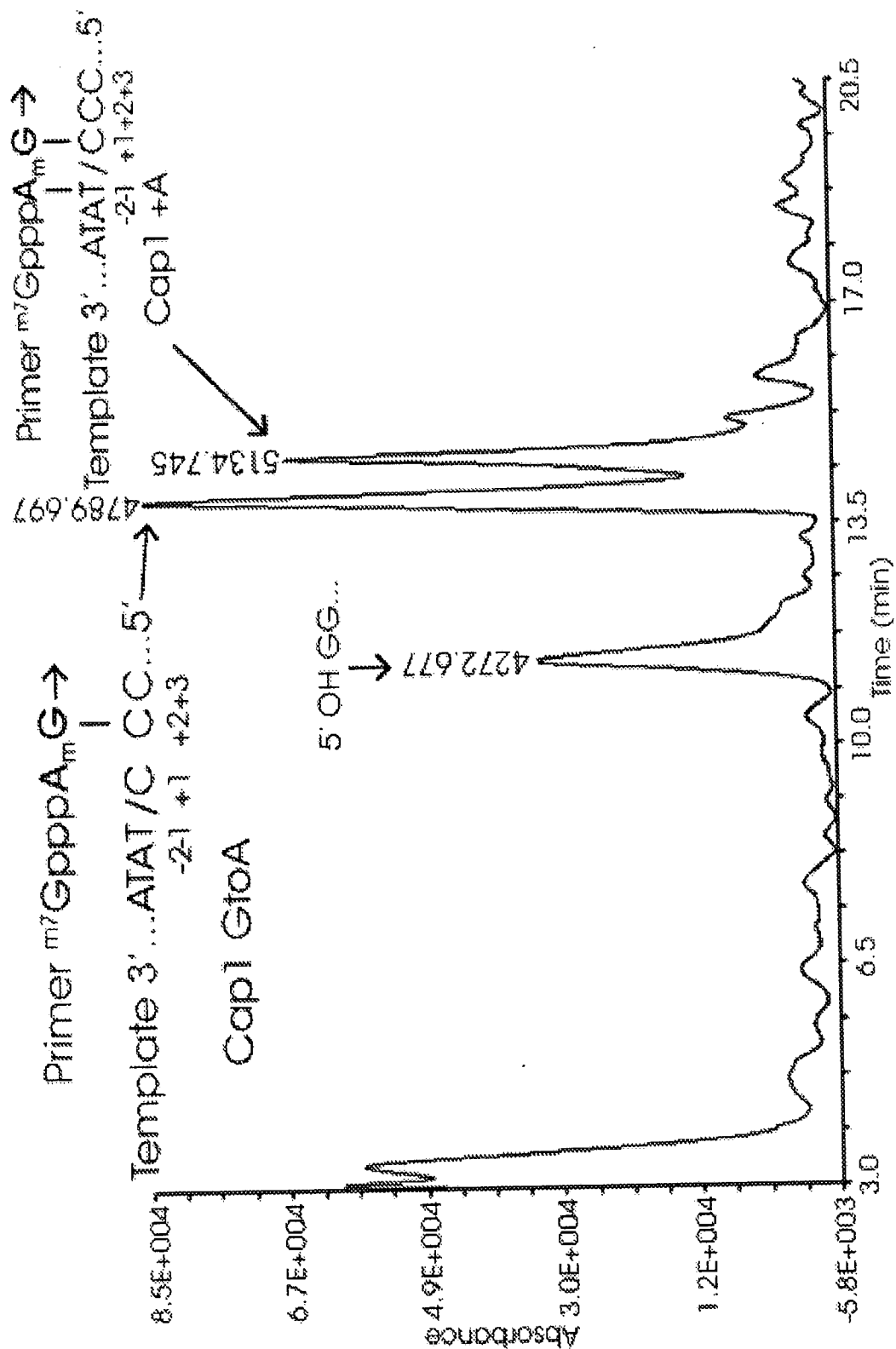
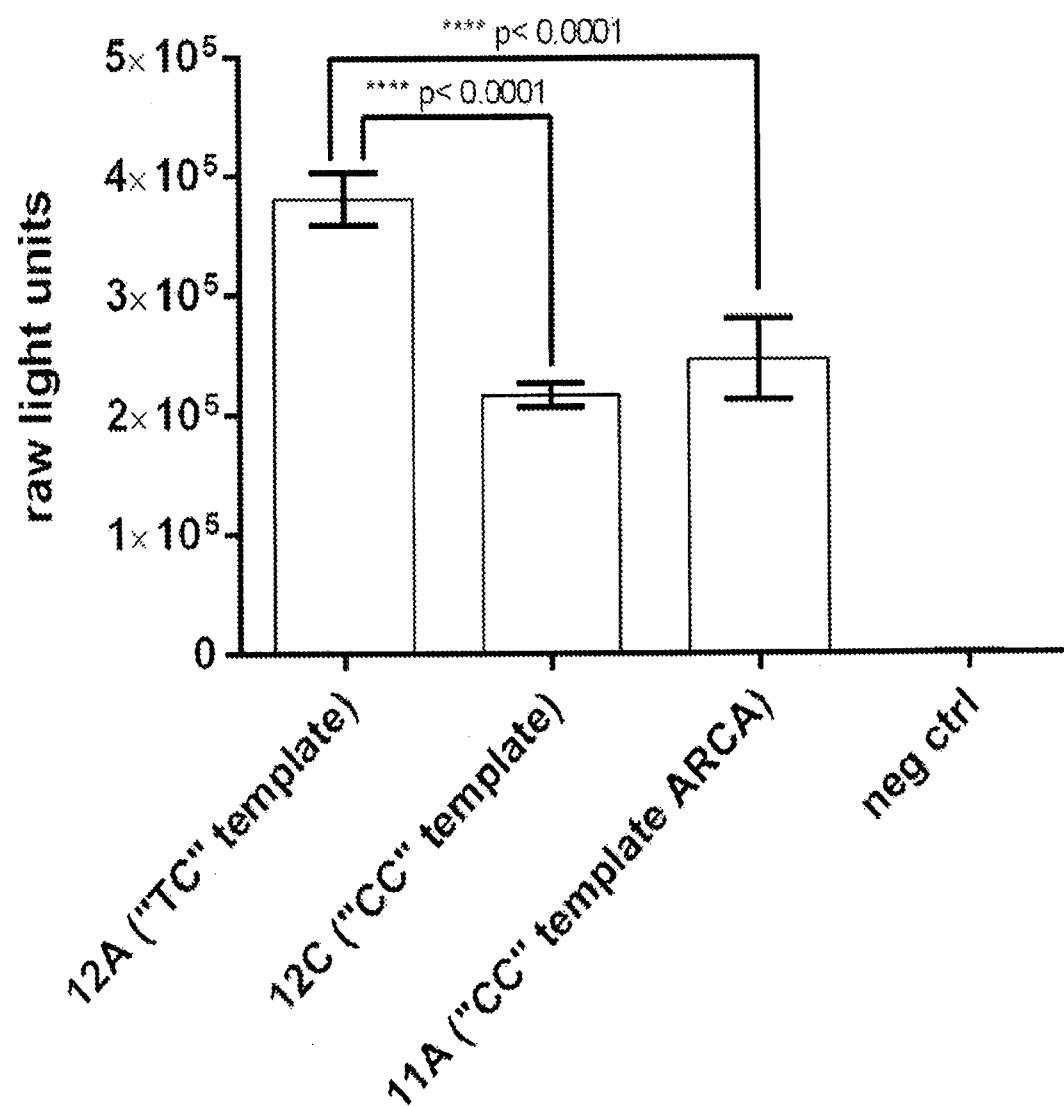
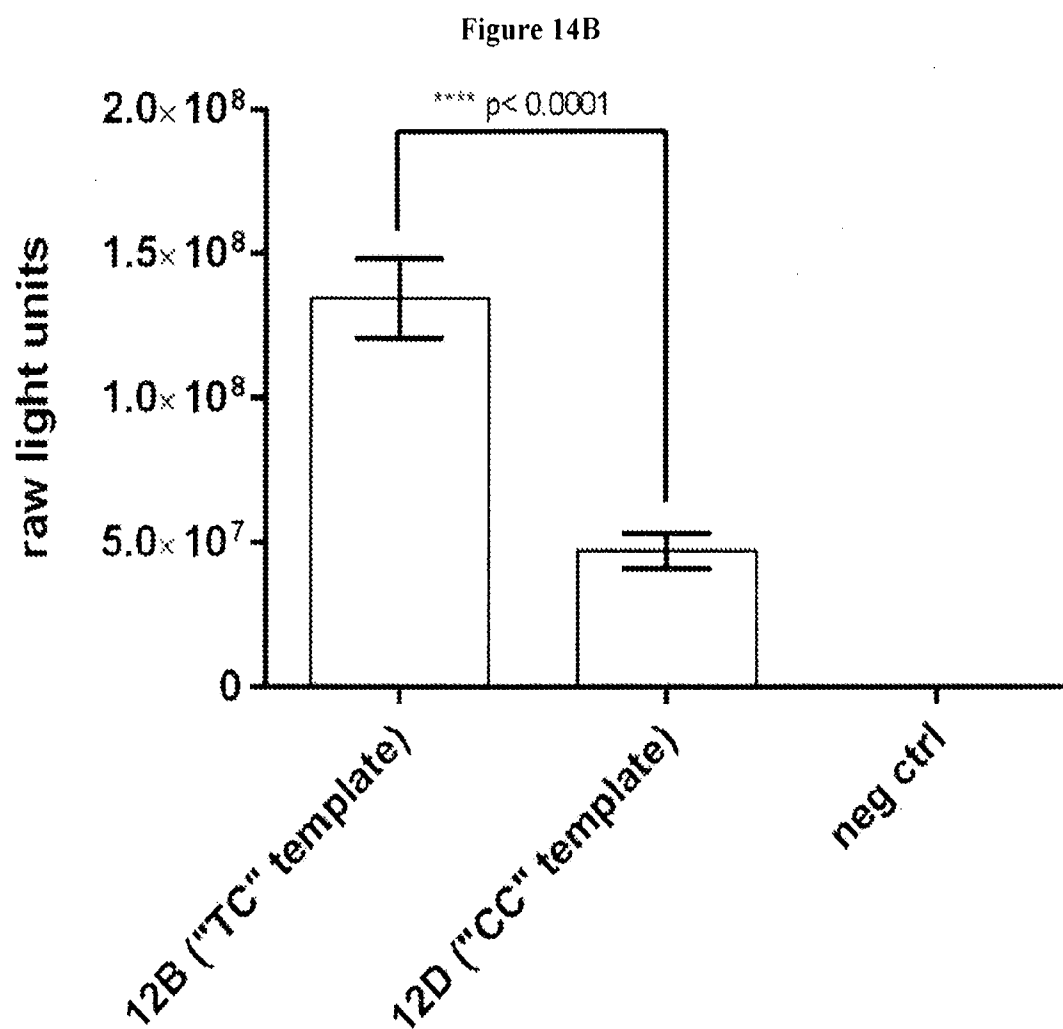


Figure 14A





INTERNATIONAL SEARCH REPORT

International application No.

PCT/US16/52670

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - A23L 33/13; A61K 31/7115, 31/712 (2016.01) CPC - A23L 33/13; A61K 31/7115, 31/712 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): A23L 33/13; A61K 31/7115, 31/712, 31/7125; C07H 21/02, 21/04 (2016.01) CPC: A23L 33/13; A61K 31/7115, 31/712, 31/7125; C07H 21/02, 21/04 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) PatSeer; Google Scholar; Pubmed; EBSco; SureChEMBL; Trilink Biotechnologies, Hogrefe, Lebedev, McCaffrey, Shin, capped oligonucleotide, m7GpppApA, m7GpppAmpG, m7GpppAmpA, m7GpppApU, m7GpppApG, m7GpppApG, dimer, trimer, tetramer, pentamer, hexamer, heptamer, octamer, 2-mer, 3-mer, 4-mer, 5-mer, 6-mer, primer, PCR, mismatch, extension, DNA, RNA, complement		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SMITH, RE et al. A Unique Class of Compound, Guanosine-Nucleoside Tetraphosphate G(5') pppp(5')N, Synthesized during the in Vitro Transcription of Cytoplasmic Polyhedrosis Virus of Bombyx mori. The Journal of Biological Chemistry, Vol. 257, No. 1, 10 January 1982, pp. 485-494; abstract; page 485, column 1, paragraphs 2-3; page 487, figure 1A; page 489, column 1, paragraph 1; page 491, column 2, paragraph 2	1-2, 3/1-2, 4, 14, 18 5-13, 15-17
Y	ALLAWI, HT et al. Mapping of RNA Accessible Sites by Extension of Random Oligonucleotide Libraries with Reverse Transcriptase. RNA, Vol. 7, 2001, pp. 314-327; abstract; page 324, column 2, paragraphs 1-2	5-6, 9, 13
Y	KOLCHINSKY, A et al. Analysis of SNPs and Other Genomic Variations Using Gel-Based Chips. Human Mutation, Vol. 19, 2002, pp. 343-360; page 348, figures 1A-B	7
Y	SOKOLOV, BP et al. A Rapid and Simple PCR-based Method for Isolation of cDNAs from Differentially Expressed Genes. Nucleic Acids Research, Vol. 22, No. 19, 1994, pp. 4009-4015; abstract; page 4012, figure 7	8
Y	EP 0081099 A2 (SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH) 15 June 1983; page 4, lines 10-25; page 12, lines 15-25; claims 1, 9, 13	10-12, 15, 17
Y	GEHBARD, LG et al. Functional RNA Elements in the Dengue Virus Genome. Viruses, Vol. 3, 2011, pp. 1739-1756; page 1740, paragraph 4	16
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* 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 20 December 2016 (20.12.2016)		Date of mailing of the international search report 27 JAN 2017
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. 571-273-8300		Authorized officer Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774



(12)发明专利申请

(10)申请公布号 CN 108366604 A

(43)申请公布日 2018.08.03

(21)申请号 201680067458.6

(74)专利代理机构 北京纪凯知识产权代理有限公司 11245

(22)申请日 2016.09.20

代理人 王永伟 赵蓉民

(30)优先权数据

62/221,248 2015.09.21 US

(51)Int.Cl.

A23L 33/13(2006.01)

(85)PCT国际申请进入国家阶段日

A61K 31/7115(2006.01)

2018.05.18

A61K 31/712(2006.01)

(86)PCT国际申请的申请数据

PCT/US2016/052670 2016.09.20

(87)PCT国际申请的公布数据

W02017/053297 EN 2017.03.30

(71)申请人 垂林克生物技术公司

地址 美国加利福尼亚

(72)发明人 R·I·霍格雷夫 A·列别捷夫

A·P·麦卡弗里 D·施恩

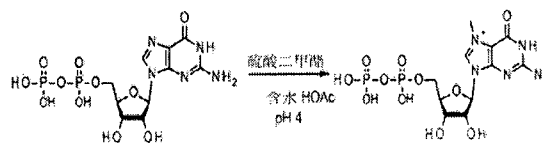
权利要求书4页 说明书33页 附图24页

(54)发明名称

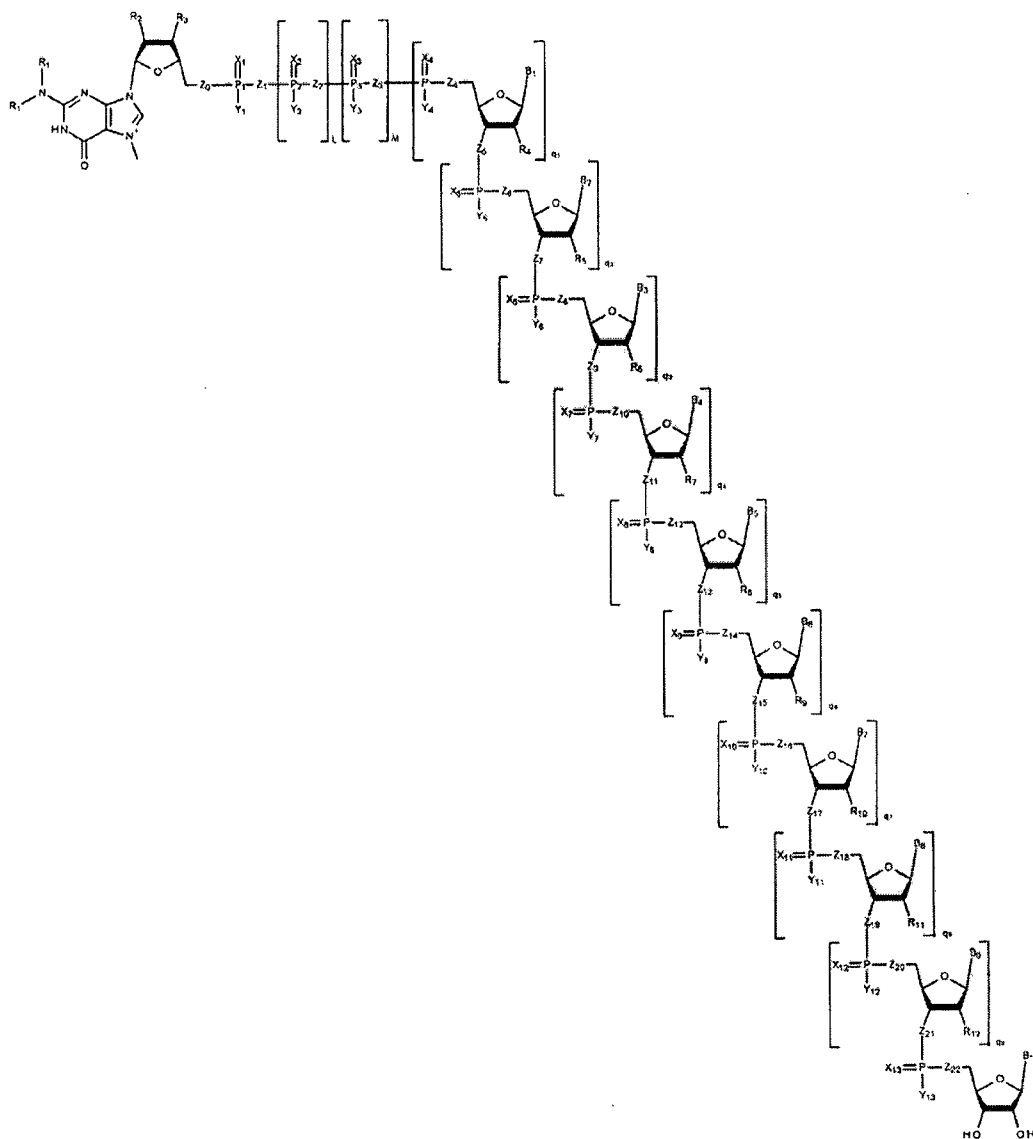
用于合成5'-加帽RNA的组合物和方法

(57)摘要

本文提供用于合成5'加帽RNA的方法和组合物,其中起始加帽寡核苷酸引物具有通式形式 $m^7\text{Gppp}[\text{N}_{2,6}\text{me}]_n[\text{N}]_m$,其中 $m^7\text{G}$ 为N7-甲基化的鸟苷或任何鸟苷类似物,N为任何天然的、修饰的或非天然的核苷,"n"可以是0至4的任何整数且"m"可以是1至9的整数。



1. 起始加帽寡核苷酸引物, 其包含以下结构:



其中:

B₁至B₁₀中的每一个独立地为天然的、修饰的或非天然的核苷碱基;

M为0或1;

L为0或1;

q₁为1并且q₂至q₉中的每一个独立地为0或1;

R₁为H或甲基;

R₂和R₃独立地为H、OH、烷基、O-烷基、卤素、连接体或可检测标志物;

X₁至X₁₃中的每一个独立地为O或S;

Y₁至Y₁₃中的每一个独立地为OH、SH、BH₃、芳基、烷基、O-烷基或O-芳基;

Z₁至Z₂₂中的每一个独立地为O、S、NH、CH₂、C(卤素)₂或CH(卤素); 以及

R₄至R₁₂中的每一个独立地为H、OH、OMe或可检测标志物。

2. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 = 1$, 以及

$q_2 - q_9$ 中的每一个 $= 0$ 或 1 。

3. 根据权利要求1或2所述的起始加帽寡核苷酸引物, 其中:

B_1 与核酸模板上转录模板位置+1处的核苷碱基完全互补,

B_2 至 B_9 , 如果存在, 与核酸模板上转录开始位点处从位置+2开始的对应核苷碱基完全互补, 以及

B_{10} , 起始加帽寡核苷酸引物的最后一个核苷碱基, 与最后一个杂交转录模板核苷酸完全互补。

4. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_2 - q_9 = 0$, 以及

B_{10} 与模板核苷酸2杂交。

5. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 - q_2 = 1$

$q_3 - q_9 = 0$, 以及

B_2 和 B_{10} 分别与模板核苷酸2和3杂交。

6. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 - q_3 = 1$,

$q_4 - q_9 = 0$, 以及

B_2 、 B_3 和 B_{10} 分别与模板核苷酸2、3和4杂交。

7. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 - q_4 = 1$,

$q_5 - q_9 = 0$, 以及

B_2 、 B_3 、 B_4 和 B_{10} 分别与模板核苷酸2、3、4和5杂交。

8. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 - q_5 = 1$,

$q_6 - q_9 = 0$, 以及

B_2 、 B_3 、 B_4 、 B_5 和 B_{10} 分别与模板核苷酸2、3、4、5和6杂交。

9. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 - q_6 = 1$,

$q_7 - q_9 = 0$, 以及

B_2 、 B_3 、 B_4 、 B_5 、 B_6 和 B_{10} 分别与模板核苷酸2、3、4、5、6和7杂交。

10. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 - q_7 = 1$,

$q_8 - q_9 = 0$, 以及

B_2 、 B_3 、 B_4 、 B_5 、 B_6 、 B_7 和 B_{10} 分别与模板核苷酸2、3、4、5、6、7和8杂交,。

11. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q_1 - q_8 = 1$,

$q_9 = 0$, 以及

B_2 、 B_3 、 B_4 、 B_5 、 B_6 、 B_7 、 B_8 和 B_{10} 分别与模板核苷酸2、3、4、5、6、7、8和9杂交。

12. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q1-q9=1$, 以及

B2、B3、B4、B5、B6、B7、B8、B9和B10分别与模板核苷酸2、3、4、5、6、7、8、9和10杂交。

13. 根据权利要求1所述的起始加帽寡核苷酸引物, 其中:

$q1-q2=1$,

$q3-9=0$, 以及

B2和B10分别与模板核苷酸2和3杂交。

14. RNA分子, 其包含根据权利要求1所述的起始加帽寡核苷酸引物。

15. 细胞, 其含有包含根据权利要求1所述的起始加帽寡核苷酸引物的RNA分子。

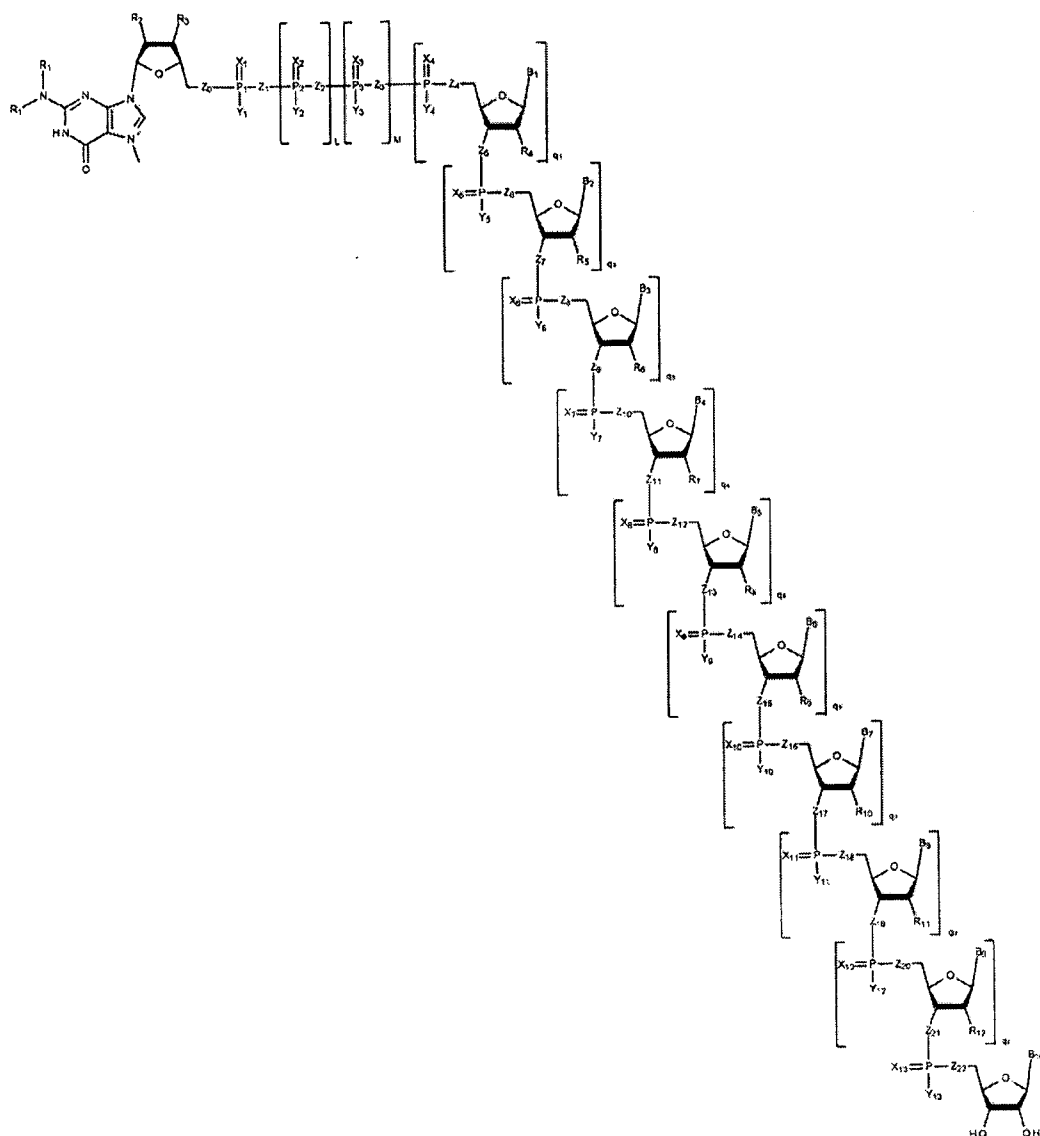
16. 细胞, 其含有由包含根据权利要求1所述的起始加帽寡核苷酸引物的RNA分子翻译的蛋白质或肽。

17. 药物组合物, 其包含RNA分子和药学上可接受的载体, 所述RNA分子包含根据权利要求1所述的起始加帽寡核苷酸引物。

18. 用于合成完全模板化的RNA分子的方法, 其包含以下步骤:

将具有以下结构的起始加帽寡核苷酸引物在有利于通过多核苷酸模板的RNA聚合酶转录的条件下引入包含所述RNA聚合酶的混合物中, 以及

温育所述混合物一段足以允许所述模板转录的时间,



其中：

B₁至B₁₀中的每一个独立地为天然的、修饰的或非天然的核苷碱基；

M为0或1；

L为0或1；

R₁为H或甲基；

R₂和R₃独立地为H、OH、烷基、O-烷基、卤素、连接体或可检测标志物；

X₁至X₁₃中的每一个独立地为O或S；

Y₁至Y₁₃中的每一个独立地为OH、SH、BH₃、芳基、烷基、O-烷基或O-芳基；

Z₁至Z₂₂中的每一个独立地为O、S、NH、CH₂、C(卤素)₂或CH(卤素)；以及

R₄至R₁₂中的每一个独立地为H、OH、OMe或可检测标志物。

用于合成5'-加帽RNA的组合物和方法

[0001] 相关申请的交叉引用

[0002] 本申请要求于2015年9月21日提交的美国临时申请号62/221,248的优先权,在此其内容其整体通过引用并入本文。

[0003] 序列表

[0004] 本申请含有已经以ASCII格式电子提交的序列表并且其整体在此通过引用并入本文。所述ASCII复本于2016年11月16日创建,其名称为095109-000500PC-1022543_SL.txt且大小为785字节。

发明领域

[0005] 本发明涉及用于合成5'加帽RNA的方法和组合物。在具体方面,本发明涉及新型含“帽”起始寡核苷酸引物,其具有天然的或修饰的5'-帽0、帽1、帽2或三甲基鸟苷-帽(TMG-帽)结构。在其他的方面,本发明涉及提供用于有效地产生其和使用其制备5'-加帽RNA的方法。

背景技术

[0006] 提供以下描述以帮助读者的理解。所提供的信息或引用的参考文献都不被承认是本发明的现有技术。

[0007] 编码用于治疗应用的生理上重要的蛋白质的信使RNA(mRNA)相对于基于DNA的质粒和病毒载体对于递送遗传物质已经显示出显著优势。这些优势中最重要的是:

[0008] (i) 高的安全水平(降低来自病毒或质粒整合的潜在的基因组损害),

[0009] (ii) mRNA递送导致立即蛋白质表达(不像一般伴随质粒发生的延迟响应),

[0010] (iii) mRNA允许对蛋白质表达的稳健剂量依赖性控制,以及

[0011] (iv) 与制造质粒和病毒载体相比,大规模合成mRNA的简单性。

[0012] 信使RNA可以编码实质上任何已知的蛋白质并可以通过本领域技术人员已知的多种方法递送至特定的组织和器官。一旦被递送,这些mRNA指导目标组织内的核糖体蛋白质表达,导致按照mRNA分子生产数百种蛋白质。

[0013] 存在于每个活性mRNA分子中的几个结构元件被利用以有效地翻译编码的蛋白质。这些元件的一个是在mRNA的5'-末端上的帽结构,其存在于所有真核生物(和一些病毒)中。天然存在的帽结构包含在鸟嘌呤碱基的N7位甲基化的核糖鸟苷残基。此7-甲基鸟苷(7^mG)通过5'-至5'-三磷酸链连接在mRNA分子的5'-末端贯穿本申请,7^m和m7可互换使用,具有同等含义。5'-末端上7^mGppp片段的存在对于mRNA成熟至关重要,其:

[0014] 保护mRNA免于被核酸外切酶降解,

[0015] 促进mRNA从核转运至细胞质,以及

[0016] 在翻译起始复合体的装配中起关键作用(Cell 9:645-653, (1976); Nature 266:235, (1977); Federation of Experimental Biologists Society Letter 96:1-11, (1978); Cell 40:223-24, (1985); Prog. Nuc. Acid Res. 35:173-207, (1988);

Ann.Rev.Biochem.68:913-963, (1999); J.Biol.Chem.274:30337-3040, (1999))。

[0017] 只有那些携带帽结构的mRNA在帽依赖性翻译中具有活性;mRNA的“去帽(decapitation)”导致其对于蛋白质合成的模板活性几乎完全失去(Nature,255:33-37, (1975); J.Biol.Chem.,vol.253:5228-5231, (1978); and Proc.Natl.Acad.Sci.USA,72:1189-1193, (1975))。

[0018] 真核mRNA的另一个元件是在转录物位置1(帽1)处存在2'-O-甲基核苷残基,以及在某些情况下在转录物位置1和2(帽2)处存在2'-O-甲基核苷残基。mRNA的2'-O-甲基化对于在体内mRNA翻译的更高效力而言是需要的(Proc.Natl.Acad.Sci.USA,77:3952-3956 (1980))并且进一步提高5'-加帽mRNA的核酸酶稳定性。具有帽1(和帽2)的mRNA为独特的标志,其允许细胞识别真正的mRNA 5'末端,并且在一些情况下,区别于源自感染性遗传元件的转录物(Nucleic Acid Research 43:482-492(2015))。

[0019] 初级(primary)mRNA转录物携带在体内始于NTP(一般GTP)的RNA合成起始(起动, initiation)导致的5'-三磷酸基团(5'-pppmRNA)。通过几个酶促步骤发生mRNA转录物的5'-三磷酸化末端向帽结构(帽0)的转化(J.Biol.Chem.250:9322, (1975); J.Biol.Chem.271:11936, (1996); J.Biol.Chem.267:16430, (1992))。这些酶促反应步骤包括:

[0020] 步骤1:RNA三磷酸酶使mRNA的5'-三磷酸基转化为5'-二磷酸基, $\text{pppN}_i(\text{pN})_x \rightarrow \text{ppN}_i(\text{pN})_x + \text{无机磷酸根}$;

[0021] 步骤2:RNA鸟苷酸转移酶利用GTP将GMP残基转移至mRNA的5'-二磷酸基, $\text{ppN}_i(\text{pN})_x + \text{GTP} \rightarrow \text{G}(5')\text{ppp}(5')\text{N}_i(\text{pN})_x + \text{无机焦磷酸根}$;以及

[0022] 步骤3:鸟嘌呤-7-甲基转移酶利用S-腺苷-甲硫氨酸(AdoMet)作为辅因子并将甲基从AdoMet转移至鸟嘌呤碱基的7-氮, $\text{G}(5')\text{ppp}(5')\text{N}_i(\text{pN})_x + \text{AdoMet} \rightarrow {}^7\text{mG}(5')\text{ppp}(5')\text{N}_i(\text{pN})_x + \text{AdoHyc}$ 。

[0023] 这些酶促活性导致的RNA被称为“5'加帽RNA”或“加帽RNA”,并且导致“加帽RNA”形成的此过程中涉及的酶组合被称为“加帽酶”。加帽酶(包括这些酶的克隆形式)已经被从许多来源鉴定和纯化并且是本领域公知的(Prog.Nucleic Acid Res.Mol.Biol.66:1-40, (2001); Prog.Nucleic Acid Res.Mol.Biol.50:101-129, (1995); 和Microbiol.Rev.44:175, (1980))。通过加帽酶将帽核苷酸添加至初级RNA的5'-末端而导致的加帽RNA已经被称为具有“帽0结构”的加帽RNA(J.Biol.Chem.269:14974-14981, (1994); J.Biol.Chem.271:11936-11944, (1996))。加帽酶已经用于在体外合成具有帽0结构的加帽RNA(J.Biol.Chem.255:11588, (1980); Proc.Natl.Acad.Sci.USA 94:9573, (1997); J.Biol.Chem.267:16430, (1992); J.Biol.Chem.269:14974, (1994); 和J.Biol.Chem.271:11936, (1996))。

[0024] 具有5'-帽0结构的加帽RNA可以在体内通过(核苷-2'-O-)甲基转移酶的作用进一步转化为“Cap1”结构(J.Biol.Chem.269:14974-14981, (1994); J.Biol.Chem.271:11936-11944, (1996); 和EMBO 21:2757-2768, (2002))。例如,牛痘mRNA(核苷-2'-O)甲基转移酶可以通过以下反应催化具有帽0结构的5'-加帽RNA的5'-倒数第二个核苷酸的2'-羟基的甲基化:

[0025] ${}^7\text{mG}(5')\text{ppp}(5')\text{N}_1\text{pN}_2(\text{pN})_x + \text{AdoMet} \rightarrow {}^7\text{mG}(5')\text{ppp}(5')\text{N}_1{}^{2'-\text{m}}\text{pN}_2(\text{pN})_x + \text{AdoHyc}$ 。

[0026] 据报道,具有帽1结构的二甲基化加帽RNA比具有帽0结构的加帽RNA更有效地被翻译(Nucleic Acids Res. 26:3208, (1998))。真核细胞利用另一种(核苷-2'-O)甲基转移酶(例如人细胞中的hMTR2(Nucleic Acids Res. 39:4756 (2011)))催化5'-加帽RNA的第二个被转录核苷酸的2'-羟基的甲基化,以通过以下反应将帽1结构转化为帽2结构:

[0027] ${}^7\text{mG}(5')\text{ppp}(5')\text{N}_1^{2'-\text{OMe}}\text{pN}_2(\text{pN})_x + \text{AdoMet} \rightarrow {}^7\text{mG}(5')\text{ppp}(5')\text{N}_1^{2'-\text{OMe}}\text{pN}_2^{2'-\text{OMe}}(\text{pN})_x + \text{Adolylc}.$

[0028] 大约50%的真核mRNA具有帽2结构。

[0029] 为了生产用于各种生物学研究的长功能性RNA,在二十世纪80年代中期开发了在体外酶促合成初级RNA的方法(Methods Enzymol. 180:51-62, (1989); Nucl. Acids Res., 10:6353-6362, (1982); Meth. Enzymol., 180:42-50 (1989); Nucl. Acids Res., 12:7035-7056, (1984) 和 Nucleic Acid Reseach 15:8783-8798, (1987))。

[0030] 在体外转录之后,携带5'-三磷酸基团的原始mRNA转录物可以进一步用加帽酶加帽。然而,体外酶促5'-加帽是昂贵、费力、低效和难以控制的。

[0031] 考虑到这些缺点,开发了另一种用于在体外合成加帽mRNA的方法,其中化学合成的二核苷酸 ${}^7\text{mG}(5')\text{ppp}(5')\text{G}$ (也被称为mCAP)用于起始转录(RNA 1:957-967, (1995))。mCAP二核苷酸含有成熟mRNA的5'-帽0结构,但不具有帽1和帽2结构特有的2'-O-甲基核苷。

[0032] 然而,存在两个主要缺点归因于利用合成mCAP二核苷酸起始体外转录。第一个是mCAP和pppG对于起始mRNA合成的强竞争性。当用pppG起始mRNA时,由于存在5'-三磷酸基,所得的ppp-mRNA无翻译活性并且有免疫原性。相应地,当用mCAP起始mRNA时,所得的5'-加帽-mRNA有翻译活性并且不那么有免疫原性。

[0033] 为了提高5'-加帽与5'-无帽(或5'-三磷酸化;pppmRNA)mRNA的比率,必须使用相对于pppG过量的 ${}^7\text{mGpppG}$ (4:1至10:1)以有利于生产5'-帽结构mRNA转录物(高达80-90%)。此方式的消极方面是在转录期间GTP供给的快速消耗和对大量可能昂贵的合成mCAP二聚体的需求导致mRNA的总产率显著降低。转录后,为了降低合成的mRNA的免疫原性,需要用碱性磷酸酶另外处理含有5'-加帽mRNA和5'-pppmRNA两者的粗混合物,以从pppmRNA去除无帽5'-三磷酸基团。磷酸酶处理后获得的无帽5'-OH形式的mRNA是无活性的并且不参与在翻译过程中。

[0034] 当使用非对称的mCAP二核苷酸时,可出现另一个缺点,双向起始。存在 ${}^7\text{mGpppG}$ 的G或 ${}^7\text{mG}$ 部分的3'-羟基以几乎相等的概率充当转录延伸的起始点的倾向。根据转录反应的条件,其一般导致 ${}^7\text{mG}(5')\text{pppG}(\text{pN})_n$ 和 $\text{G}(5')\text{ppp}{}^7\text{mG}(\text{pN})_n$ 形式的两种异构RNA以大约相等的比例合成(RNA 1:957-967, (1995))。

[0035] 为了消除用mCAP二核苷酸进行的mRNA合成的双向起始,开发了新型的修饰的mCAP类似物,其中 ${}^7\text{mG}$ 残基的3'-OH基团被OCH₃("OMe")取代: ${}^7\text{mG}(3'-\text{O-Me})\text{pppG}$ (也被称作抗反向帽类似物(Anti-Reverse Cap Analog)(ARCA))。ARCA仅以恰当的正向方向起始mRNA合成(RNA 7:1486-1495 (2001))。本领域已知几种类型的ARCA类似物(参见例如美国专利号7,074,596)。然而,仍然需要相对于pppG的大幅摩尔过量的ARCA以确保大部分mRNA转录物分子具有5'-帽结构。进一步缺点是具有帽1结构的mRNA不能用 ${}^7\text{mGpppG}^{2'-\text{OMe}}$ 帽二聚体(RNA 1:957, (1995))或其ARCA类似物合成。

[0036] 目前,生产含有帽1结构的活性长mRNA的已知途径由下列组成:5'-三磷酸化mRNA

转录物的酶促加帽和酶促2'-O-甲基化或mCAP-加帽型或ARCA-加帽型mRNA前体的酶促2'-O-甲基化 (Nucleosides, Nucleotides, and Nucleic Acids, 25:337-340, (2006) 和 Nucleosides, Nucleotides, and Nucleic Acids 25(3):307-14, (2006))。两种方式均是相当费力、难以控制的,并且即使有了实质上的优化,两种方式都不能保证加帽并且甲基化的mRNA前体的高产率 (J. Gen. Virol., 91:112-121, (2010))。而且,用于制备具有帽2结构的mRNA的方法甚至更难并且结果是不可预测的。用于将帽1转化成帽2的酶不是当前可商业获得的。

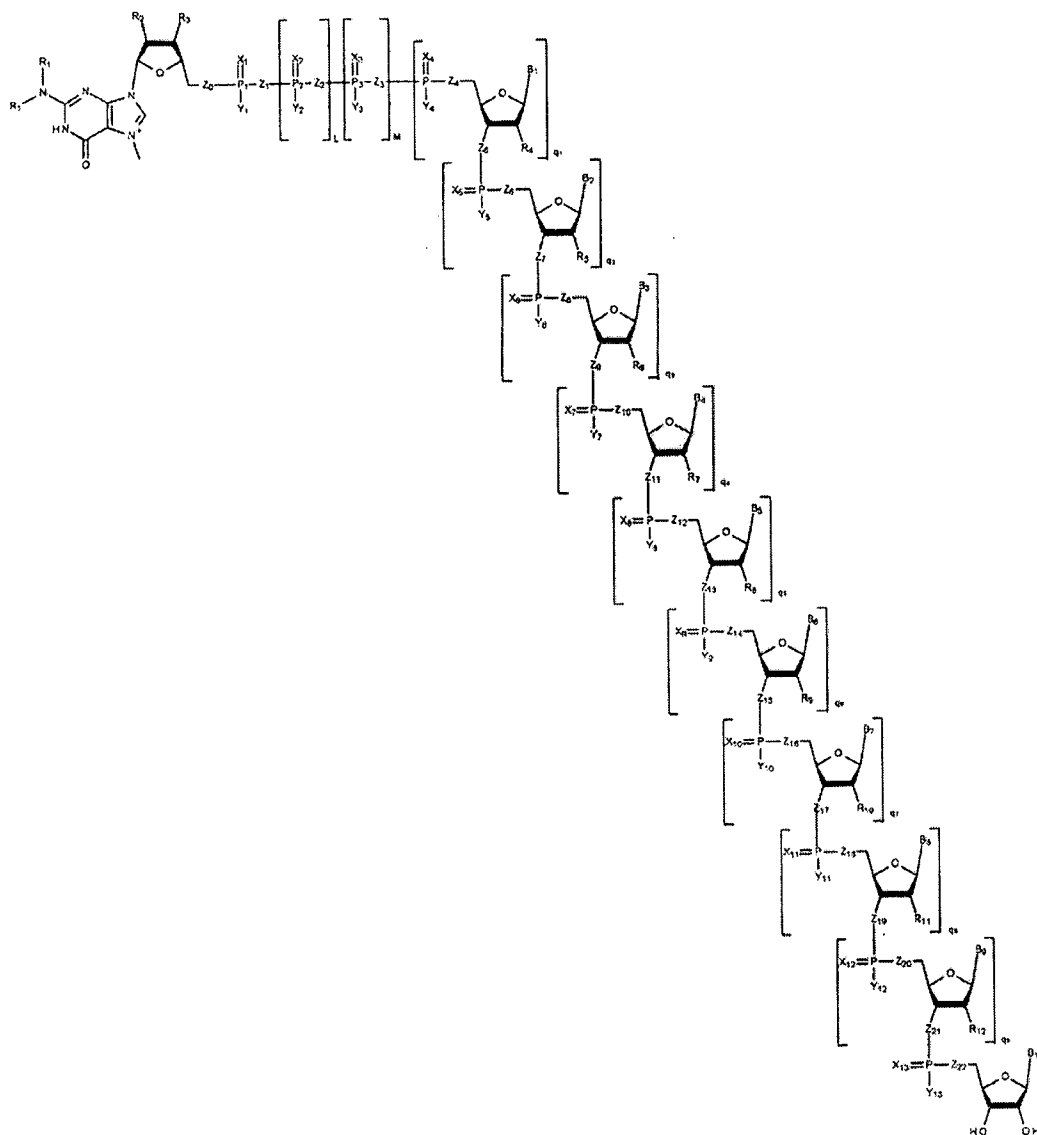
[0037] 在体外合成mRNA的另一个显著复杂性,尤其在大规模制造中,是需要从所有无活性的和在一些情况下有免疫原性的无帽mRNA形式分离和纯化带帽的活性mRNA分子。不幸的是,这些方法不是微不足道的并且通常需要合成具有缀合亲和标签部分的修饰的mCAP类似物,该缀合亲和标签部分允许更容易地分离和纯化加帽RNA转录物。合成具有亲和标签作为报道部分/亲和部分的mCAP类似物的方法和用于从转录反应混合物分离加帽RNA的新方案是本领域已知的 (参见例如美国专利8,344,118)。虽然这些方式是有效的,但它们需要使用更昂贵的mCAP类似物并且它们仅允许制备和分离含有帽0结构的mRNA。

[0038] 在体外合成天然的和修饰的RNA用于多种应用中,包括核酶、反义、生物物理和生物化学研究。另外,加帽mRNA转录物用于需要蛋白质合成的应用,如体内表达实验 (使用微注射、转染和感染)、体外翻译实验和试验以及在治疗学、诊断学、疫苗开发、标记和检测中的各种应用。

[0039] 因此,在工业中存在对如下允许大规模合成mRNA的组合物和方法的需要:(a) 比常规方法少费力,(b) 消除或减少转录期间的双向起始,(c) 导致mRNA的产率更高,(d) 与当前的方法相比成本减少,(e) 减少具有不同5'-序列的异种产物的产生,以及(f) 不需要另外的酶促反应以将帽1和帽2结构并入到合成的mRNA内。还存在合成含有修饰的和/或非天然的核苷的、携带特定的修饰和/或亲和标签的各种mRNA的需要,该特定的修饰和/或亲和标签诸如荧光染料、放射性同位素、质量标签和/或分子结合对的一个伙伴 (partner) 如生物素——在分子的5'末端或在分子的5'末端附近。

发明内容

[0040] 本文提供了用于合成5'-加帽RNA的方法和组合物。在本发明的一方面,起始加帽寡核苷酸引物包含式I的通式结构:



[0041]

[0042] 其中

[0043] B₁至B₁₀中的每一个独立地为天然的、修饰的或非天然的核苷碱基；

[0044] M为0或1；

[0045] L为0或1；

[0046] q₁为1；[0047] q₂至q₆中的每一个独立地为0或1；[0048] R₁为H或甲基；[0049] R₂和R₃独立地为H、OH、烷基、O-烷基、卤素、胺、叠氮、连接体或可检测标志物；[0050] X₁至X₁₃中的每一个独立地为O或S；[0051] Y₁至Y₁₃中的每一个独立地为OH、SH、BH₃、芳基、烷基、O-烷基或O-芳基；[0052] Z₁至Z₂₂中的每一个独立地为O、S、NH、CH₂、C(卤素)₂或CH(卤素)；以及[0053] R₄至R₁₂中的每一个独立地为H、OH、OMe、连接体或可检测标志物。

[0054] 在本发明的另一方面，提供包含式I的含帽起始寡核苷酸引物的RNA分子、包含这

种RNA的药物组合物、含有这种RNA的细胞和含有从这种RNA翻译的蛋白质或肽的细胞。

[0055] 在本发明的又另一方面,提供下列方法:在有利于通过多核苷酸模板的RNA聚合酶转录的条件下,将包含式I的含帽起始寡核苷酸引物的RNA分子合成到包含多核苷酸模板和RNA聚合酶的混合物中,并且随后温育得到的混合物足以允许所述模板转录的时间。

附图说明

[0056] 图1示出了从鸟苷5'-二磷酸制备7-甲基鸟苷5'-二磷酸(pp^{7m}G)的示例性方法;

[0057] 图2示出了从pp^{7m}G制备7-甲基鸟苷5'-二磷酸咪唑(imidazolidine)(Im-pp^{7m}G)的示例性方法;

[0058] 图3示出了从3'-O-甲基鸟苷制备3'-O-甲基鸟苷5'-磷酸(pG3' Ome)的示例性方法;

[0059] 图4示出了从pG3' Ome制备3'-O-甲基鸟苷5'-磷酸咪唑(phosphorimidazolidine)(Im-pG3' Ome)的示例性方法;

[0060] 图5示出了从Im-pG3' Ome制备3'-O-甲基鸟苷5'-二磷酸(ppG3' Ome)的示例性方法;

[0061] 图6示出了用于从ppG3' Ome制备7-甲基-3'-O-甲基鸟苷5'-二磷酸(pp^{7m}G3' Ome)的示例性方法;

[0062] 图7示出了用于从pp^{7m}G3' Ome制备7-甲基-3'-O-甲基鸟苷5'-二磷酸咪唑(Im-pp^{7m}G3' Ome)的示例性方法;

[0063] 图8示出了用于制备pN_{2'-m}ppN寡核苷酸(R₁=H或Me)的一般程序;

[0064] 图9示出了用于合成具有帽0、帽1或帽2结构的起始寡核苷酸的一般程序;

[0065] 图10A示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为鸟嘌呤;M为0;L为1;q₁至q₉为0;R₁为H;R₂为H;R₃为O-甲基;X₁为0;X₂为0;X₁₃为0;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₂₂为0;

[0066] 图10B示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为鸟嘌呤;B₁₀为鸟嘌呤;M为0;L为1;q₁为1;q₂至q₉为0;R₁为H;R₂为H;R₃为H;X₁为0;X₂为0;X₄为0;X₁₃为0;Y₁为OH;Y₂为OH;Y₄为OH;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₄为0;Z₅为0;Z₂₂为0;R₄为O-甲基;

[0067] 图10C示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为鸟嘌呤;B₁₀为鸟嘌呤;M为0;L为1;q₁为1;q₂至q₉为0;R₁为H;R₂为H;R₃为O-甲基;X₁为0;X₂为0;X₄为0;X₁₃为0;Y₁为OH;Y₂为OH;Y₄为OH;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₄为0;Z₅为0;Z₂₂为0;R₄为O-甲基;

[0068] 图10D示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为腺嘌呤;B₁₀为鸟嘌呤;M为0;L为1;q₁为1;q₂至q₉为0;R₁为H;R₂为H;R₃为H;X₁为0;X₂为0;X₄为0;X₁₃为0;Y₁为OH;Y₂为OH;Y₄为OH;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₄为0;Z₅为0;Z₂₂为0;R₄为O-甲基;

[0069] 图10E示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为腺嘌呤;B₁₀为鸟嘌呤;M为0;L为1;q₁为1;q₂至q₉为0;R₁为H;R₂为H;R₃为O-甲基;X₁为0;X₂为0;X₄为0;X₁₃为0;Y₁为OH;Y₂为OH;Y₄为OH;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₄为0;Z₅为0;Z₂₂为0;R₄为O-甲基;

[0070] 图10F示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为

胞嘧啶;B₁₀为鸟嘌呤;M为0;L为1;q₁为1;q₂至q₉为0;R₁为H;R₂为H;R₃为H;X₁为0;X₂为0;X₃为0;X₄为0;X₁₃为0;Y₁为OH;Y₂为OH;Y₄为OH;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₄为0;Z₅为0;Z₂₂为0;R₄为0-甲基;

[0071] 图10G示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为胞嘧啶;B₁₀为鸟嘌呤;M为0;L为1;q₁为1;q₂至q₉为0;R₁为H;R₂为H;R₃为0-甲基;X₁为0;X₂为0;X₄为0;X₁₃为0;Y₁为OH;Y₂为OH;Y₄为OH;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₄为0;Z₅为0;Z₂₂为0;R₄为0-甲基;

[0072] 图10H示出了根据式I的实例中采用的起始加帽寡核苷酸引物的结构,其中:B₁为腺嘌呤;B₂为鸟嘌呤;B₁₀为鸟嘌呤;M为0;L为1;q₁为1;q₂为1;q₃至q₉为0;R₁为H;R₂为H;R₃为0-甲基;X₁为0;X₂为0;X₃为0;X₅为0;X₁₃为0;Y₁为OH;Y₂为OH;Y₄为OH;Y₅为OH;Y₁₃为OH;Z₀为0;Z₁为0;Z₂为0;Z₄为0;Z₅为0;Z₆为0;Z₂₂为0;R₄为0-甲基;R₅为0-甲基;

[0073] 图11示出了mRNA共转录地加帽的mRNA的荧光素酶活性;

[0074] 图12A-12H共同示出了用12A) ARCA、12B) m⁷GpppG_{2'}-Ome-pG、12C) m⁷G_{3'}-Ome-PPP-G_{2'}-Ome-pG、12D) m⁷GpppA_{2'}-Ome-pG、12E) m⁷G_{2'}-Ome-pppA_{2'}-Ome-pG、12F) m⁷GpppC_{2'}-Ome-pG、12G) m⁷G_{3'}-Ome-pppC_{2'}-Ome-pG、12H) m⁷G_{3'}-Ome-PppA_{2'}-Ome-PG_{2'}-Ome-pG共转录地加帽的mRNA的加帽效率。图11B-11G中的加帽效率等于或显著超过在图11A中观察到的加帽效率;以及

[0075] 图13A-13D共同示出了转录模板上用m⁷GpppA_{2'}-Ome-pG共转录地加帽mRNA的加帽效率和起始保真度,其中图13A示例了利用引物/NTP配制物(配方, formulation)2、在模板位置+1和+2处2'-脱氧胸苷和2'-脱氧胞苷残基的应用;图13B示例了利用引物/NTP配制物3、在模板位置+1和+2处2'-脱氧胸苷和2'-脱氧胞苷残基的应用;图13C示例了利用引物/NTP配制物2、在模板位置+1和+2处2'-脱氧胞苷残基的应用;以及图13D示例了利用引物/NTP配制物3在模板位置+1和+2处2'-脱氧胞苷残基的应用;

[0076] 图14A和14B共同示出了在模板位置+1和+2处具有2'-脱氧胸苷和2'-脱氧胞苷残基的转录模板上vs在模板位置+1和+2处具有胞苷残基的转录模板上用m⁷GpppA_{2'}-Ome-pG起始加帽寡核苷酸制成的mRNA在分化THP-1细胞中的翻译的比较。

具体实施方式

[0077] 除非另外定义,本文所使用的所有术语具有与该发明所属领域技术人员通常理解的相同的含义。整个公开所引用的所有专利、专利申请和出版物其整体通过引用并入本文。如果本文的术语存在多个定义,以本节中的那些定义为准。

[0078] 如本文结合数值所使用的,术语“大约”或“约”意为指示值加或减30%,包括限定范围内的所有值,包括所述的值。

[0079] 如本文所使用的,术语“核酸”、“核苷酸序列”或“核酸序列”指代寡核苷酸、多核苷酸或其任何片段、任何核糖或脱氧核糖衍生物以及指代含有天然的和/或修饰的核苷酸残基和核苷酸间连接体的天然存在的或合成的分子。这些短语也指代天然的(例如,基因组的)或合成的来源的DNA或RNA,该DNA或RNA可以是单链的、双链的、三链的或四链的并且可以代表有义链或反义链,或指代任何DNA样或RNA样物质。参照DNA序列,“RNA等价物”由与参照DNA序列相同的线性核苷酸序列组成,除了所有或大部分出现的含氮碱基胸腺嘧啶被尿嘧啶替代,并且糖骨架由核糖代替2'-脱氧核糖组成。适于本文提供的方法和组合物的另外

的可选的核酸骨架包括但不限于硫代磷酸酯、硒代磷酸酯、烷基磷酸三酯、芳基磷酸三酯、烷基磷酸酯、芳基磷酸酯、磷硼酸酯(phosphoboronate)、吗啉代核酸(MNA)、锁定核酸(LNA)、肽核酸(PNA)。

[0080] 如本文所使用的,术语“引物”或“寡核苷酸引物”指代核糖-或脱氧核糖-或嵌合核糖/脱氧核糖-寡核苷酸,单链的,可以是天然地存在的或合成的,并且常常包括约2至约10个核苷酸,约3至约8个核苷酸或约3至约5个核苷酸之间的序列。寡核苷酸引物可以含有一个或多个修饰基团。寡核苷酸引物可以包括RNA、DNA和/或其他修饰的核苷。技术人员能够设计并制备适合于转录DNA模板序列的寡核苷酸引物。

[0081] 如本文所使用的,术语“起始加帽寡核苷酸类似物”或“起始加帽寡核苷酸引物”指代在引物的5'-端上含有帽0、帽1、帽2或TMG-帽结构的起始寡核苷酸引物。加帽引物具有未修饰的或开放的3'-OH基团,并且其可以通过将NTP并入到引物的3'端上而通过RNA聚合酶延伸。它能够在含有必要的组分(DNA模板(例如DNA质粒)、RNA聚合酶、核苷5'-三磷酸和合适的缓冲液)的转录系统中在启动子控制下起始体外转录。本文还使用的“起始引物”或“起始寡核苷酸引物”指代作为RNA聚合酶的有效底物的携带末端3'-OH基团的寡核苷酸。在某些实施方式中,起始寡核苷酸引物是RNA聚合酶的底物并且可以通过将NTP并入到引物的3'端而延伸。起始寡核苷酸引物在起始位点与DNA模板互补。

[0082] 如本文所使用的,在起始加帽寡核苷酸引物和NTP的上下文中,术语“未取代的”或“未修饰的”指代还未被修饰的起始加帽寡核苷酸引物和NTP。

[0083] 如本文所使用的,术语“修饰的起始加帽寡核苷酸引物”指代含有一个或多个另外的修饰基团的起始加帽寡核苷酸引物。

[0084] 如本文所使用的,术语“修饰基团”指代可以在一些位置处附接到起始引物的任何化学部分,该位置包括但不限于糖、核苷碱基、三磷酸桥和/或核苷酸间磷酸(例如,美国专利申请号20070281308)。起始加帽寡核苷酸引物的修饰基团可以是与转录的过程相容的任何性质的基团。

[0085] 如本文所使用的,术语“核苷酸间连接体”指代连接寡核苷酸引物或核酸的两个核苷的一个或多个键,并且可以是天然的磷酸二酯连接体或修饰的连接体。

[0086] 如本文所使用的,术语“标记”或“可检测标记”指代可以附接或以其他方式与分子缔合以便可以通过检测该标记直接地或间接地检测该分子的任何化合物或化合物组合。可检测标记可以是放射性同位素(例如,碳、磷、碘、镅、硫、氟等)、质量同位素(例如, H^2 、 C^{13} 或 N^{15})、染料或荧光团(例如,花菁、荧光素或香豆素)、半抗原(例如,生物素)或可以被直接地或间接地检测的任何其他试剂。

[0087] 如本文所使用的,术语“杂交”或“特异性杂交”指代在转录反应期间适当严格的条件下使起始加帽寡核苷酸引物退火至DNA模板的过程。与DNA的杂交通过起始加帽寡核苷酸引物进行,在某些实施方式中,该引物的长度为3-10个核苷酸——包括5'-5'反向帽结构。核酸杂交技术是本领域公知的(例如,Sambrook,et.al,Molecular Cloning:A Laboratory Manual,Second Edition,Cold Spring Harbor Press,Plainview,N.Y.(1989);Ausubel,F.M.,et.al,Current Protocols in Molecular Biology,John Wiley&Sons,Secaucus,N.J.(1994))。

[0088] 如本文所使用的,在起始加帽寡核苷酸引物和DNA模板的复合体的上下文中,术语

“互补”、“互补的”或“互补性”指代标准的沃森/克里克 (Watson/Crick) 碱基配对规则。例如, 序列“5'-A-G-T-C-3'”与序列“3'-T-C-A-G-5'”互补。某些非天然的或合成的核苷酸可以被包括在本文所描述的核酸中; 这些包括但不限于碱基和糖修饰的核苷、核苷酸和核酸, 诸如肌苷、7-脱氮鸟苷、2'-O-甲基鸟苷、2'-氟-2'-脱氧胞苷、假尿苷、锁定核酸 (LNA) 和肽核酸 (PNA)。互补性不需要是完美的; 双链体可以含有错配的碱基对、简并的或不匹配的核苷酸。本领域技术人员可以凭经验考虑到多种变量来确定双链体稳定性, 所述变量包括, 例如, 寡核苷酸的长度、碱基组成和寡核苷酸的序列、错配碱基对的发生率、离子强度、杂交缓冲液的组分和反应条件。

[0089] 互补性可以是“完全的”或“全部的”, 其中两条核酸链的所有核苷酸碱基按照公认的碱基配对规则匹配, 其可以是“部分的”, 其中仅起始加帽寡核苷酸引物的一些核苷酸碱基与DNA靶按照公认的碱基配对规则匹配, 或者其可以是“不存在的”, 其中两条核酸链的核苷酸碱基都不按照公认的碱基配对规则匹配。起始加帽寡核苷酸引物与DNA模板之间的互补性程度可以对起始加帽寡核苷酸与DNA模板之间的杂交强度和相应地反应效率具有显著影响。术语互补性也可以涉及个别核苷酸使用。例如, 可提示寡核苷酸内的一个具体核苷酸的其与另一条链内的一个核苷酸的互补性或其缺乏 (lack) (相对于或相比于起始加帽寡核苷酸引物的其余部分与DNA链之间的互补性)。

[0090] 如本文所使用的术语“完全”、“全部”或“完美”互补意为起始加帽寡核苷酸引物的每一个核苷酸碱基均按照公认的碱基配对规则与DNA靶精确地匹配。

[0091] 如本文所使用的, 术语“基本上互补”指代在严格杂交条件下杂交的两个序列。本领域技术人员将理解基本上互补的序列不需要沿着其整个长度杂交。具体地, 基本上互补的序列可以包含不与靶序列杂交的连续碱基序列, 并且可以定位在严格杂交条件下与靶序列杂交的连续碱基序列的3'或5'侧。

[0092] 如本文所使用的, 当涉及起始加帽寡核苷酸引物序列和其与DNA模板杂交的能力使用时, 术语“特异性的”是当比对起始加帽寡核苷酸引物与DNA链时, 与DNA模板的一部分具有至少50%序列同一性的序列。可以优选的更高水平的序列同一性包括至少75%、至少80%、至少85%、至少90%、至少95%、至少99%, 以及最优选100%的序列同一性。

[0093] 如本文所使用的, 术语“核苷”包括所有天然地存在的核苷, 包括在自然界中发现的所有形式的核苷碱基和呋喃糖苷。天然地存在的核苷中最常发现的碱基环是嘌呤环和嘧啶环。天然地存在的嘌呤环包括, 例如, 腺嘌呤、鸟嘌呤和N⁶-甲基腺嘌呤。天然地存在的嘧啶环包括, 例如, 胞嘧啶、胸腺嘧啶、5-甲基胞嘧啶、假尿嘧啶 (pseudouracyl)。天然地存在的核苷例如包括但不限于腺苷、鸟苷、胞苷、胸苷、尿苷、肌苷、7-甲基鸟苷或假尿苷的核糖、2'-O-甲基或2'-脱氧核糖衍生物。

[0094] 如本文所使用的, 术语“核苷类似物”、“修饰的核苷”或“核苷衍生物”包括如本文所描述的合成的核苷。核苷衍生物还包括具有或不具有保护基的具有修饰的碱基或/和糖部分的核苷, 并且包括, 例如, 2'-脱氧-2'-氟尿苷、5-氟尿苷和类似物。本文提供的化合物和方法包括这种碱基环和其合成类似物、以及非天然的杂环取代的碱基糖和无环 (acyclic) 取代的碱基糖。可以用于本发明的其他核苷衍生物包括, 例如, LNA核苷、卤素取代的嘌呤 (例如, 6-氟嘌呤)、卤素取代的嘧啶、N⁶-乙基腺嘌呤、N¹-(烷基)-胞嘧啶、5-乙基胞嘧啶和类似物 (美国专利号6,762,298)。

[0095] 如本文所使用的,术语“通用碱基”、“简并碱基”、“通用碱基类似物”和“简并碱基类似物”包括,例如,具有人工碱基的核苷类似物,在某些实施方式中,该人工碱基可被RNA聚合酶识别为天然NTP(例如,ATP、UTP、CTP和GTP)之一或其他特定NTP的替身。通用碱基或简并碱基公开在Loakes,D.,*Nucleic Acids Res.*,29:2437-2447(2001);Crey-Desbiolles,C.,et.al.,*Nucleic Acids Res.*,33:1532-1543(2005);Kincaid,K.,et.al.,*Nucleic Acids Res.*,33:2620-2628(2005);Preparata,FP,Oliver,JS,J.*Comput. Biol.*753-765(2004);以及Hill,F.,et.al.,*Proc. Natl Acad. Sci. USA*,95:4258-4263(1998))中。

[0096] 如本文所使用的,术语“修饰的NTP”指代具有结合在任何位置处的化学部分的核苷5'-三磷酸,该位置包括糖、碱基、三磷酸链或这三种位置的任何组合。这种NTP的实例可以在例如“*Nucleoside Triphosphates and Their Analogs: Chemistry, Biotechnology and Biological Applications*”,Vaghefi,M.,ed.,Taylor and Francis,Boca Raton(2005)中找到。

[0097] 如本文所使用的,术语“修饰的寡核苷酸”包括,例如,含有修饰的核苷、修饰的核苷酸间连接体或具有修饰的核苷和核苷酸间连接体的任何组合的寡核苷酸。寡核苷酸的核苷酸间连接体修饰的实例包括硫代磷酸酯、磷酸三酯和甲基膦酸酯衍生物(Stec,W.J.,et.al.,*Chem. Int. Ed. Engl.*,33:709-722(1994);Lebedev,A.V.,et.al.,E.,*Perspect. Drug Discov. Des.*,4:17-40(1996);以及Zon,et.al.,美国专利申请号20070281308)。核苷酸间连接体修饰的其他实例可以在Waldner,et.al.,*Bioorg. Med. Chem. Letters*6:2363-2366(1996)中找到。

[0098] 如本文所使用的,术语“启动子”指代指导和控制具体DNA序列(例如基因)的转录起始的dsDNA模板的区域。启动子位于DNA上的相同链和上游上(接近有义链的5'区域)。启动子一般与待转录的DNA序列紧邻(或与之部分地重叠)。启动子中的核苷酸位置相对于DNA转录开始的转录开始位点被指定(位置+1)。起始寡核苷酸引物与启动子序列的起始位点(其在某些实施方式中,在位置+1和+2处,并且在起始四聚体的情况下,在位置+1、+2和+3处)互补。

[0099] 如本文所使用的,术语“转录”或“转录反应”指代本领域已知的用于酶促制备与DNA模板互补的RNA从而生产DNA序列的多个RNA拷贝的方法。在转录反应中合成的RNA分子称为“RNA转录物”、“原始转录物”或“转录物”。涉及本文提供的组合物和方法的转录反应采用“起始加帽寡核苷酸引物”。DNA模板的转录可以是指数性的、非线性的或线性的。DNA模板可以是双链线性DNA、部分双链线性DNA、环状双链DNA、DNA质粒、PCR扩增子,与RNA聚合酶相容的修饰的核酸模板。

[0100] 如本文所使用的,术语“酰基”表示基团-C(O)R^a,其中R^a为氢、低级烷基、环烷基、杂环基、芳基、杂芳基和类似基团。

[0101] 如本文所使用的,术语“取代的酰基”表示基团-C(O)R^a,其中R^a为取代的低级烷基、取代的环烷基、取代的杂环基、取代的芳基、取代的杂芳基和类似基团。

[0102] 如本文所使用的,术语“酰氧基”表示基团-OC(O)R^b,其中R^b为氢、低级烷基、取代的低级烷基、环烷基、取代的环烷基、杂环基、取代的杂环基、芳基、取代的芳基、杂芳基、取代的杂芳基和类似基团。

[0103] 如本文所使用的,术语“烷基”指代烃的单键链,在一些实施方式中1-20个碳原子

范围内的,并且在一些实施方式中1-8个碳原子范围内的烃的单键链;实例包括包括甲基、乙基、丙基、异丙基、正丁基、仲丁基、异丁基、叔丁基、戊基、异戊基、己基、辛基、十二烷基和类似基团。

[0104] 如本文所使用的,术语“低级烷基”指代直链或支链的烃,在一些实施方式中,1-6个碳原子范围内的,并且在一些实施方式中2-5个碳原子范围内的直链或支链的烃。实例包括乙基、丙基、异丙基和类似基团。

[0105] 如本文所使用的,术语“链烯基”指代直链或支链烃基,其具有一个或多个双键,并且,除非另外指定,含有约2至约20个碳原子,并且在一些实施方式中约2至约10个范围内的碳原子,并且在一些实施方式中约2至约8个范围内的碳原子,并且在一些实施方式中约2至约6个范围内的碳原子。链烯基的实例包括乙烯基、烯丙基、1,4-丁二烯基、异丙烯基和类似基团。

[0106] 如本文所使用的,术语“链烯基芳基”指代链烯基取代的芳基,并且“取代的链烯基芳基”指代进一步带有如本文所述的一个或多个取代基的链烯基芳基。

[0107] 如本文所使用的,术语“亚链烯基(alkenylene)”指代这样的二价直链或支链烃基:具有至少一个碳-碳双键,并且一般含有2-20个碳原子,并且在一些实施方式中,2-12个范围内的碳原子,并且在一些实施方式中,2-8个范围内的碳原子,并且“取代的亚链烯基”指代进一步带有如本文所述的一个或多个取代基的亚链烯基。

[0108] 如本文所使用的,术语“亚烷基”指代这样的二价烃基:含有1-20个碳原子,并且在一些实施方式中1-15个范围内的碳原子,直链或支链,其中两个氢原子从相同的碳原子或从不同的碳原子被移除。亚烷基的实例包括但不限于亚甲基(-CH₂-)、亚乙基(-CH₂CH₂-)和类似基团。

[0109] 如本文所使用的,术语“炔基”指代这样的直链或支链烃基:具有一个或多个三键,并且含有约2-20个碳原子,并且在一些实施方式中约2-10个范围内的碳原子,并且在一些实施方式中约2-8个范围内的碳原子,并且在一些实施方式中约2-6个范围内的碳原子。炔基的实例包括乙炔基、丙炔基(炔丙基)、丁炔基和类似基团。

[0110] 如本文所使用的,术语“炔基芳基”指代炔基取代的芳基,并且“取代的炔基芳基”指代进一步带有如本文所述的一个或多个取代基的炔基芳基。

[0111] 如本文所使用的,术语“烷氧基”表示基团-OR^c,其中R^c为如定义的低级烷基、取代的低级烷基、芳基、取代的芳基、芳烷基、取代的芳烷基、杂烷基、杂芳基烷基、环烷基、取代的环烷基、环杂烷基或取代的环杂烷基。

[0112] 如本文所使用的,术语“低级烷氧基”表示基团-OR^d,其中R^d为低级烷基。

[0113] 如本文所使用的,术语“烷基芳基”指代烷基取代的芳基,并且“取代的烷基芳基”指代进一步带有如本文所述的一个或多个取代基的烷基芳基。

[0114] 如本文所使用的,术语“烷硫基(alkylthio)”指代基团-S-R^h,其中R^h为烷基。

[0115] 如本文所使用的,术语“取代的烷硫基”指代基团-S-Rⁱ,其中Rⁱ为取代的烷基。

[0116] 如本文所使用的,术语“亚炔基”指代这样的二价直链或支链烃基:具有至少一个碳-碳三键的,并且一般具有约2-12个范围内的碳原子,并且在一些实施方式中约2-8个范围内的碳原子,并且“取代的亚炔基”指代进一步带有如本文所述的一个或多个取代基的亚炔基。

[0117] 如本文所使用的,术语“酰氨基”表示基团 $-C(O)NR^JR^J$,其中 R^J 和 R^J 可以独立地为氢、低级烷基、取代的低级烷基、烷基、取代的烷基、芳基、取代的芳基、杂芳基或取代的杂芳基。

[0118] 如本文所使用的,术语“取代的酰氨基”表示基团 $-C(O)NR^kR^k$,其中 R^k 和 R^k 独立地为氢、低级烷基、取代的低级烷基、芳基、取代的芳基、杂芳基或取代的杂芳基,然而,前提是 R^k 和 R^k 中的至少一个不是氢。 R^kR^k 与氮组合可以形成任选地取代的杂环或杂芳基环。

[0119] 如本文所使用的,术语“氨基”或“胺”表示基团 $-NR^nR^n$,其中 R^n 和 R^n 可以独立地为如本文定义的氢、低级烷基、取代的低级烷基、烷基、取代的烷基、芳基、取代的芳基、杂芳基或取代的杂芳基。“二价胺”表示基团 $-NH-$ 。“取代的二价胺”表示基团 $-NR-$,其中 R 为低级烷基、取代的低级烷基、烷基、取代的烷基、芳基、取代的芳基、杂芳基或取代的杂芳基。

[0120] 如本文所使用的,术语“取代的氨基”或“取代的胺”表示基团 $-NR^pR^p$,其中 R^p 和 R^p 独立地为氢、低级烷基、取代的低级烷基、烷基、取代的烷基、芳基、取代的芳基、杂芳基、取代的杂芳基,然而,前提是 R^p 和 R^p 中的至少一个不是氢。 R^pR^p 与氮组合可以形成任选地取代的杂环或杂芳基环。

[0121] 如本文所使用的,术语“芳酰基”指代芳基羰基类,如苯甲酰基,并且“取代的芳酰基”指代进一步带有如本文所述的一个或多个取代基的芳酰基。

[0122] 如本文所使用的,单独的或组合中的术语“芳基”指代苯基、萘基或稠合芳族杂环,其任选地具有5-10个环成员并且在一些实施方式中5-6个环成员,和/或任选地被下列中的1至3个基团或取代基取代:卤素、羟基、烷氧基、烷硫基、烷基亚磺酰基、烷基磺酰基、酰氧基、芳氧基、杂芳氧基、任选地被烷基、芳基或杂芳基单取代或二取代的氨基、脒基、任选地被烷基、芳基、杂芳基或杂环基取代的尿素、任选地被烷基、芳基或杂芳基 N -单取代或 N,N -二取代的氨基磺酰基、烷基磺酰基氨基、芳基磺酰基氨基、杂芳基磺酰基氨基、烷基羰基氨基、芳基羰基氨基、杂芳基羰基氨基或类似基团。

[0123] 如本文所使用的,术语“芳氧基”表示基团 $-OAr$,其中 Ar 为芳基或取代的芳基。

[0124] 如本文所使用的,术语“碳环”指代具有由连接的碳原子构成的单环或多个稠环的饱和的、不饱和的或芳族的基团。该环(一个或多个)可以任选地是未取代的或被以下基团取代的:例如,卤素、低级烷基、烷氧基、烷硫基、乙炔($-C\equiv CH$)、氨基、酰氨基、叠氮基、羧基、羟基、芳基、芳氧基、杂环、杂芳基、取代的杂芳基、硝基($-NO_2$)、氰基($-CN$)、巯基($-SH$)、磺酰胺基($-S(O)_2NH_2$)和类似基团。

[0125] 如本文所使用的,术语“胍基”表示基团 $-N=C(NH_2)_2$,并且“取代的胍基”表示基团 $-N=C(NR_2)_2$,其中每个 R 独立地为如本文所述的H、烷基、取代的烷基、芳基或取代的芳基。

[0126] 如本文所使用的,术语“卤”或“卤素”指代所有卤素,即,氯(Cl)、氟(F)、溴(Br)和碘(I)。

[0127] 如本文所使用的,术语“杂芳基”指代这样的含有5或6个环原子的单环芳族环结构或具有8-10个原子的二环芳族基团;含有独立地选自基团O、S和N的一个或多个,并且在一些实施方式中1-4个,并且在一些实施方式中1-3个,并且在一些实施方式中1-2个杂原子,并且任选地被下列中的1-3个基团或取代基取代:卤素、羟基、烷氧基、烷硫基、烷基亚磺酰基、烷基磺酰基、酰氧基、芳氧基、杂芳氧基、任选地被烷基、芳基或杂芳基单取代或二取代的氨基、脒基、任选地被烷基、芳基、杂芳基或杂环基取代的尿素、任选地被烷基、芳基或杂

芳基N-单取代或N,N-二取代的氨基磺酰基、烷基磺酰基氨基、芳基磺酰基氨基、杂芳基磺酰基氨基、烷基羰基氨基、芳基羰基氨基、杂芳基羰基氨基或类似基团。杂芳基还旨在包括氧化的S或N,诸如亚磺酰基、磺酰基和叔环氮的N-氧化物。碳原子或氮原子是杂芳基环结构的附接点,从而保持稳定的芳族环。杂芳基的实例为邻苯二甲酰亚胺、吡啶基、哒嗪基、吡嗪基、噻唑啉基、嘌呤基、吡咯基、噻吩基、嘧啶基、吡咯基、噁唑基、噁唑基、噁吩基、异噁唑基、氧杂噻二唑基(oxathiadiazolyl)、异噻唑基、四唑基、咪唑基、三嗪基、呋喃基、苯并呋喃基、吡啶基和类似基团。取代的杂芳基含有附接在可利用的碳或氮上以产生稳定化合物的取代基。

[0128] 如本文所使用的,术语“取代的杂芳基”指代任选地被一个或多个官能团单取代或多取代的杂环,该官能团为例如,卤素、低级烷基、低级烷氧基、烷硫基、乙炔、氨基、酰氨基、羧基、羟基、芳基、芳氧基、杂环、取代的杂环、杂芳基(hetaryl)、取代的杂芳基(substituted hetaryl)、硝基、氰基、巯基、磺酰胺基和类似基团。

[0129] 如本文所使用的,术语“杂环”指代具有单环(例如,吗啉代、吡啶基或呋喃基)或多个稠环(例如,蔡基吡啶基、喹啉基、喹啉基、吡啶基(indoliziny)或苯并[b]噻吩基)并且在环内具有碳原子和至少一个杂原子(诸如N、O或S)的饱和的、不饱和的或芳族的基团,其可以任选地是未取代的或被以下基团取代的:例如,卤素、低级烷基、低级烷氧基、烷硫基、乙炔、氨基、酰氨基、羧基、羟基、芳基、芳氧基、杂环、杂芳基、取代的杂芳基、硝基、氰基、巯基、磺酰胺基和类似基团。

[0130] 如本文所使用的,术语“取代的杂环”指代被1或多个(例如,1、2或3个)选自以下基团的取代基取代的杂环:任选地取代的烷基、任选地取代的链烯基、任选地取代的炔基、卤素、羟基、烷氧基、烷硫基、烷基亚磺酰基、烷基磺酰基、酰氧基、芳基、取代的芳基、芳氧基、杂芳氧基、氨基、酰氨基、脒基、任选地被烷基、芳基、杂芳基或杂环基取代的尿素、任选地被烷基、芳基或杂芳基N-单取代或N,N-二取代的氨基磺酰基、烷基磺酰基氨基、芳基磺酰基氨基、杂芳基磺酰基氨基、烷基羰基氨基、芳基羰基氨基、杂芳基羰基氨基、酰基、羧基、杂环、取代的杂环、杂芳基、取代的杂芳基、硝基、氰基、巯基、磺酰氨基和氧代,该取代基附接在任何可利用的点处以产生稳定化合物。

[0131] 如本文所使用的,术语“烃基”指代任何这样的有机基团:其中其骨架仅包含碳和氢。因此,烃基包括烷基、环烷基、链烯基、环链烯基、炔基、芳基、烷基芳基、芳基烷基、芳基链烯基、链烯基芳基、芳基炔基、炔基芳基和类似基团。

[0132] 如本文所使用的,术语“取代的烃基”指代进一步带有一个或多个选自以下的取代基的任何上面提到的烃基:羟基、烷氧基、取代的烷氧基、烷硫基、取代的烷硫基、芳硫基(arylthio)、取代的芳硫基、氨基、烷基氨基、取代的烷基氨基、羧基、 $-C(S)SR$ 、 $-C(O)SR$ 、 $-C(S)NR_2$,其中每个R独立地为氢、烷基或取代的烷基、硝基、氰基、卤素、 $-SO_2M$ 或 $-OSO_2M$,其中M为H、Na、K、Zn、Ca,或葡甲胺(meglumine)、胍基、取代的胍基、烃基、取代的烃基、烃基羰基、取代的烃基羰基、烃氧基羰基、取代的烃氧基羰基、烃基羰氧基(hydrocarbylcarbonyloxy)、取代的烃基羰氧基、酰基、酰氧基、杂环、取代的杂环、杂芳基、取代的杂芳基、杂芳基羰基、取代的杂芳基羰基、氨基甲酰基、单烷基氨基甲酰基、二烷基氨基甲酰基、芳基氨基甲酰基、氨基甲酰基、二硫代氨基甲酰基、芳酰基、取代的芳酰基、有机磺酰基、取代的有机磺酰基、有机亚磺酰基、取代的烷基亚磺酰基、烷基磺酰基氨基、取代的

烷基磺酰基氨基、芳基磺酰基氨基、取代的芳基磺酰基氨基、磺酰氨基、砜基和类似基团,包括上述基团中的两个和或更多个通过诸如下列的连接体/间隔体(spacer)部分附接到烃基部分: -O-、-S-、-NR-——其中R为氢、烷基或取代的烷基、-C(O)-、-C(S)-、-C(=NR')-、-C(=CR'₂)-——其中R'为烷基或取代的烷基、-O-C(O)-、-O-C(O)-O-、-O-C(O)-NR- (或-NR-C(O)-O-)、-NR-C(O)-、-NR-C(O)-NR-、-S-C(O)-、-S-C(O)-O-、-S-C(O)-NR-、-O-S(O)₂-、-O-S(O)₂-O-、-O-S(O)₂-NR-、-O-S(O)-、-O-S(O)-O-、-O-S(O)-NR-、-O-NR-C(O)-、-O-NR-C(O)-O-、-O-NR-C(O)-NR-、-NR-O-C(O)-、-NR-O-C(O)-O-、-NR-O-C(O)-NR-、-O-NR-C(S)-、-O-NR-C(S)-O-、-O-NR-C(S)-NR-、-NR-O-C(S)-、-NR-O-C(S)-O-、-NR-O-C(S)-NR-、-O-C(S)-、-O-C(S)-O-、-O-C(S)-NR- (或-NR-C(S)-O-)、-NR-C(S)-、-NR-C(S)-NR-、-S-S(O)₂-、-S-S(O)₂-O-、-S-S(O)₂-NR-、-NR-O-S(O)-、-NR-O-S(O)-O-、-NR-O-S(O)-NR-、-NR-O-S(O)₂-、-NR-O-S(O)₂-O-、-NR-O-S(O)₂-NR-、-O-NR-S(O)-、-O-NR-S(O)-O-、-O-NR-S(O)-NR-、-O-NR-S(O)₂-O-、-O-NR-S(O)₂-NR-、-O-NR-S(O)₂-、-O-P(O)R₂-、-S-P(O)R₂-或-NR-P(O)R₂-, 其中每个R独立地为氢、烷基或取代的烷基、和类似基团。

[0133] 如本文所使用的, 术语“羟基(hydroxyl或hydroxy)”指代基团-OH。

[0134] 如本文所使用的, 术语“氧代”指代与附接的碳双键键合的氧取代基。

[0135] 如本文所使用的, 术语“亚磺酰基”表示基团-S(O)-。

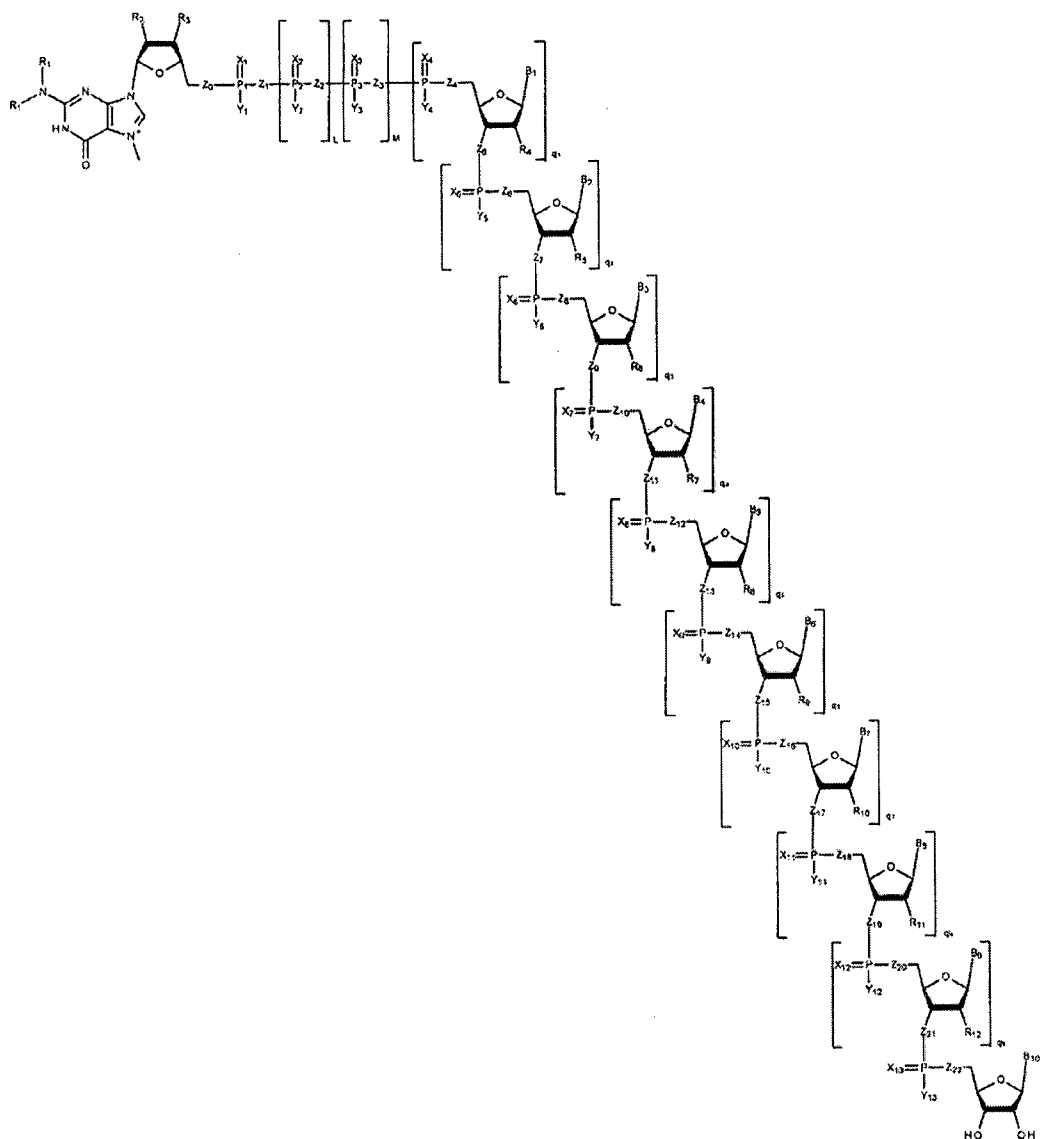
[0136] 如本文所使用的, 术语“取代的亚磺酰基”表示基团-S(O)R¹, 其中R¹为低级烷基、取代的低级烷基、环烷基、取代的环烷基、环烷基烷基、取代的环烷基烷基、杂环基、取代的杂环基、杂环基烷基、取代的杂环基烷基、芳基、取代的芳基、杂芳基、取代的杂芳基、杂芳烷基、取代的杂芳烷基、芳烷基或取代的芳烷基。

[0137] 如本文所使用的, 术语“磺酰基”表示基团-S(O)₂-。

[0138] 如本文所使用的, 术语“取代的磺酰基”表示基团-S(O)₂R¹, 其中R¹为低级烷基、取代的低级烷基、环烷基、取代的环烷基、环烷基烷基、取代的环烷基烷基、杂环基、取代的杂环基、杂环基烷基、取代的杂环基烷基、芳基、取代的芳基、杂芳基、取代的杂芳基、杂芳烷基、取代的杂芳烷基、芳烷基或取代的芳烷基。

[0139] 如本文所使用的, 术语“硫酰基”表示基团-S(O)₂-。

[0140] 本发明提供用于合成5'加帽RNA的方法和组合物, 其中起始加帽寡核苷酸引物具有通式形式^mGppp[N_{2'}omc]_n[N]_m, 其中^mG为N7-甲基化的鸟苷或任何鸟苷类似物, N为任何天然的、修饰的或非天然的核苷, “n”可以是1至4的任何整数并且“m”可以是1至9的整数。在本发明的一方面, 起始加帽寡核苷酸引物具有式I的结构:



[0141]

式 I

- [0142] 其中：
- [0143] B_1 至 B_{10} 的每一个独立地为天然的、修饰的或非天然的核苷碱基；
- [0144] M 为0或1；
- [0145] L 为0或1；
- [0146] q_1 为1；
- [0147] q_2 至 q_9 的每一个独立地为0或1；
- [0148] R_1 为H或甲基；
- [0149] R_2 和 R_3 独立地为H、OH、烷基、O-烷基、胺、叠氮基、卤素、连接体或可检测标志物；
- [0150] X_1 至 X_{13} 的每一个独立地为O或S；
- [0151] Y_1 至 Y_{13} 的每一个独立地为OH、SH、BH₃、芳基、烷基、O-烷基或O-芳基；
- [0152] Z_1 至 Z_{13} 的每一个独立地为O、S、NH、CH₂、C(卤素)₂或CH(卤素)，以及
- [0153] R_5 至 R_{12} 的每一个独立地为H、OH、OMe或可检测标志物。

[0154] 在某些实施方式中,起始加帽寡核苷酸引物为三聚体 ($q^2-q^9=0$)、四聚体 ($q^3-q^9=0$)、五聚体 ($q^4-q^9=0$)、六聚体 ($q^5-q^9=0$)、七聚体 ($q^6-q^9=0$)、八聚体 ($q^7-q^9=0$)、九聚体 ($q^8-q^9=0$)、十聚体 ($q^9=0$) 或十一聚体。起始加帽寡核苷酸三聚体引物的多个实例在下表 I 中列出:

[0155] 表 I

[0156]

起始加帽寡核苷酸引物的序列	
	m ⁷ GpppApA
	m ⁷ GpppApC
	m ⁷ GpppApG
	m ⁷ GpppApU
	m ⁷ GpppCpA
	m ⁷ GpppCpC
	m ⁷ GpppCpG
	m ⁷ GpppCpU
	m ⁷ GpppGpA
	m ⁷ GpppGpC
	m ⁷ GpppGpG
	m ⁷ GpppGpU
	m ⁷ GpppUpA
	m ⁷ GpppUpC
	m ⁷ GpppUpG
	m ⁷ GpppUpU
	m ⁷ G _{3'} OmepppApA
	m ⁷ G _{3'} OmepppApC
	m ⁷ G _{3'} OmepppApG
	m ⁷ G _{3'} OmepppApU
	m ⁷ G _{3'} OmepppCpA
	m ⁷ G _{3'} OmepppCpC
	m ⁷ G _{3'} OmepppCpG
	m ⁷ G _{3'} OmepppCpU
	m ⁷ G _{3'} OmepppGpA
	m ⁷ G _{3'} OmepppGpC
	m ⁷ G _{3'} OmepppGpG
	m ⁷ G _{3'} OmepppGpU
	m ⁷ G _{3'} OmepppUpA
	m ⁷ G _{3'} OmepppUpC
	m ⁷ G _{3'} OmepppUpG
	m ⁷ G _{3'} OmepppUpU
	m ⁷ G _{3'} OmepppA _{2'} OmepA
	m ⁷ G _{3'} OmepppA _{2'} OmepC
	m ⁷ G _{3'} OmepppA _{2'} OmepG
	m ⁷ G _{3'} OmepppA _{2'} OmepU
	m ⁷ G _{3'} OmepppC _{2'} OmepA
	m ⁷ G _{3'} OmepppC _{2'} OmepC
	m ⁷ G _{3'} OmepppC _{2'} OmepG
	m ⁷ G _{3'} OmepppC _{2'} OmepU
	m ⁷ G _{3'} OmepppG _{2'} OmepA
	m ⁷ G _{3'} OmepppG _{2'} OmepC

[0157]

$m^7G_3'OmepppG_2'Ome$	pG
$m^7G_3'OmepppG_2'Ome$	pU
$m^7G_3'OmepppU_2'Ome$	pA
$m^7G_3'OmepppU_2'Ome$	pC
$m^7G_3'OmepppU_2'Ome$	pG
$m^7G_3'OmepppU_2'Ome$	pU
$m^7GpppA_2'Ome$	pA
$m^7GpppA_2'Ome$	pC
$m^7GpppA_2'Ome$	pG
$m^7GpppA_2'Ome$	pU
$m^7GpppC_2'Ome$	pA
$m^7GpppC_2'Ome$	pC
$m^7GpppC_2'Ome$	pG
$m^7GpppC_2'Ome$	pU
$m^7GpppG_2'Ome$	pA
$m^7GpppG_2'Ome$	pC
$m^7GpppG_2'Ome$	pG
$m^7GpppG_2'Ome$	pU
$m^7GpppU_2'Ome$	pA
$m^7GpppU_2'Ome$	pC
$m^7GpppU_2'Ome$	pG
$m^7GpppU_2'Ome$	pU

[0158] 本发明涵盖的其他起始加帽寡核苷酸引物包括具有已知的或新的碱基类似物的那些引物。用于合成起始加帽寡核苷酸引物的方法在下面的实例中举例说明。

[0159] 转录

[0160] 在真核细胞中,信使RNA(mRNA)的转录由RNA聚合酶II完成。这是具有复杂调控的复杂多亚基(亚单位,subunit)酶。为了在体外进行大规模转录,研究通常使用源自T7、T3、SP6、K1-5、K1E、K1F或K11噬菌体的单亚基噬菌体聚合酶。该族聚合酶具有~17个核苷酸的简单最小启动子序列,其不需要辅助蛋白并且具有对起始核苷酸序列的最小限制。虽然本申请重点讨论T7RNA聚合酶(T7RNAP),但本领域技术人员将理解本发明可以用其他RNA聚合酶实践。

[0161] T7RNAP以至少两种蛋白质状态存在。第一种被称为“无效(顿挫型,abortive)复合体”并且与转录起始相关。第二种是非常进行性的构象,称为“延伸复合体”。体外转录可以分成六个步骤:1)RNA聚合酶结合启动子序列,2)转录起始,3)非进行性延伸,被称为无效转录,在此期间聚合酶频繁地释放DNA模板和短小的无效转录物,4)开放复合体转化为封闭复合体,5)进行性延伸,以及6)转录终止。转录期间产生的大量RNA由长度为2-8个核苷酸的短小无效片段组成(Biochemistry 19:3245-3253(1980);Nucleic Acids Res.9:31-45(1981);Nucleic Acids Res-15:8783-8798(1987);Biochemistry 27:3966-3974(1988))。合成约10-14个碱基后,RNA聚合酶从无效循环逃脱,同时失去与启动子DNA的序列特异性接触,并且形成进行性延伸复合体,其中RNA链以序列非依赖性方式延伸(J.Mol.Biol.183:165-177(1985);Proc.Natl.Acad.Sci.U.S.A.83:3614-3618(1986);Mol.Cell Biol.7:3371-3379(1987))。

[0162] 最具活性的III类T7启动子的共有序列包括转录起始位点上游的17bp序列,和下游的6bp序列(Cell 16:815-25.(1979))。第一个被转录核苷酸的位置通常被称为RNA的+1

转录物核苷酸,第二个被转录核苷酸为+2转录物核苷酸,依此类推(表2)。在转录期间,两条链解链(熔解,melted)以形成转录泡并且双链体的下链(表2中的3'至5'所示)为用于转录的模板。对于转录物核苷酸+3及以后,模板链主要通过沃森-克里克碱基配对相互作用来限定被转录核苷酸的特性(identity)。此处,编码第一个RNA转录物核苷酸的核苷酸被限定为模板的+1核苷酸。在表2所示的实例中,+1转录物核苷酸为G,且+1模板核苷酸为C。同样地,+4转录物核苷酸为A,且+4模板核苷酸为T。

[0163] 表2

[0164]

转录物中的位置

转录物序列



启动子上链

5'-TAATACGACTCACTATAGGGAGA...-3' SEQ ID NO. 1

启动子下链

3'-ATTATGCTGAGTGATATCCCTGT...-5' SEQ ID NO. 2

模板中的位置

[0165] 不同于DNA聚合酶,T7RNAP在不存在引物的情况下起始RNA合成。起始中的第一步骤被称为从头(de novo) RNA合成,其中RNA聚合酶识别DNA模板上的特定序列,选择与在位置+1和+2处的模板残基互补的第一对核苷酸三磷酸,并且催化形成磷酸二酯键以形成二核苷酸。起始核苷酸对聚合酶具有的亲和力比延伸期间使用的那些低。对于第一起始NTP,Kd值为2mM,且对于第二起始NTP,Kd值为80μM,而在延伸期间,对于NTP,Kd为大约5μM(J.Mol.Biol.(2007) 370,256-268)。已经发现从头合成是转录期间的限速步骤。T7 RNAP展现出对GTP作为起始核苷酸的强烈偏向(J.Biol.Chem.248:2235-2244(1973))。在基因组中的17个T7启动子中,15个以GTP(并且13个以pppGpG)起始,而在转录延伸期间不存在明显的NTP偏好(J.Mol.Biol.370:256-268(2007))。T7RNA聚合酶很少在位置+1处编码A的启动子上起始;取而代之,转录主要以位置+2处编码的G起始(J.Biol.Chem.278:2819-2823(2003))。

[0166] 在从头RNA合成期间,起始核苷酸的结合主要通过由碱基堆叠创造的自由能、聚合酶残基、起始核苷酸的鸟嘌呤部分之间的特异性相互作用和碱基互补性相互作用实现(J.Mol.Biol.370:256-268(2007))。

[0167] 已知T7RNAP还可以用短小寡核苷酸引物起始。例如,已知T7基因组中的13个启动子用pppGpG起始(J.Mol.Biol.370:256-268(2007))。数个小组示出T7RNAP可以从二核苷酸引物起始(Biochemistry 24:5716-5723(1985))。Axelrod等证实了无帽GpA二核苷酸可以从分别为2'-脱氧胞苷和2'-脱氧胸苷的+1和+2模板核苷酸("CT"模板)起始。其反应条件为200微摩尔(μM)二聚体和100μM ATP、CTP、GTP和UTP。其反应混合物还含有100μM 3' dATP、3' dCTP、3' dUTP或50μM 3' dGTP。他们仅观察到GpA起始的RNA,而没有观察到GpA起始的RNA和来自GTP起始的5'三磷酸RNA的混合物。这可能是由于所采用的反应条件。对于第一起始鸟苷,100μM GTP远低于2mM Kd的T7聚合酶(J.Mol.Biol.(2007) 370,256-268)。由于GTP与起始寡核苷酸竞争起始,使用低GTP浓度有利于GpA起始,但导致低转录产率(估计最大计算产率为<150ug/ml的反应)。当在"CT"模板上用ApG、CpG、UpG或GpG起始转录时,他们观察到具有另外的非模板化的5'核苷酸(分别为A、C、U或G)的RNA转录物的形成。

[0168] Axelrod等还使用无帽GpG二核苷酸在模板核苷酸+1和+2是2'-脱氧胞苷("CC"模板)的启动子上起始RNA合成。他们观察到起始的低保真度并观察到三种不同的转录产物。他们陈述,"放射自显影照片的检查指示每个三联体的一个成员由在正常(+1)位置用GpG起始得到,每个三联体的第二个成员由在异常(-1)位置用GpG起始得到,并且每个三联体的第三个成员由在正常位置用鸟苷三磷酸起始得到。因此,GpG对于 ϕ 10启动子("CC"模板),以及对于 ϕ 1.1A启动子("CT"模板)是相对弱的起始子,并且不能阻止在所使用的浓度下通过鸟苷三磷酸的正常起始"。他们没有观察到通过 ϕ 10启动子("CC"模板)用GpA二核苷酸起始。CpA、ApC和ApA没有在上述"TC"和"CC"模板中的任何一个上充当起始子,大概因为这些不能在位置+1和+2处与模板核苷酸杂交。Axelrod等描述的方法被设计用于生产非常少量的测序用的放射性的转录物,并且不适用于大规模生产有用的药物量的RNA。如果使用更大浓度($\sim 5\text{mM}$)的起始二聚体和NTP(包括GTP)以增加RNA的产率,则预期的结果是低比例的以起始二聚体开始的RNA,因为在更接近于Kd(2mM)的NTP浓度下,GTP与二聚体有效地竞争从+1核苷酸起的起始,产生大比例的用pppG开始的RNA。

[0169] Pitulle等示出,用T7RNAP转录可以用无帽寡核苷酸(2聚体至6聚体)起始(Gene, 112:101-105(1992))。这些寡核苷酸具有5'-OH或5'单磷酸。他们还用生物素-ApG结构的寡核苷酸起始转录。在该研究中所使用的所有寡核苷酸均含有3'端G。Pitulle等还示出,2'-O-甲基残基和脱氧残基可以被包括在引物序列内以产生在RNA的5'端处或5'端附近具有2'-O-甲基化或2'-脱氧残基的RNA转录物。在该出版物中清楚的是,任何起始寡核苷酸的3'端鸟苷残基与+1模板核苷酸配对。这导致非模板化的核苷酸附加到被转录RNA的5'端。具体而言,作者陈述,"由于除了Y-末端的G之外不需要与模板DNA碱基配对,此区段中的序列变化也是很可能的"。因此没有一个起始寡核苷酸引物仅在位置+1处与模板核苷酸"C"完全互补,而不与之后位置(+2、+3等)处的任何核苷酸互补。这在该小组随后的方法论文中得到确认(Methods Mol Biol. 74:99-110(1997), Methods Mol. Biol. 252:9-17, (2004))。他们的方法不同于本文描述的方法——其中起始加帽寡核苷酸引物的所有核苷酸与位置+1和之后位置处的模板核苷酸完全互补。Kleineidam等通过用2'-脱氧或2'-O-甲基糖修饰的二聚体或三聚体起始创建了5'修饰的tRNA转录物(Nucleic Acids Research 21:1097-1101(1993))。再次,作者陈述引物的3'端鸟苷在+1模板核苷酸"C"处起始。

[0170] Ishikawa等的另一个研究示出, $m^7\text{GpppApG}$ 、 $m^7\text{Gppp}^{m^5}\text{ApG}$ 、 $m^7\text{GpppA}_{2'-\text{om}}\text{pG}$ 或 $m^7\text{Gppp}^{m^5}\text{A}_{2'-\text{om}}\text{pG}$ 结构的加帽起始寡核苷酸三聚体可以在模板位置+1和+2处具有2'-脱氧胞苷残基的模板上起始转录("CC"模板;Nucleic Acids Symposium Series No.53:129(2009))。作者陈述,"来自使用 $m^7\text{G}5'$ pppG案例的不同结果可以由 $m^7\text{G}5'$ pppN1pG中的额外腺苷(N1)和T7启动子中-1位置处的2'-脱氧胸苷之间的碱基配对造成的。"该方法明确区别于在本发明中所描述的方法——其中起始加帽寡核苷酸三聚体的+1和+2核苷酸与模板核苷酸的+1和+2配对。Ishikawa等使用了6mM起始寡核苷酸三聚体、0.9mM GTP和各7.5mM的ATP、CTP和UTP。作者使用了相对于竞争性GTP大于6倍过量的加帽起始寡核苷酸引物(转录反应中最昂贵的核苷酸组分),以驱动转录反应相对于pppRNA倾向加帽RNA,这增加了合成RNA的总成本。在另一方面,低浓度的GTP(0.9mM)限制了转录反应中RNA总产率(理论上到少于1.4mg/mL)。相反,本文所描述的方法不需要限制任何NTP的浓度以获得有效的RNA加帽和RNA的更高产率(2至6mg/mL)两者,并且因此允许以商业上有用的成本生产高质量的mRNA。

[0171] 以上讨论的出版物均没有直接地测量RNA加帽效率,所以那些研究中的加帽程度是未知的。

[0172] 重要的是,以上所描述的所有研究中,+1模板核苷酸是2'-脱氧胞苷(Biochemistry 24:5716-5723(1985),Gene,112:101-105(1992),Methods Mol.Biol.74:99-110(1997),Methods Mol.Biol.252:9-17,(2004),Nucleic Acids Research 21:1097-1101(1993),Nucleic Acids Symposium Series No.53:129(2009))。

[0173] 自用寡核苷酸引物进行转录起始研究公开以来超过20年中,还没有发表过用含有5'-至5'反向帽结构的起始寡核苷酸引物进行转录起始的实例,直到在Nucleic Acids Symposium Series No.53:129(2009)中的短报告公开。

[0174] 本文提供的用于制备5'-加帽RNA的方法和组合物包括但不限于mRNA、小核RNA(snRNA)、小核仁RNA(snoRNA)、小卡哈尔体(cajal body)特异性RNA(scaRNA)。这些方法涉及使用含帽的寡核苷酸引物、核苷5'-三磷酸(NTP)和RNA聚合酶进行DNA模板化的和启动子控制的RNA合成。在某些方面,该方法使用起始加帽寡核苷酸引物,其在RNA合成中,特别地在加帽mRNA合成中提供效用。起始寡核苷酸引物具有类似天然RNA分子的帽0、帽1、帽2或TMG-帽的结构,其包括在RNA的倒数第二帽1和相邻于倒数第二帽2的5'-位置处的2'-O-甲基化的核苷单元。天然的帽0结构不具有2'-O-甲基化的核苷单元。

[0175] 用于制备RNA的方法和组合物,所述RNA包括但不限于mRNA、snRNA、snoRNA、scaRNA、转运RNA(tRNA)、核糖体RNA(rRNA)和转运-信使RNA(tmRNA)——其携带在分子的5'-端处或5'-端附近的修饰。这些方法涉及使用具有或不具有帽的起始寡核苷酸引物、核苷5'-三磷酸(NTP)和RNA聚合酶进行DNA模板化的和启动子控制的RNA合成。在某些方面,该方法使用修饰的起始寡核苷酸引物,所述修饰的起始寡核苷酸引物携带在RNA合成中;特别地在5'-修饰的RNA的合成中提供效用的结构修饰。

[0176] 起始加帽寡核苷酸引物具有开放的3'-OH基团,其通过向引物的3'端添加核苷酸单元而允许在DNA模板上RNA聚合酶介导的RNA合成起始。起始加帽寡核苷酸引物在转录起始位点处(即,起始位点位置更接近于启动子序列的3'-端并且可以与启动子序列重叠)与模板DNA序列基本上互补。在某些实施方式中,起始加帽寡核苷酸引物主要在始于引物的3'端的一个方向("正向")指导RNA的合成。在某些方面和实施方式中,起始加帽寡核苷酸引物对于RNA的合成起始胜过任何核苷5'-三磷酸,由此最大化用起始加帽寡核苷酸引物开始的RNA的生产,并且最小化用5'-三磷酸核苷(一般GTP)开始的RNA的生产。

[0177] 通过体外转录的mRNA制造利用了高活性噬菌体RNA聚合酶(T3、T7、SP6以及其他)。RNA聚合酶在特定启动子的控制下工作,该启动子被并入DNA质粒构建体中模板核苷酸序列前。转录过程通常从嘌呤核苷5'-三磷酸(一般GTP)开始,并继续持续直到RNA聚合酶遇到终止序列或完成DNA模板。

[0178] 如上讨论的,含有帽0的mCAP二核苷酸类似物($^7\text{mG}(5')\text{ppp}(5')\text{N}$)已经被用于起始体外转录(例如,RNA 1:957-967(1995))。使用这些二核苷酸类似物产生的加帽RNA分子含有帽0。然而,只有约50%的合成加帽RNA分子具有正确的帽0"正向"定向。为了将具有帽0的RNA转化为具有帽1的RNA,必须使用(核苷-2'-O)甲基转移酶进行另外的酶促反应。然而这种转化可能不是定量的;不容易控制并且难以将剩余的帽0 RNA与帽1 RNA分离。另外,来自NTP(具体地GTP)对起始转录的竞争进一步减少了活性加帽RNA分子的产量。

[0179] 另外,诸如 ${}^7\text{mG}^{\text{ome}}(5')\text{ppp}(5')\text{N}$ 的修饰的二核苷酸类似物和携带核糖上3'和/或2'位置被阻断的修饰 ${}^7\text{mG}$ 残基的其他相关ARCA类似物已经被用于在体外转录的起始(例如, RNA 7:1486-1495 (2001))。这些ARCA帽类似物仅以“正向”方向指导RNA合成,并且因此产生在5'-端上具有(天然的)帽0的RNA分子(具有对 ${}^7\text{mG}$ 残基的2'和/或3'修饰)。与使用标准二核苷酸类似物(${}^7\text{mG}(5')\text{ppp}(5')\text{N}$)制备的RNA相比,这种RNA在翻译系统中更具活性。为了将具有ARCA帽0的RNA转化为具有ARCA帽1的RNA,必须用类似于先前讨论的二核苷酸类似物所需的(核糖-2'-O)甲基转移酶进行另外的酶促反应。该方法与关于mCAP二核苷酸类似物详述的方法具有相同的缺点:将具有帽0的RNA转化为具有ARCA帽1的RNA可能不是定量的;该反应不容易控制,难以将剩余的帽0 RNA与帽1 RNA分离,并且来自NTP(具体地GTP)对转录起始的竞争进一步减少了活性加帽RNA分子的产量。

[0180] 具有3'-末端鸟苷残基的短小寡核苷酸引物(2至6聚体)已经被用于体外转录的起始(Pitulle, C. et al., Gene, 112:101-105 (1992))。这些寡核苷酸引物含有修饰的和未修饰的核糖核苷残基(例如,修饰的核糖核苷残基包括2'-O-甲基化的核苷残基和2'-脱氧核糖核苷残基)。较短的寡核苷酸引物(二聚体至四聚体)对于转录的起始基本上胜过GTP,而与GTP相比,较长的引物(五聚体至六聚体)在转录的起始中效率低得多。这可能是因为这些较长的引物(如它们所被设计)与DNA模板在起始位点具有低百分比的互补性。相比之下,二聚体AG(如所被设计)与DNA模板在起始位点互补。在本节中讨论的使用寡核苷酸引物产生的RNA分子具有内部2'-O-甲基化的核苷,但不含有5'-帽0、帽1、帽2或TMG-帽。为了将不具有帽结构的RNA转化为具有帽1、帽2或TMG-帽结构的RNA,将进行使用加帽酶的另外的酶促反应。然而,这种转化与上述那些具有相同的缺点。

[0181] 已经化学制备了含有帽结构和内部2'-O-甲基化核苷残基的其他短RNA寡核苷酸(Ohkubo et al., Org. Letters 15:4386-4389 (2013))。使用T4DNA连接酶和互补DNA碎片(splinter)寡核苷酸将这些短加帽寡核苷酸与长RNA的“去帽(decapitated)”(不具有5'-帽结构)片段连接。使用该化学-酶促方法合成的最终RNA具有内部2'-O-甲基化核苷残基和5'-TMG-帽结构两者。然而,使用该连接方式仅制备出短加帽RNA(<200聚体)。而且,产率低(15-30%)。不容易控制和优化T4DNA连接反应,并且需要使用聚丙烯酰胺凝胶电泳的费力分离过程和从剩余无帽RNA中分离加帽RNA。通过PAGE方法从剩余无帽mRNA中分离长(500-10000个碱基)加帽mRNA是不可行的。

[0182] 最后,5'-修饰的核苷或5'-修饰的单核苷酸或5'-修饰的二核苷酸(一般是鸟苷的衍生物),已经被用于起始体外RNA转录(Gene, 112:101-105 (1992) 和 Bioconjug. Chem., 10:371-378 (1999))。这些起始核苷和核苷酸可以携带标记或亲和基团(例如生物素),并且当在RNA的5'-端上并入时,将允许容易的检测、分离和纯化合成的RNA。该5'标记的或标签化的RNA对于一些应用可能是必要的。然而,该策略不用于制备具有帽0、帽1、帽2或TMG-帽结构的mRNA。

[0183] 在本发明的某些方面,提供了式I的起始加帽寡核苷酸引物的组合物。在相关方面,提供了使用式I的起始加帽寡核苷酸引物合成RNA的方法。

[0184] 起始加帽寡核苷酸引物

[0185] 本发明的起始加帽寡核苷酸引物具有可以在起始位点与DNA模板上的序列互补的杂交序列。用于本文提供的方法和组合物中的引物的杂交序列的长度取决于几个因素,包

括模板核苷酸序列的特性和该引物与DNA模板杂交或在体外转录期间使用的温度。用于转录的起始加帽寡核苷酸引物的特定核苷酸序列的期望长度的确定可以被本领域的普通技术人员或通过常规实验法确定。例如,核酸或寡核苷酸的长度可以基于期望的杂交特异性或选择性确定。

[0186] 在一些实施方式中,起始加帽寡核苷酸引物(包括反向5'-5'帽核苷酸)的核苷酸长度在3个至约9个之间,在一些实施方式中,起始加帽寡核苷酸引物(包括帽)的核苷酸长度在3个至约7个之间,在一些实施方式中,起始加帽寡核苷酸引物(包括帽)的核苷酸长度在3个至约5个之间,并且在一些实施方式中,起始加帽寡核苷酸引物(包括帽)的核苷酸长度为约3个。起始加帽寡核苷酸引物内的杂交序列的长度可以等于或短于起始加帽寡核苷酸引物的总长度。

[0187] 杂交序列的存在促使起始加帽寡核苷酸引物主要仅以期望的方向(即,“正向”方向)在起始位点与DNA模板的互补序列对齐。在正向方向,RNA转录物始于反向鸟苷残基(即, $^7\text{mG}(5')\text{ppp}(5')\text{N}\cdots$)。DNA模板(图1)上引物对齐的正向方向相对于不正确的“反向”方向的优势地位通过杂交复合体的热力学维持。后者是通过起始加帽寡核苷酸引物的杂交序列的长度和与DNA模板的杂交所涉及的碱基的特性来确定的。在期望的正向方向下的杂交也可以取决于DNA模板和起始加帽寡核苷酸引物杂交的或用于体外转录期间的温度和反应条件。

[0188] 与用标准GTP、ATP、CTP或UTP起始的效力相比,本发明的起始加帽寡核苷酸引物增强了转录起始的效力。在一些实施方式中,当RNA的合成主要始于起始加帽寡核苷酸引物而非始于转录混合物中的任何NTP时,认为转录的起始增强。转录起始的增强效率导致RNA转录物的更高的产率。转录起始的增强的效率可以相对于用无起始加帽引物的常规方法进行的RNA合成增加约10%、约20%、约40%、约60%、约80%、约90%、约100%、约150%、约200%或约500%。在某些实施方式中,“起始加帽寡核苷酸引物”对于转录的起始胜过任何NTP(包括GTP)。本领域普通技术人员能够容易地确定底物活性的水平和起始加帽寡核苷酸引物的效力。确定底物效力的方法的一个实例在实施例13中说明。在某些实施方式中,起始从加帽寡核苷酸引物而不是从NTP发生,其导致转录的mRNA的更高的加帽水平。

[0189] 在一些方面,提供了利用具有取代或修饰的起始加帽寡核苷酸引物合成RNA的方法。在一些方面,起始加帽寡核苷酸引物的取代和修饰不显著损害RNA的合成。可以进行常规测试合成以确定用修饰的起始加帽寡核苷酸引物是否可以获得期望的合成结果。本领域技术人员可以进行这种常规实验法以确定是否可以获得期望的结果。起始加帽寡核苷酸引物的取代或修饰包括,例如,一个或多个修饰的核苷碱基、一个或多个修饰的糖、一个或多个修饰的核苷酸间连接体和/或一个或多个修饰的三磷酸桥。

[0190] 可以包括本文提供的方法和组合物中的一个或多个修饰基团的修饰起始加帽寡核苷酸引物可以通过将NTP并入到开放的3'-OH基团上而在DNA模板上通过RNA聚合酶延伸。起始加帽寡核苷酸引物可以包括天然的RNA核苷和DNA核苷、修饰的核苷或核苷类似物。起始加帽寡核苷酸引物可以含有天然的核苷酸间磷酸二酯连接体或其修饰形式或其组合。

[0191] 在一个实施方式中,修饰基团可以是热不稳定的基团,其以随着酶反应介质的温度升高而增加的速率从修饰的起始加帽寡核苷酸引物解离。寡核苷酸和NTP的热不稳定基团的实例被描述于Nucleic Acids Res.,36:e131(2008),Collect.Symp.Ser.,10:259-263

(2008) 和 *Analytical Chemistry*, 81:4955-4962 (2009) 中。

[0192] 在一些方面, 提供了合成RNA的方法, 其中至少一个或多个NTP被添加到转录反应, 可以具有如本文所公开的修饰。在一些方面, 该至少一个NTP的修饰不显著损害RNA聚合酶介导的RNA合成。NTP的修饰可以包括, 例如, 一个或多个修饰的核苷碱基、一个或多个修饰的糖、一个或多个修饰的5'-三磷酸。修饰的NTP可以并入到起始加帽寡核苷酸引物的3'端上并且其并不阻断转录并支持引物的进一步延伸。

[0193] 在另一个实施方式中, 起始加帽寡核苷酸引物的修饰基团可以是可检测标记或可检测标志物。因此, 在转录之后, 含有可检测标记或标志物的靶RNA可以通过尺寸、质量、颜色和/或亲和捕获来鉴定。在一些实施方式中, 可检测标记或标志物为荧光染料; 并且亲和捕获标记为生物素。在某些实施方式中, 转录反应的一个或多个组分(起始加帽寡核苷酸引物和/或NTP)可以用可检测标记或标志物进行标记。因此, 在转录后, RNA分子可以例如通过尺寸、质量、亲和捕获或颜色来鉴定。在一些实施方式中, 可检测标记为荧光染料; 并且亲和捕获标记为生物素。

[0194] 标准化学和酶促合成方法可以被用于合成本发明的“起始加帽寡核苷酸引物”, 并且在本文的实施例章节中被公开。

[0195] 试剂盒

[0196] 还考虑包括用于实施转录的“起始加帽寡核苷酸引物”的试剂盒。例如, 试剂盒可以含有用于合成普通RNA(例如, FLuc mRNA)的所有转录试剂。更具体地, 试剂盒可以含有: “起始加帽寡核苷酸引物”; 标记用于转录的容器; 实施RNA合成的说明; 和选自一种或多种修饰的或未修饰的起始加帽寡核苷酸引物、一种或多种未修饰的NTP、一种或多种修饰的NTP(例如, 假尿苷5'-三磷酸)、RNA聚合酶、其他酶、反应缓冲液、镁和DNA模板的一种或多种试剂。

[0197] 本发明的起始加帽寡核苷酸引物相对于当前的方法和组合物具有显著优势, 涉及各种起始核苷、核苷酸和寡核苷酸的使用或含有帽0结构的聚磷酸二核苷酸衍生物(诸如mCAP和ARCA)的使用。起始加帽寡核苷酸引物是与现有的转录系统和试剂相容的, 并且不需要另外的酶或试剂。另外, 使用起始加帽寡核苷酸引物使得若干非酶促和酶促步骤(诸如加帽和2'-O-甲基化)是不必要的, 因此减少了RNA合成的过程复杂性和成本。

[0198] 虽然本文描述的示例性方法涉及T7 RNA聚合酶介导的转录反应, 但本领域已知的用于转录反应的多种其他RNA聚合酶可以与本发明的组合物和方法一起应用。包括可以利用的天然的或突变的变体在内的其他酶包括, 例如, SP6和T3 RNA聚合酶和来自其他来源的RNA聚合酶(包括热稳定的RNA聚合酶)。

[0199] 一些核酸复制和扩增方法可以包括转录作为该过程的一部分。在这些方法中有: 转录介导的扩增(TMA)和基于核酸序列的扩增(NASBA)、DNA和RNA测序以及本领域已知的其他核酸延伸反应。技术人员将理解其他方法可以代替转录方法或与转录方法(包括在将来开发的转录反应的变型)一起使用。

[0200] 治疗用途

[0201] 本发明还考虑生产含有起始加帽寡核苷酸引物的mRNA以用作药物组合物中的治疗剂、在细胞中引入含有起始加帽寡核苷酸引物的RNA以治疗细胞的医学状况或者将含有起始加帽寡核苷酸引物的RNA引入到利用那些RNA产生可以对宿主细胞具有治疗作用的蛋

白质的细胞中。

[0202] 一种用于利用含有起始加帽寡核苷酸引物的RNA治疗状况的方法包括向患有或疑似患有状况的对象给予含有式I的起始加帽寡核苷酸引物的RNA或包含这种RNA的组合物的步骤,该状况的症状/症状学可以严重程度降低或被消除。

[0203] 当在4mg/ml或更低的浓度下在药学上可接受的载体和/或药学上可接受的盐中配制时,含有式I“化合物”的起始加帽寡核苷酸引物的RNA,相对于单独使用该药学上可接受的载体未治疗的个体,有效地产生至少20%、30%、40%、50%、60%、70%、80%、90%或更多的症状和/或症状学降低。

[0204] 药物组合物可以被配制用于通过注射或本领域技术人员已知的用于治疗具体状况的其他合适的途径进行给予。用于肠胃外给予的可注射组合物一般含有在适当的溶液和/或药物载体(如无菌生理盐水)中的活性化合物。该组合物还可以被配制为在脂质或磷脂中的悬浮液、在脂质体悬浮液中或在含水乳液中。

[0205] 本领域技术人员已知用于制备多种组合物和/或配制物的方法。参见Remington's pharmaceutical Sciences (19th Ed., Williams&Wilkins, 1995)。待给予的组合物将含有对于增加靶细胞或组织中期望蛋白质的表达而言在药学上安全且有效量的一定量选定化合物。

[0206] 在一些实施方式中,药物组合物含有至少0.1% (w/v) 的如上描述的化合物,在一些实施方式中,药物组合物含有大于0.1%,在一些实施方式中,药物组合物含有上至约10%的化合物,在一些实施方式中,药物组合物含有上至约5%的化合物,并且在一些实施方式中,药物组合物含有上至约1% (w/v) 的化合物。适当浓度的选择取决于诸如活性剂递送的期望剂量、频率和方法的因素。

[0207] 对于对象(诸如哺乳动物或人)的治疗,基于诸如对象的体重和整体健康情况、所治疗的状况、症状的严重程度等的因素确定剂量。确定剂量和浓度以产生期望的好处,同时避免任何不期望的副作用。对于人类患者,主题化合物的一般剂量在约0.0005至500mg/天的范围内,并且在一些实施方式中,在约1-100mg/天之间的范围内。例如,较高剂量方案包括,例如50-100mg/天、75-100mg/天或50-75mg/天,并且较低剂量方案包括,例如1-50mg/天、25-50mg/天或1-25mg/天。

[0208] 通过以下非限制性实施例说明本发明的各个方面。这些实施例是用于说明的目的,而不是对本发明的任何实践的限制。应该理解,在不脱离本发明的精神和范围的情况下可以做出变化和修改。本领域普通技术人员容易知道如何合成或商业上获得本文所描述的试剂和组分。

[0209] 实施例1

[0210] 由鸟苷5'-二磷酸制备7-甲基鸟苷5-二磷酸(pp^{7m}G) (图1)

[0211] 向搅拌中的鸟苷5'-二磷酸(2.5mmol)在40.0mL水中的溶液,添加乙酸以将溶液的pH调节至4.0。经30分钟的时间向该混合物滴加硫酸二甲酯(4.0mL),并且将该反应混合物在室温下搅拌4小时,同时使用0.1M NaOH溶液将反应混合物的pH维持在约4.0。4小时后,用CH₂Cl₂ (3×50mL) 萃取反应混合物以去除未反应的硫酸二甲酯。将水层用水稀释至500mL,用1M TEAB调节至pH 6.5,并且装载于DEAE Sephadex柱(3×50cm)上。使用0-1M TEAB, pH 7.5 (3L) 的线性梯度洗脱产物。将含有纯pp^{7m}G(三乙铵盐)的部分合并,蒸发,并在高真空下干燥

以得到精细白色粉末(产率:80%)。类似的程序公开在Bioorgan.Med.Chem.Letters 17: 5295-5299 (2007) 中。

[0212] 实施例2

[0213] 由pp^{7m}G制备7-甲基鸟苷5'-二磷酸咪唑(Im-pp^{7m}G)

[0214] (图2)

[0215] 将pp^{7m}G的三乙铵盐(0.4mmol)与咪唑(4mmol)、三苯膦(2mmol)和2,2'-二吡啶基二硫化物(2mmol)在干燥DMF(20ml)中反应8小时。通过将反应混合物倒入250ml的0.2M高氯酸钠溶液中沉淀粗Im-pp^{7m}G。将混合物冷却至-20℃并通过离心收集得到的沉淀,用丙酮(3×50mL)洗涤并在高真空下干燥。Im-pp^{7m}G的分离产率为大约100%。类似的程序公开在Nucleosides,Nucleotides,and Nucleic Acids 24:1131-1134 (2005) 和J.Org.Chem.64: 5836-5840 (1999) 中。

[0216] 实施例3

[0217] 由3'-O-甲基鸟苷制备3'-O-甲基鸟苷5'-磷酸(pG(x_{ome}))

[0218] (图3)

[0219] 在60-70℃下,将3'-O-甲基鸟苷(10mmol)溶解在磷酸三乙酯(40mL)中。将混合物在冰水浴中冷却至0℃,添加氯氧化磷(30mmol)并将混合物在氩气下在室温下搅拌3小时。通过在搅拌下缓慢添加1M TEAB(100mL;pH 8.5)猝灭反应。将该混合物搅拌8小时并用1L的水稀释。将得到的溶液装载到DEAE Sephadex柱(3×40cm)上并用线性梯度0.05至1.0M(3L)的TEAB(pH 7.5)洗脱产物。将含有纯产物的部分组合,蒸发至固体残余物并与甲醇(4×50mL)共蒸发以得到呈白色固体的pG(x_{ome})(三乙铵盐)(产率60%)。类似的程序公开在美国专利申请序列号2012/0156751中。

[0220] 实施例4

[0221] 由pG(x_{ome})制备3'-O-甲基鸟苷5'-磷酸咪唑(Im-pG(x_{ome}))

[0222] (图4)

[0223] 将pG(x_{ome})的三乙铵盐(0.5mmol)与咪唑(5mmol)、三苯膦(2.5mmol)和2,2'-二吡啶基二硫化物(2.5mmol)在干燥DMF(25mL)中反应5小时。通过将反应混合物倒入400mL的0.2M高氯酸钠的丙酮溶液中沉淀粗Im-pG(x_{ome})。将混合物冷却至-20℃并通过离心收集得到的沉淀,用丙酮(3×60mL)洗涤并在高真空下干燥(产率:100%)。类似的程序公开在美国专利申请序列号2012/0156751中。

[0224] 实施例5

[0225] 由Im-pG(x_{ome})制备3'-O-甲基鸟苷5'-二磷酸(ppG(x_{ome}))

[0226] (图5)

[0227] 将固体氯化锌(14.0mmol)以小份添加到Im-pG(x_{ome})(7.0mmol)的干燥DMF(40mL)溶液。将混合物在氩气下搅拌15分钟直到所有固体溶解。添加1M三丁基磷酸铵(tributylammonium phosphate)的DMF(40mL)溶液并将混合物在室温下搅拌。5小时后,将混合物用200mL的水稀释并用二氯甲烷(2×200mL)萃取。将水层用水(1L)稀释,装载到DEAE Sephadex柱(5×40cm)上并用0.05至1.0M(6L)TEAB(pH 7.5)的线性梯度洗脱。将含有纯产物的部分组合,蒸发并与甲醇(4×50mL)共蒸发以得到呈白色固体的ppG(x_{ome})(三乙铵盐)(产率:60%)。类似的程序公开在美国专利申请序列号2012/0156751中。

[0228] 实施例6

[0229] 由ppG (3'-ome) 制备7-甲基-3'-O-甲基鸟苷5-二磷酸 (pp^{7m}G (3'-ome))

[0230] (图6)

[0231] 制备ppG (3'-ome) (三乙铵盐; 3.0mmol) 在50mL水中的溶液并添加冰乙酸以将溶液的pH调节至4.0。经30分钟时间向该混合物滴加硫酸二甲酯 (10.0mL) 并将反应混合物在室温下搅拌4小时, 同时使用0.1MNaOH溶液维持pH为4.0±0.5。4小时后, 将反应混合物用CH₂Cl₂ (3×150mL) 萃取以去除未反应的硫酸二甲酯。将水层调节至pH 5.5, 用水 (500mL) 稀释并装载到DEAE Sephadex柱 (3×50cm) 上。将产物使用0-1.0M TEAB, pH 7.5 (3L) 的线性梯度洗脱。将含有纯pp^{7m}G (3'-ome) (三乙铵盐) 的部分合并, 蒸发, 并在高真空下干燥以得到精细白色粉末 (产率: 80%)。类似的程序公开在RNA 9:1108-1122 (2003); Nucleoside Nucleotides & Nucleic acids 25:337-340 (2006); 和美国专利申请序列号2012/0156751中。

[0232] 实施例7

[0233] 由pp^{7m}G (3'-ome) 制备7-甲基-3'-O-甲基鸟苷5-二磷酸咪唑 (Im-pp^{7m}G (3'-ome))

[0234] (图7)

[0235] 利用实施例1中描述的方案制备Im-pp^{7m}G (3'-ome)。类似的程序公开在RNA14:1119-1131 (2008) 中。

[0236] 实施例8

[0237] 用于制备pN_{2'-OR1} pN二核苷酸 (R₁=H或Me) 的一般程序

[0238] (图8)

[0239] 将亚磷酰胺单体 (i) (1.0mmol) 和2', 3', N-保护的核苷 (ii) (1.0mmol) 在含有2.5摩尔当量的活化剂 (四唑) 的10mL乙腈中反应。在室温下搅拌60分钟后, 将中间产物用碘从P (III) 氧化成P (V) 状态并且用二氯甲烷 (200mL) 和盐水 (200mL) 萃取。将有机层用硫酸钠干燥并蒸发至固体泡沫 (中间体 (iii))。

[0240] 为去除DMT-保护基团, 将中间体 (iii) 溶解在10mL的80%乙酸中, 并且在反应完成 (约1-2小时) 后, 将混合物蒸发并与甲醇 (5×30mL) 共蒸发以去除乙酸。将粗5'-OH二聚体 (iv) 使用在二氯甲烷中的5%甲醇作为洗脱剂通过硅胶色谱法分离并纯化。

[0241] 将5'-OH二聚体 (iv) (1.0mmol) 用2当量的双-氰乙基-N,N-二异丙基-亚磷酰胺和2当量的活化剂 (四唑) 在10mL的乙腈中亚磷酸化 (phosphitylated)。在室温下搅拌30分钟后, 将5'-亚磷酸化二聚体用碘从P (III) 氧化至P (V) 状态并用二氯甲烷 (150mL) 和盐水 (150mL) 萃取。将有机层蒸发至油状残留物, 与甲醇 (2×30mL) 共蒸发, 溶解在12mL甲醇中并添加浓氨水 (12mL)。将混合物在室温下保持超过48小时, 直到完成pN_{2'-OR1}pN二聚体 (v) 的脱保护。将混合物蒸发并与甲醇 (2×30mL) 共蒸发。

[0242] 当R=甲基时, 将粗二聚体 (v) 直接通过阴离子交换和反相色谱法纯化 (步骤5)。将各部分蒸发以得到呈白色固体的最终的pN_{2'-OR1}pN二聚体 (v) (v; R₁=甲基; 三乙铵盐) (35%总产率)。

[0243] 当R=TBDMs时, 将粗二聚体 (v) 用HF-3TEA混合物处理 (步骤4B) 以去除2'-OTBDMs保护基团 (Org. Biomol. Chem. 3:3851-3868 (2005) 和Nucl. Acids Res. 22:2430-2431 (1994))。当反应完成时, 将混合物用0.05M TEAB稀释并通过阴离子交换和反向色谱法纯化 (步骤5)。将各部分蒸发以得到呈白色固体的最终的pN_{2'-OR1}pN二聚体 (v) (R₁=H; 三乙铵盐)

(30%总产率)。

[0244] 实施例9

[0245] 用于合成具有帽0、帽1或帽2结构的起始寡核苷酸的一般程序(实施例中使用5'-磷酸化二核苷酸)

[0246] (图9)

[0247] A方式1:

[0248] 向Im-pp^{7m}G (x_{0mc}) 或Im-pp^{7m}G (2mmol; 钠盐形式) 和5'-磷酸化二核苷酸(1mmol; 三乙铵盐) 在DMF (50mL) 中的悬浮液缓慢添加无水的ZnCl₂ (1g), 同时将混合物在35℃下搅拌。24小时后, 通过添加25mM的EDTA水溶液(500mL) 终止反应并通过1M碳酸氢钠溶液中和。将混合物用水稀释至1L并装载到DEAE Sephadex柱(3×50cm) 上。将产物使用0-1M碳酸氢铵, pH 7.2 (2L) 的线性梯度洗脱。将含有纯产物的部分合并, 蒸发, 并在高真空下干燥以得到精细白色粉末(产率: 60%)。类似的方式公开在美国专利申请序列号2012/0156751; Bioorg.Med.Chem.Lett. 17:5295-5299 (2007) 和RNA14:1119-1131 (2008) 中。

[0249] B. 方式2:

[0250] 将Im-pp^{7m}G (x_{0mc}) 或Im-pp^{7m}G (2mmol) 溶解在含有MnCl₂ (2mmol) 的N-甲基吗啉缓冲液(0.2M, pH 7.0, 10mL) 中, 并添加到固体5'-磷酸化二核苷酸(1mmol; 三乙铵盐)。在室温下搅拌反应。24-40小时后, 用10mL的0.25M EDTA溶液终止反应。将混合物装载到DEAE Sephadex柱(3×50cm) 上。将产物使用0-1.0M碳酸氢铵, pH 7.2 (2L) 的线性梯度洗脱。将含有纯产物的部分合并, 蒸发, 并在高真空下干燥以得到精细白色粉末(产率50%)。类似的方式公开在Bioorganic&Medicinal Chemistry 21:7921-7928 (2013), Nucleic Acids Research 37:1925-1935 (2009); J.Org.Chem. 64:5836-5840 (1999) 中。(注: 1. 可以使用5'-磷酸化的三聚体、四聚体、五聚体、六聚体、七聚体、八聚体、九聚体或十聚体寡核苷酸代替在实施例9中说明的5'-磷酸化二核苷酸。2. 为了制备具有帽0的起始加帽寡核苷酸, 5'-磷酸化寡核苷酸在最先的5'-核苷残基上不具有2'-O-甲基, 例如pApG。3. 为了制备具有帽1的起始加帽寡核苷酸, 5'-磷酸化寡核苷酸在第一个5'-核苷残基上携带一个2'-O-甲基, 例如pA (x_{0mc}) pG。4. 为了制备具有帽2的起始加帽寡核苷酸, 5'-磷酸化寡核苷酸在第一个和第二个5'-核苷残基上携带两个2'-O-甲基, 例如pA (x_{0mc}) pG (x_{0mc}) pG)。

[0251] 实施例10

[0252] 根据式I的起始加帽寡核苷酸引物的结构

[0253] 图10示出了根据结构I的实例中使用的起始加帽寡核苷酸引物的结构。A) ^{m7}Gx_{0mc}pppG、B) ^{m7}GpppGx_{0mc}pG、C) ^{m7}Gx_{0mc}pppGx_{0mc}pG、D) ^{m7}GpppAx_{0mc}pG、E) ^{m7}Gx_{0mc}pppAx_{0mc}pG、F) ^{m7}GpppCx_{0mc}pG、G) ^{m7}Gx_{0mc}pppCx_{0mc}pG、H) ^{m7}GpppAx_{0mc}pGx_{0mc}pG。

[0254] 实施例11

[0255] 用ARCA引物在体外转录

[0256] 编码萤火虫荧光素酶的双链DNA转录模板通过聚合酶链反应生成。+1模板核苷酸为2'-脱氧胞苷。转录反应与25ug/mL转录模板、40mM Tris-HCl (pH 8.0)、27mM MgCl₂、2mM 亚精胺、10mM DTT、0.002% Triton x-100、1000单位/mL鼠RNA酶抑制剂(New England Biolabs目录#M0314)、2单位/mL无机焦磷酸酶(New England Biolabs目录#M2403)、4000单位/mL T7RNA聚合酶(New England Biolabs目录#M0251)、6mM ARCA(^{m7}Gx_{0mc}pppGG)、1.5mM

GTP、7.5mM ATP、7.5mM CTP和7.5mM UTP组合。转录在37°C下温育2小时。用10mM Tris-HCl (pH 7.6)、2.5mM MgCl₂、0.5mM CaCl₂和100单位/mL DNA酶I (New England Biolabs目录#M0303) 补充反应并在37°C下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱并通过将溶液调节至50mM Bis-Tris-Propane HCl (pH 6.0)、1mM MgCl₂、0.1mM ZnCl₂和250单位/mg南极磷酸酶 (New England Biolabs目录#M0289) 而去磷酸化。将反应在37°C下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱。

[0257] 实施例12

[0258] 用三聚体起始加帽寡核苷酸引物在体外转录

[0259] 表1指示了用于转录的一些起始加帽寡核苷酸引物。通过聚合酶链反应生成每个三聚体的特定双链萤火虫荧光素酶DNA转录模板。这些模板的不同之处在于模板核苷酸+1和+2。对于给定的三聚体,模板核苷酸+1与B₁ (该核苷酸注定成为转录物核苷酸+1) 互补。同样地,模板核苷酸+2与B₁₀ (该核苷酸注定成为转录物核苷酸+2;参见实施例10) 互补。转录反应与25ug/mL转录模板、40mM Tris-HCl (pH 8.0)、27mM MgCl₂、2mM亚精胺、10mM DTT、0.002% Triton x-100、1000单位/mL鼠RNA酶抑制剂 (New England Biolabs目录#M0314)、2单位/mL无机焦磷酸酶 (New England Biolabs目录#M2403)、4000单位/mL T7RNA聚合酶 (New England Biolabs目录#M0251) 和6mM起始加帽寡核苷酸引物、1.5mM GTP和各7.5mM的ATP、CTP和UTP组合。在随后的实施例中,该引物/NTP配制物将被称为引物/NTP配制物1。本领域技术人员清楚,其他聚合酶,诸如T7、T3或SP6RNA聚合酶,可以用于代替T7RNA聚合酶,通过使用其对应的启动子而执行相同的功能。转录反应混合物在37°C下温育2小时。用10mM Tris-HCl (pH 7.6)、2.5mM MgCl₂、0.5mM CaCl₂和100单位/mL DNA酶I (New England Biolabs目录#M0303) 补充反应,并且在37°C下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱,并通过将溶液调节至50mM Bis-Tris-Propane HCl (pH 6.0)、1mM MgCl₂、0.1mM ZnCl₂和250单位/mg南极磷酸酶 (New England Biolabs目录#M0289) 而去磷酸化。将反应在37°C下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱。

[0260] 实施例13

[0261] 用四聚体起始加帽寡核苷酸引物^{7m}G3' omc pppA2' omc pG2' omc pG在体外转录

[0262] 起始加帽寡核苷酸引物^{7m}G3' omc pppA2' omc pG2' omc pG用于转录。按照具有下列改变的实施例11进行转录。通过聚合酶链反应生成四聚体的特定双链DNA转录模板。模板核苷酸+1 (2'-脱氧胸苷) 与腺苷 (注定成为转录物核苷酸+1的引物第一个核苷酸) 互补。同样地,模板核苷酸+2 (2'-脱氧胞苷) 与鸟苷 (注定成为转录物核苷酸+2的引物第二核苷酸) 互补。同样地,模板核苷酸+3 (2'-脱氧胞苷) 与鸟苷 (注定成为转录物核苷酸+3的引物第三个核苷酸) 互补。转录反应与25ug/mL转录模板、40mM Tris-HCl (pH 8.0)、27mM MgCl₂、2mM亚精胺、10mM DTT、0.002% Triton x-100、1000单位/mL鼠RNA酶抑制剂 (New England Biolabs目录#M0314)、2单位/mL无机焦磷酸酶 (New England Biolabs目录#M2403)、4000单位/mL T7 RNA聚合酶 (New England Biolabs目录#M0251)、6mM起始加帽寡核苷酸引物^{7m}G3' omc pppA2' omc pG2' omc pG、1.5mM的GTP、7.5mM的ATP、7.5mM的UTP和7.5mM的CTP组合。本领域技术人员清楚,

其他聚合酶,诸如T7、T3或SP6RNA聚合酶,可以用于代替T7RNA聚合酶,通过使用其对应的启动子执行相同的功能。转录反应混合物在37℃下温育2小时。用10mM Tris-HCl (pH 7.6)、2.5mM MgCl₂、0.5mM CaCl₂和100单位/mL DNA酶I (New England Biolabs目录#M0303) 补充反应,并且在37℃下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱,并通过将溶液调节至50mM Bis-Tris-Propane HCl (pH 6.0)、1mM MgCl₂、0.1mM ZnCl₂和250单位/mg南极磷酸酶 (New England Biolabs目录#M0289) 而去磷酸化。将反应在37℃下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱。

[0263] 实施例14

[0264] 在模板位置+1和+2处分别具有2'-脱氧胸苷和2'-脱氧胞苷残基的转录模板上,用^{m7}GpppA_{2'}OmcpG起始加帽寡核苷酸引物在体外转录。

[0265] 使用分别在模板位置+1和+2处的2'-脱氧胸苷和2'-脱氧胞苷残基的双链萤火虫荧光素酶DNA转录模板一起。两个转录反应与25ug/mL转录模板、40mM Tris-HCl (pH 8.0)、27mM MgCl₂、2mM亚精胺、10mM DTT、0.002% Triton X-100、1000单位/mL鼠RNA酶抑制剂 (New England Biolabs目录#M0314)、2单位/mL无机焦磷酸酶 (New England Biolabs目录#M2403)、4000单位/mL T7RNA聚合酶 (New England Biolabs目录#M0251) 组合。两个转录的不同之处在于起始加帽寡核苷酸的量、NTP的量和NTP的特性。第一个转录设计为模拟Ishicawa等 (Nucleic Acids Symposium Series No.53:129 (2009)) 的转录条件。该转录反应含有6mM的^{m7}GpppA_{2'}OmcpG起始加帽寡核苷酸引物、0.9mM GTP和7.5mM的ATP、CTP和UTP。在随后的实施例中,该引物/NTP配制物将被称为引物/NTP配制物2。在该配制物中,起始寡核苷酸引物相对于GTP大于6倍过量。第二个转录使用5mM的^{m7}GpppA_{2'}OmcpG起始加帽寡核苷酸引物、5mM GTP、ATP、CTP和假尿苷三磷酸 (ΨTP)。在随后的实施例中,该引物/NTP配制物将被称为引物/NTP配制物3。在引物/NTP配制物3中,为了产生商业上有用量的RNA,增加了GTP浓度。本领域技术人员清楚,其他聚合酶,诸如T7、T3或SP6RNA聚合酶,可以用于代替T7RNA聚合酶,通过使用其对应的启动子执行相同的功能。将转录反应混合物在37℃下温育2小时。用10mM Tris-HCl (pH 7.6)、2.5mM MgCl₂、0.5mM CaCl₂和100单位/mL DNA酶I (New England Biolabs目录#M0303) 补充反应并在37℃下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱,并通过将溶液调节至50mM Bis-Tris-Propane HCl (pH 6.0)、1mM MgCl₂、0.1mM ZnCl₂和250单位/mg南极磷酸酶 (New England Biolabs目录#M0289) 而去磷酸化。对于引物/NTP配制物2,将反应在37℃下温育1小时,并且对于引物/NTP配制物3,将反应在37℃下温育3小时。将得到的mRNA根据制造商的说明使用RNeasyMaxi试剂盒 (Qiagen目录#75162) 纯化。将mRNA在水中洗脱。关于引物/NTP配制物1和2的纯化转录产率分别为0.7毫克/毫升转录反应 (mg/mL) 和3.9g/mL转录反应。关于引物/NTP配制物2的产率大大优于关于引物/NTP配制物1获得的产率。

[0266] 实施例15

[0267] 在模板位置+1和+2处具有胞苷残基的转录模板上,用^{m7}GpppA_{2'}OmcpG起始加帽寡核苷酸引物在体外转录

[0268] 使用双链萤火虫荧光素酶DNA转录模板,其中+1和+2模板核苷酸为胞苷,并且因此

不完全与^{m7}GpppA_{2'} omcpG起始加帽寡核苷酸引物互补。两个转录反应与25ug/mL转录模板、40mM Tris-HCl (pH 8.0)、27mM MgCl₂、2mM亚精胺、10mM DTT、0.002% Triton X-100、1000单位/mL鼠RNA酶抑制剂(New England Biolabs目录#M0314)、2单位/mL无机焦磷酸酶(New England Biolabs目录#M2403)、4000单位/mL T7RNA聚合酶(New England Biolabs目录#M0251)组合。两个转录的不同之处在于起始加帽寡核苷酸和NTP的量。第一个转录设计为模拟Ishicawa等(Nucleic Acids Symposium Series No.53:129(2009))的转录条件。该转录反应含有引物/NTP配制物2。第二个转录使用引物/NTP配制物3。本领域技术人员清楚,其他聚合酶,诸如T7、T3或SP6 RNA聚合酶,可以用于代替T7 RNA聚合酶,通过使用其对应的启动子执行相同的功能。将转录反应混合物在37°C下温育2小时,用10mM Tris-HCl (pH 7.6)、2.5mM MgCl₂、0.5mM CaCl₂和100单位/mL DNA酶I(New England Biolabs目录#M0303)补充反应并在37°C下温育1小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒(Qiagen目录#75162)或通过反相高效液相色谱法来纯化。通过将溶液调节至50mM Bis-Tris-Propane HCl (pH 6.0)、1mM MgCl₂、0.1mM ZnCl₂和250单位/mg南极磷酸酶(New England Biolabs目录#M0289),将mRNA去磷酸化。对于引物/NTP配制物2,将反应在37°C下温育1小时,并且对于引物/NTP配制物3,将反应在37°C下温育3小时。将得到的mRNA根据制造商的说明使用RNeasy Maxi试剂盒(Qiagen目录#75162)纯化。将mRNA在水中洗脱。关于引物/NTP配制物1和2的纯化转录产率分别为0.6毫克/毫升转录反应(mg/mL)和3.9g/mL转录反应。关于引物/NTP配制物2的产率大大优于关于引物/NTP配制物1获得的产率。

[0269] 实施例16

[0270] 在Huh-7细胞中翻译mRNA

[0271] (图11)

[0272] 在培养的肝细胞中评估用起始加帽寡核苷酸引物生成的荧光素酶mRNA的翻译活性,一式三份。将Huh-7细胞在37°C、5%CO₂的气氛下,在补充有10%FBS、L-谷氨酰胺、必需氨基酸和青霉素/链霉素的DMEM中培养。将细胞用400ng的mRNA转染。为了比较,将细胞也用Cap0荧光素酶mRNA转染,该Cap0荧光素酶mRNA通过用ARCA起始而生成。在20小时时,收获细胞并根据制造商的建议使用ONE-Glo荧光素酶试验系统试剂盒(Promega目录#E6120)测量荧光素酶活性。根据制造商的建议使用GloMax-Multi+检测系统仪器测量发光。检测所有测试的mRNA的荧光素酶活性(图11)。用起始加帽寡核苷酸引物或3'-O-甲基起始加帽寡核苷酸引物生成的mRNA的翻译效率类似于ARCA加帽的RNA的事实表明它们被有效地共转录地加帽了。

[0273] 实施例17

[0274] 确定通过使用ARCA或采用引物/NTP配制物1的起始加帽寡核苷酸引物共转录加帽而生成的mRNA的加帽效率的试验

[0275] (图12A-12H)

[0276] 对于测试的每种起始加帽寡核苷酸引物,将足以通过液相色谱质谱法(LC-MS)检测的量的mRNA进行加帽试验。在该试验中,从全长mRNA的5'端切割小片段并通过LC-MS分析。切割mRNA之前,将其用南极磷酸酶(New England Biolabs目录#M0289)处理以将无帽单磷酸、二磷酸和三磷酸转化为5' OH以促进分析。然后将磷酸酶处理过的mRNA切割并纯化。将纯化的RNA进行LC-MS分析。图12示出了LC迹线(traces)。将相应于无帽(磷酸酶处理后的5'

OH) 和帽1的LC峰以观察到的质量指示。插图示意图示出了起始寡核苷酸引物在转录模板上的比对。请注意,在示意图中“|”指示加帽起始核苷酸与模板核苷酸的一个碱基对。下标“m”指示2'-O-甲基且上标“m7”指示碱基甲基化。为了比较,将用ARCA共转录加帽的mRNA进行加帽试验。使用下式进行加帽效率的估计: (加帽峰的强度) / [(加帽峰的强度) + (5' OH峰的强度)]。在图12中观察到的%加帽为A) $m^7G_{3'}-OmepppG=79\%$ 、B) $m^7GpppG_{2'}-OmePG=89\%$ 、C) $m^7G_{3'}-OmepppG_{2'}-OmePG=87\%$ 、D) $m^7GpppA_{2'}-OmePG=99\%$ 、E) $m^7G_{3'}-OmepppA_{2'}-OmePG=99\%$ 、F) $m^7GpppC_{2'}-OmePG=98\%$ 、G) $m^7G_{3'}-OmepppC_{2'}-OmePG=97\%$ 、H) $m^7GpppA_{2'}-OmePG_{2'}-OmePG=50\%$ 。在每种情况中,用起始三聚体加帽寡核苷酸引物共转录加帽的转录物的加帽效率大于关于 $m^7G_{3'}-OmepppG$ (ARCA) 观察到的加帽效率。

[0277] 实施例18

[0278] 在模板位置+1和+2处具有2'-脱氧胸苷和2'-脱氧胞苷残基的转录模板上vs在模板位置+1和+2处具有胞苷残基的转录模板上用 $m^7GpppA_{2'}-OmePG$ 起始加帽寡核苷酸加帽的比较

[0279] (图13A-13D)

[0280] 将在实施例14和15中制成的mRNA进行加帽试验以确定在模板位置+1和+2处具有2'-脱氧胸苷和2'-脱氧胞苷残基vs在模板位置+1和+2处具有胞苷残基的转录模板的转录物的加帽相对效率和特异性。将足以通过液相色谱质谱法 (LC-MS) 检测的足量的mRNA进行加帽试验。在该试验中,从全长mRNA的5'端切割小片段并通过LC-MS分析。切割mRNA之前,将其用南极磷酸酶 (New England Biolabs目录#M0289) 处理以将无帽单磷酸、二磷酸和三磷酸转化为5' OH以促进分析。然后将磷酸酶处理过的mRNA切割并纯化。将纯化的RNA进行LC-MS分析。图13示出了LC迹线。将相应于无帽和帽1的LC峰以观察到的质量指示。插图示意图示出了起始寡核苷酸引物在转录模板上的比对。下标“m”指示2'-O-甲基且上标“m7”指示碱基甲基化。将图13A和13B转录物分别使用引物/NTP配制物2和引物/NTP配制物3、从在模板核苷酸+1和+2处分别具有2'-脱氧胸苷和2'-脱氧胞苷的模板转录。图13C和13D分别是使用引物/NTP配制物2和引物/NTP配制物3、从在模板核苷酸+1和+2处具有2'-脱氧胞苷的模板转录的转录物。当 $m^7GpppA_{2'}-OmePG$ 起始寡核苷酸与+1和+2核苷酸完全互补时,观察到的主要产物为用 $m^7GpppA_{2'}-OmePG$ 起始的期望的模板化的帽1转录物 (图12A和12B)。对于图13A和13B,使用以下式进行加帽效率的估计: (加帽峰的强度) / [(加帽峰的强度) + (5' OH峰的强度)],并且用引物/NTP配制物2和用引物/NTP配制物3的加帽效率分别为99%和96%。仅检测到少量异常的起始产物。相反,当模板核苷酸+1和+2为胞苷时,在 $m^7GpppA_{2'}-OmePG$ 起始寡核苷酸和这些模板核苷酸之间不存在完美的互补性。图12C和12D示出了以两种对位 (register) 起始的起始加帽寡核苷酸。在第一种对位中,3' 鸟苷起始加帽寡核苷酸残基与+1模板胞苷配对以产生具有另外的非模板化的5' 腺苷的转录物 (帽1+A)。请注意,在示意图中“|”指示加帽起始核苷酸与模板核苷酸的一个碱基对。指定的模板核苷酸位置被指示。在第二种对位中,3' 鸟苷起始加帽寡核苷酸残基与+2模板胞苷配对并且+1起始加帽寡核苷酸腺苷不与+1模板核苷酸形成完全的杂合体。这产生了模板化的5' 鸟苷已经被非模板化的腺苷替代的转录物 (帽1,G至A)。使用下式进行加帽效率的估计: (加帽峰“Cap1+A”+“Cap1 G至A”的强度) / [(加帽峰“帽1+A”+“帽1G至A”的强度) + (5' OH峰的强度)]。图13C和13D中计算的加帽效率为97%和77%。注意以下惊人的发现:当起始加帽寡核苷酸引物与相应的模板+1

和+2核苷酸完全互补时,加帽效率和模板化的转录起始保真度要高得多。另外,用我们的方法,可以在不降低驱动加帽的NTP的浓度的情况下实现有效的加帽,这允许高得多的转录产率。因此,我们描述的加帽方法不同于且优于Ishikawa等的加帽方法。

[0281] 实施例19

[0282] (图14A-14B)

[0283] 在模板位置+1和+2处具有2'-脱氧胸苷和2'-脱氧胞苷残基的转录模板上vs在模板位置+1和+2处具有胞苷残基的转录模板上用^{m7}GpppA_{2'} omc pG起始加帽寡核苷酸制成的mRNA在分化的THP-1细胞中的翻译的比较

[0284] 为了评估在实施例14和15中生成的荧光素酶mRNA的表达,将mRNA转染到THP-1细胞(ATCC,目录#TIB-202)中,一式六份(复本)。在补充有10%FBS、丙酮酸钠和青霉素/链霉素的ATCC配制的RPMI-1640(ATCC,目录#30-2001)中,在37℃、5%CO₂的气氛下培养THP-1细胞。在佛波醇酯12-O-十四烷酰佛波醇-13-乙酸酯(TPA;Cell Signaling Technologies,目录#4174)的存在下,将细胞接种在24-孔板格式(每孔2E+05个细胞)中以诱导分化。接种后72小时,将细胞用每孔100ng的mRNA转染。为了比较,还用Cap0荧光素酶mRNA转染细胞,该Cap0荧光素酶mRNA通过用ARCA起始而生成。在转染后20小时,收获细胞并根据制造商的建议使用ONE-Glo荧光素酶试验系统试剂盒(Promega目录#E6120)测量荧光素酶活性。根据制造商的建议使用GloMax-Multi+检测系统仪器测量发光。将作为六份复本的平均值+/-平均值标准偏差的数据制图。使用非配对t-检验分析数据以生成p值作为显著性的度量。指示了在对之间的p-值。比较了在THP-1细胞中用^{m7}GpppA_{2'} omc pG起始加帽寡核苷酸并且用在模板位置+1和+2处包含2'-脱氧胸苷和2'-脱氧胞苷残基的模板("TC"模板)vs在模板位置+1和+2处具有胞苷残基的转录模板("CC"模板)生成的转录物的翻译。用A)引物/NTP配制物2或B)引物/NTP配制物3进行了转录。通过这两种配制物(配方),当^{m7}GpppA_{2'} omc pG起始加帽寡核苷酸引物如本发明中所述与模板核苷酸+1和+2("TC"模板)完全互补时,在培养的THP-1细胞中的翻译都是显著更优的。在图14B中,我们还评估了用"TC"模板制成的并用ARCA(帽0)加帽的mRNA的活性。这是用于产生共转录加帽mRNA的当前工业标准。ARCA加帽的RNA具有比在^{m7}GpppA_{2'} omc pG起始加帽寡核苷酸引物与模板核苷酸+1和+2("TC"模板)完全互补情况下制成的转录物显著更低的活性。

[0285] 总之,这里所示的实施例证明,相对于先前公开的方法,这里所描述的方法生成具有以下组合的RNA:1)高产率、2)高加帽程度、3)模板化的转录的高保真度,以及4)在细胞中更优的活性。

[0286] 本文说明性描述的发明可以在不存在本文没有具体地公开的任何一个或多个要素、一个或多个限制的情况下适当地实践。因此,例如,在描述本发明的上下文中(尤其在所附权利要求的上下文中),术语"一个"、"一种"和"该(所述,the)"和类似的语言将被解释为覆盖单数和复数,除非本文另外指示或与上下文明显矛盾。术语"包含"、"具有"、"包括"、"含有"等应当被广义地解读而没有限制(例如,意为"包括但不限于")。本文中数值范围的列举仅仅是旨在充当个体地指代落入该范围的每个单独值的速记法,除非本文另外指示,并且每个单独值并入说明书中,如同其在本文中被个体地列举一样。本文所描述的所有方法可以以任何适当的顺序进行,除非本文另外指示或另外与上下文明显矛盾。本文提供的任何和所有实施例或示例性语言(例如,"诸如")的使用仅仅旨在更好阐明本发明而不是造

成对本发明的范围的限制,除非另外声明。说明书中的任何语言都不应该被解释为表示任何未声明的要素是对本发明实践所必需的。另外,本文所采用的术语和表达已经用作描述而不是限制的术语,并且在使用这种术语和表达时不意排除示出和描述的特征的任何等价物或其部分,而是要知道本发明所要求保护的范围内可以进行各种修改。因此,应该理解虽然已经参照某些实施方式和任选的特征具体地公开了本发明,但本领域技术人员可以采取本文所公开的其中实施的发明的修改和变化,并且这种修改和变化被认为是在本发明的范围内。因此,应该理解虽然已经通过参照某些实施方式和任选的特征具体地公开了本发明,但本领域技术人员可以采取本文所公开的其中实施的发明的修改、改进和变化,并且这种修改、改进和变化被认为是在本发明的范围内。本文所提供的材料、方法和实例是某些实施方式的代表,是示例性的,而不是旨在作为对本发明的范围的限制。

[0287] 本文已经宽泛地和总体地描述了本发明。落入上位公开的每个窄类和下位组也构成本发明的部分。这包括具有将任何主题从大类排除的附带条件或消极限制的一般性描述——无论该排除物是否在本文中被具体列举。

[0288] 另外,在就马库什组描述本发明的特征或方面的情况下,本领域技术人员将认识到,本发明还由此就马库什组的任何个体成员或成员子组被描述。

[0289] 本文提及的所有出版物、专利申请、专利和其他参考文献其整体均通过引用明确地并入,如同每一个均个体地通过引用而并入。在冲突的情况下,以本说明书(包括定义)为准。

[0290] 申请人保留从任何这种文章、专利、专利申请或其他物理的或电子的文件将任何和所有材料和信息物理上并入该申请的权利。

[0291] 其他实施方式在所附权利要求中提出。

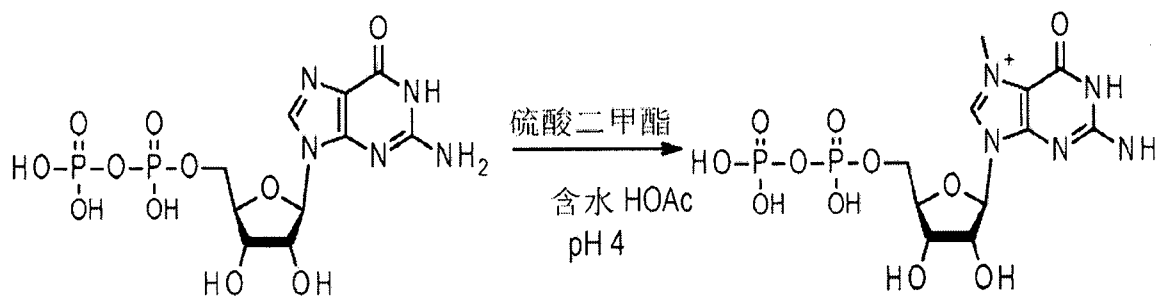


图1

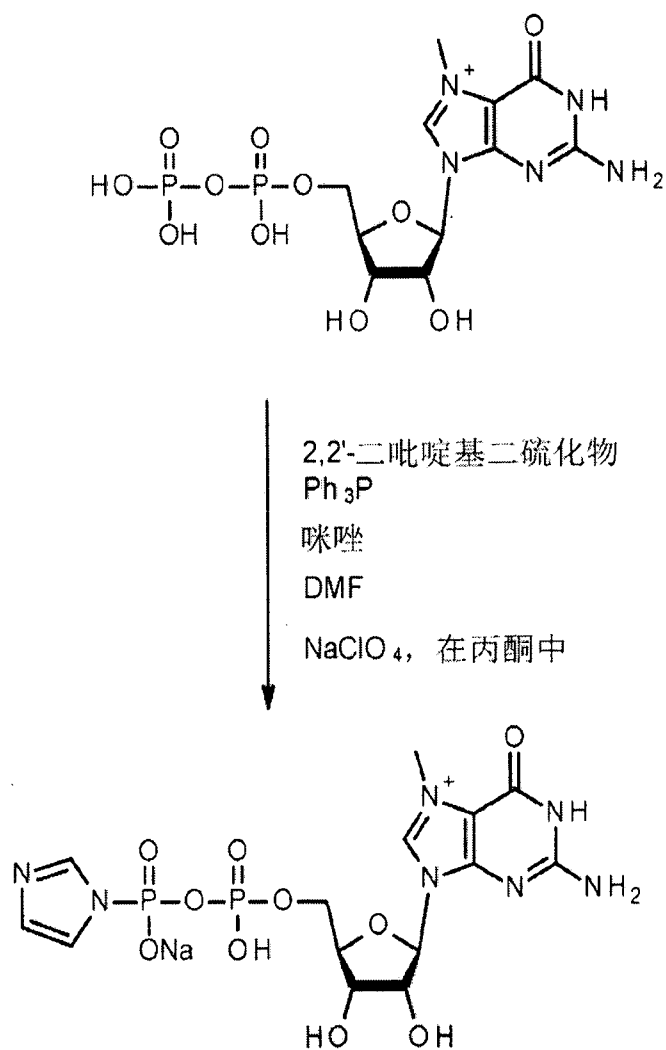


图2

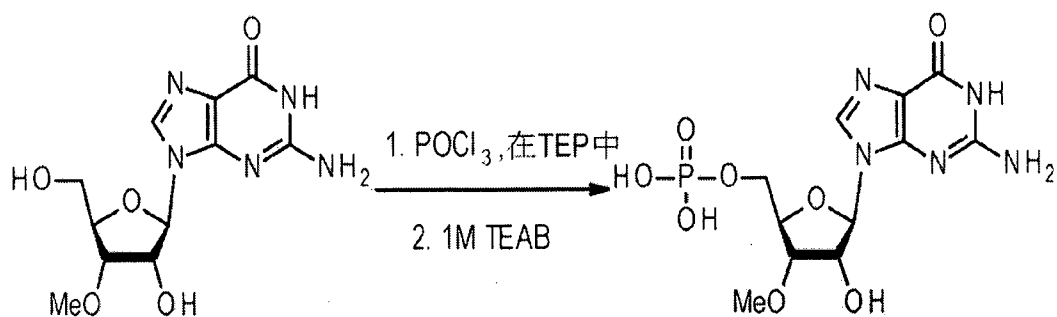


图3

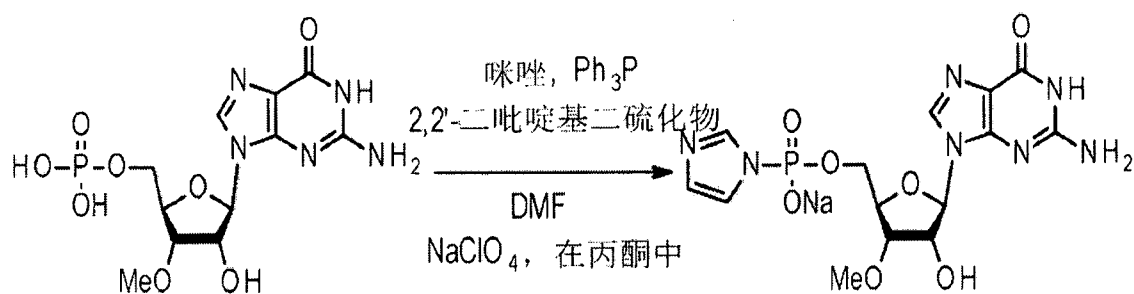


图4

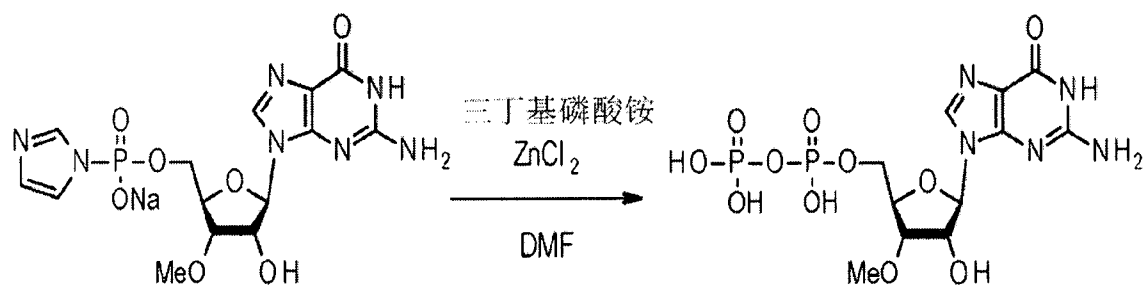


图5

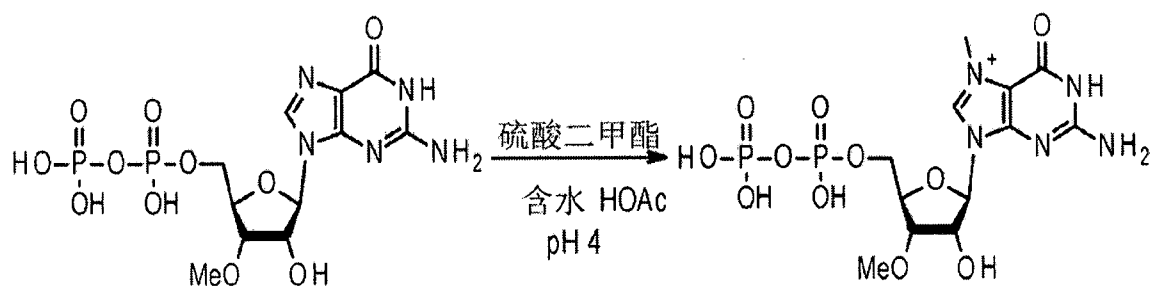


图6

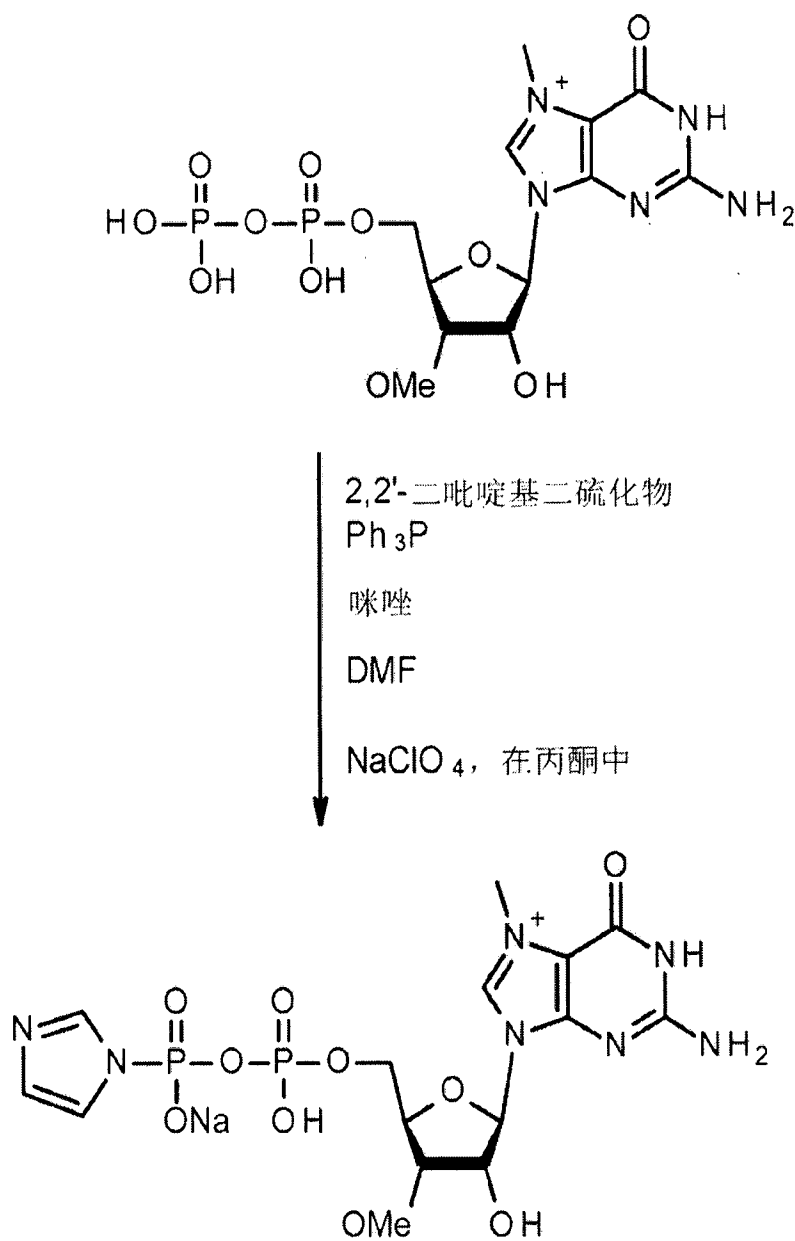
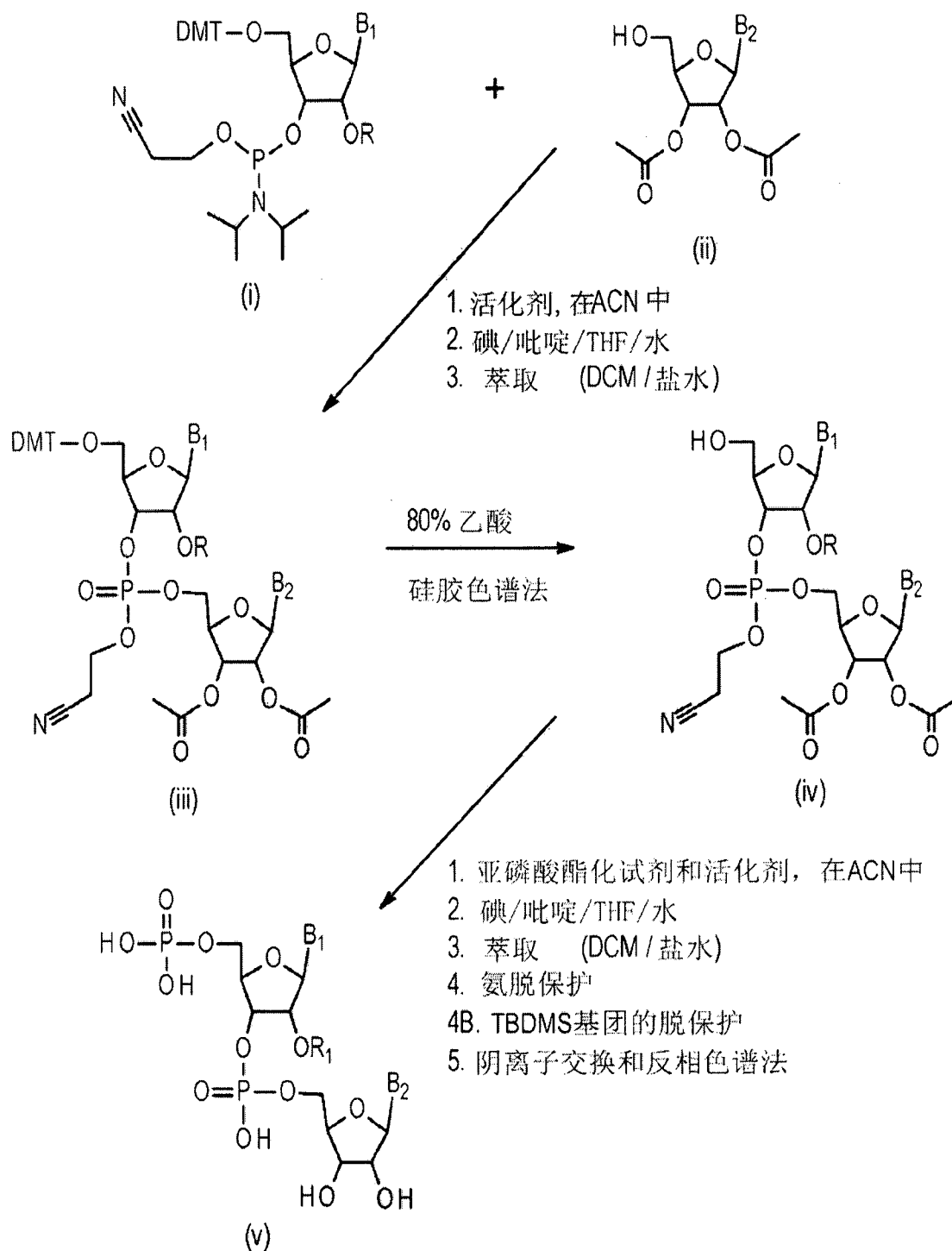


图7



B1 和 B2 是适当地保护的或未保护的核苷碱基
 R = Me 或 TBDMS; R₁ = Me 或 H

图8

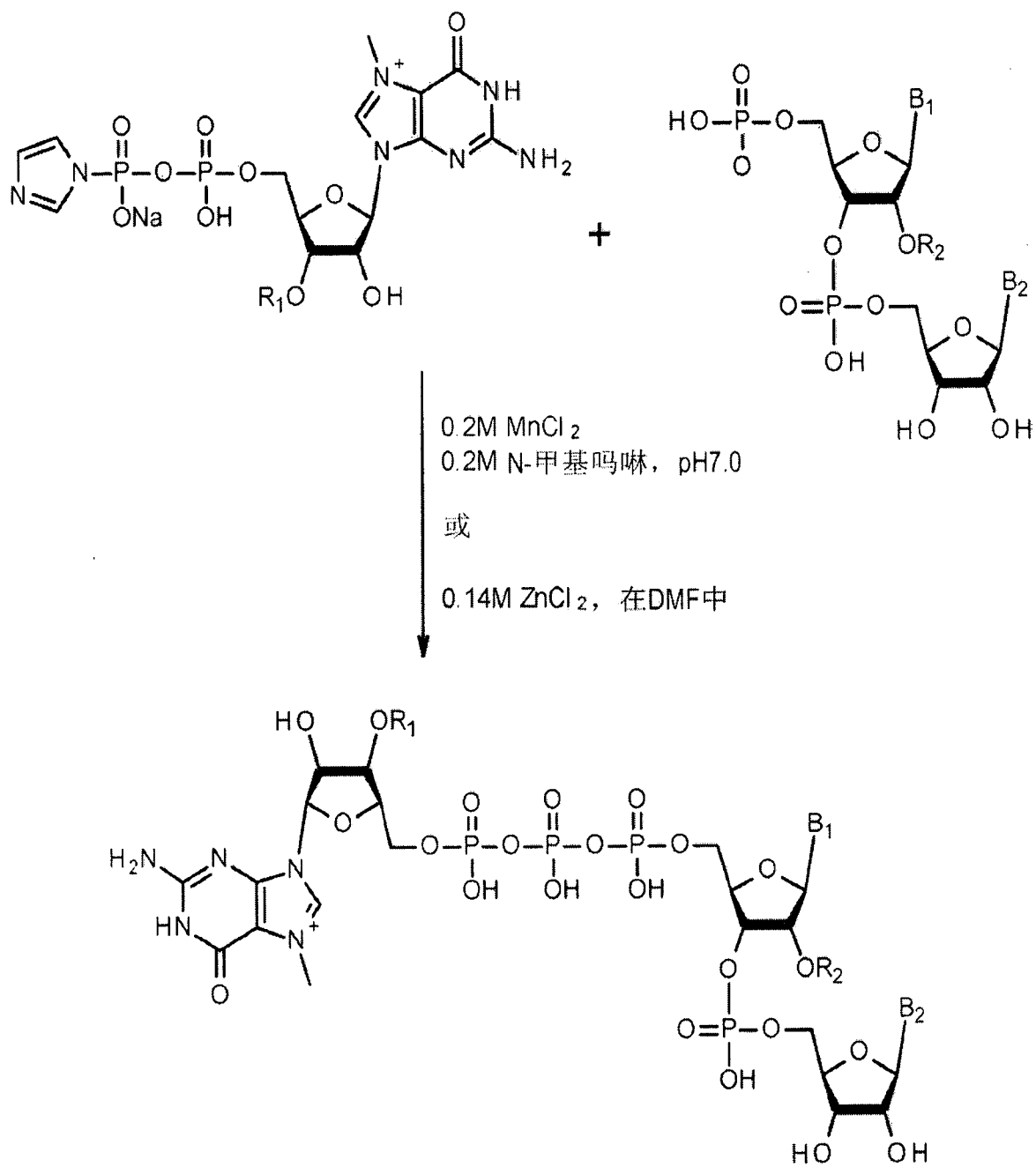


图9

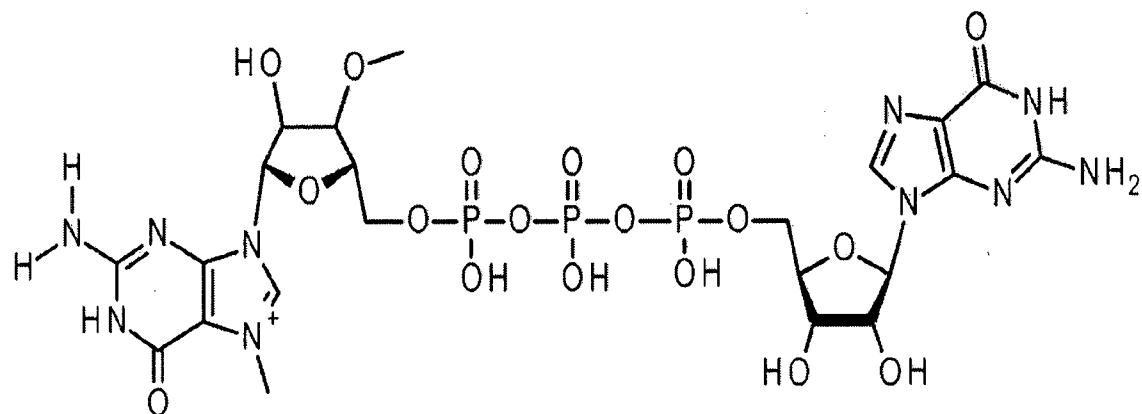


图10A

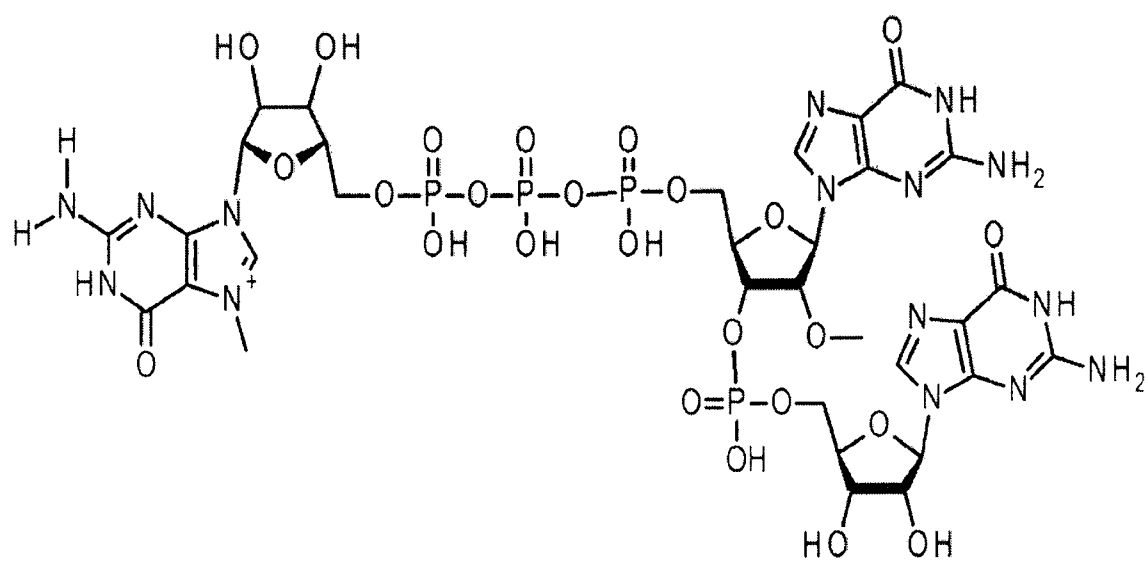


图10B

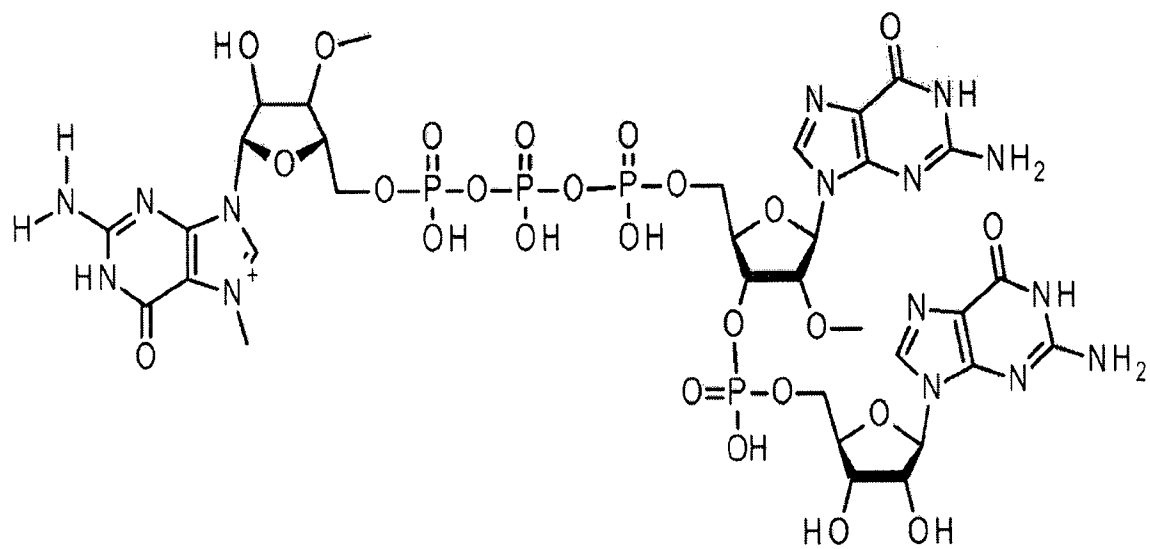


图10C

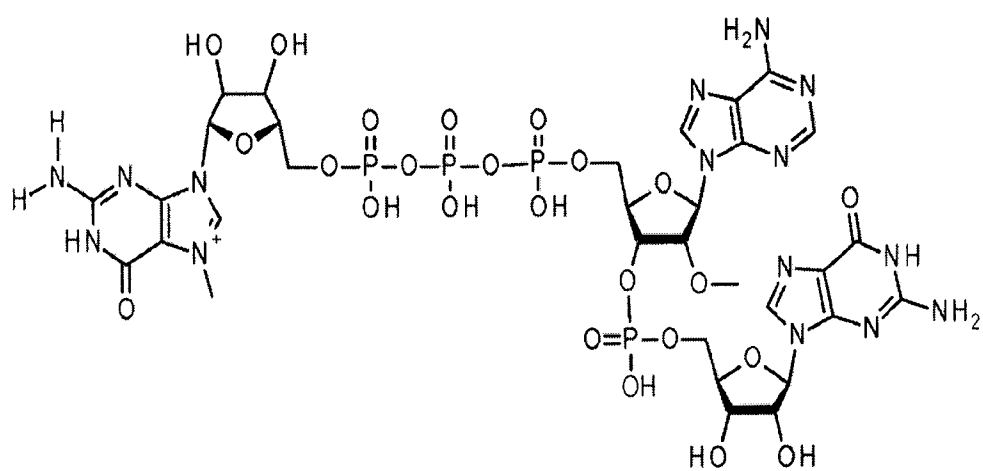


图 10D

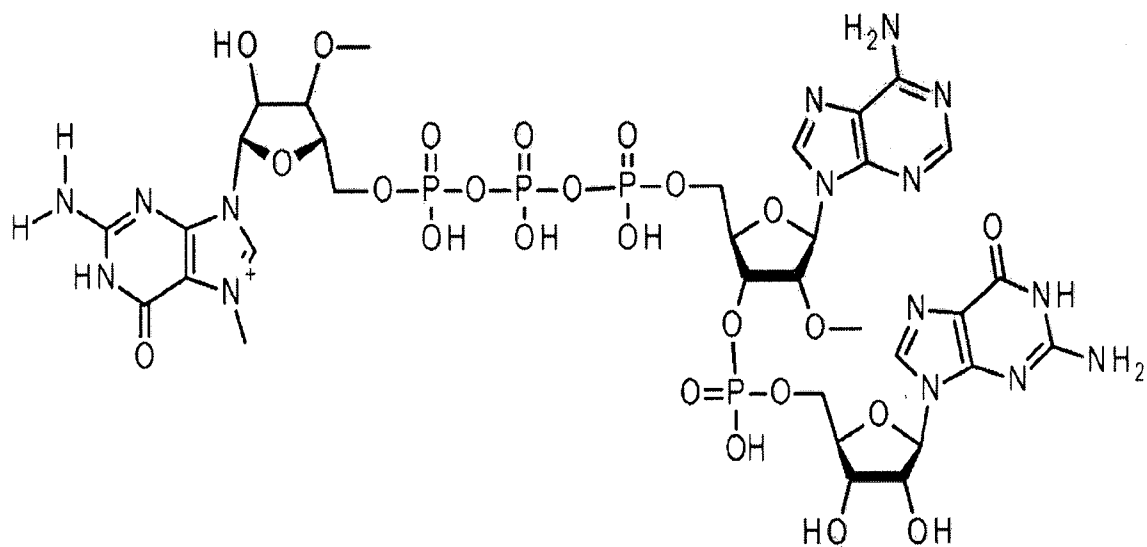


图10E

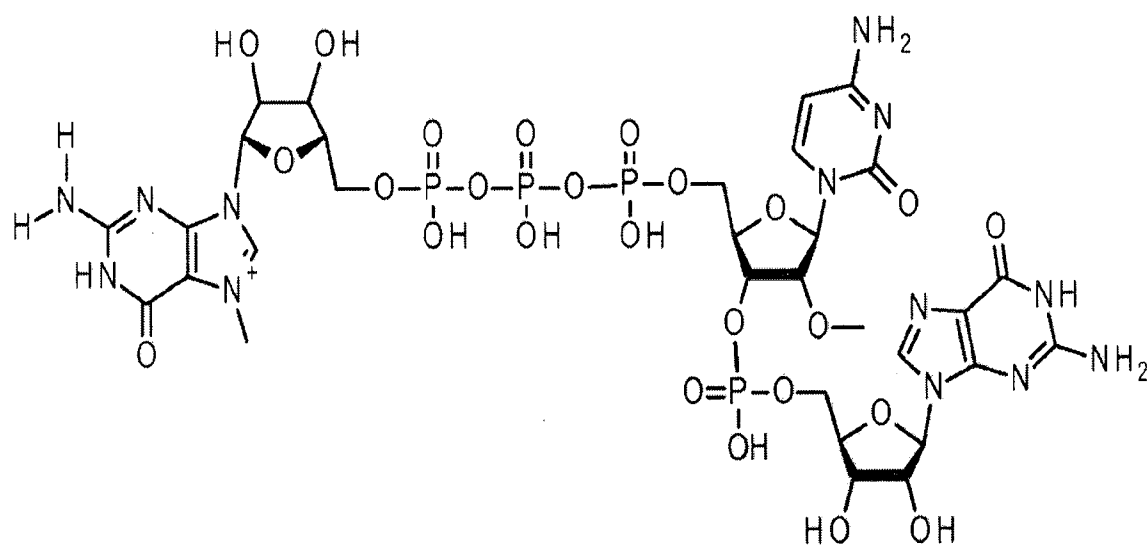


图10F

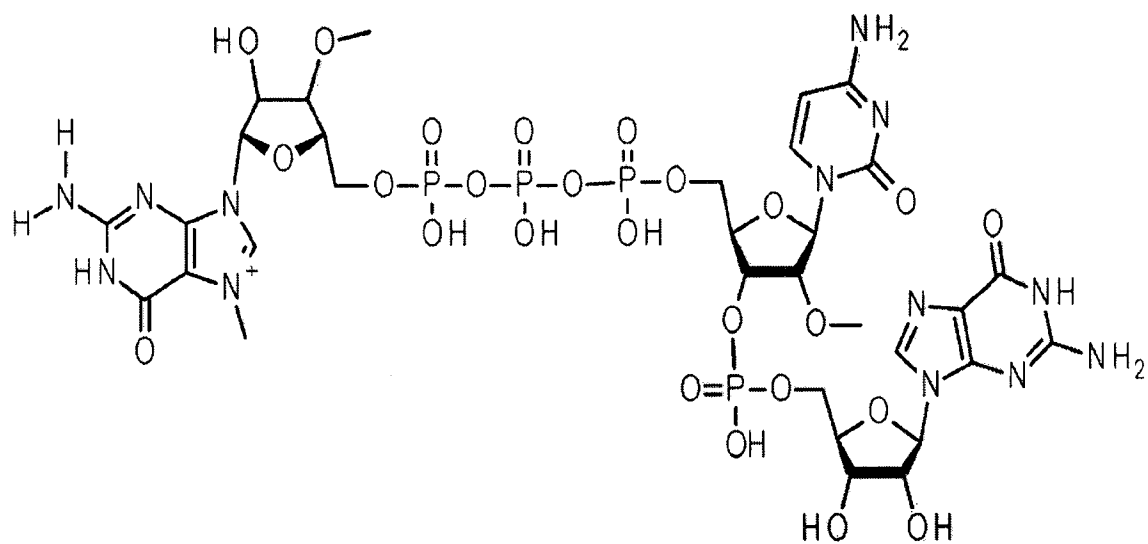


图10G

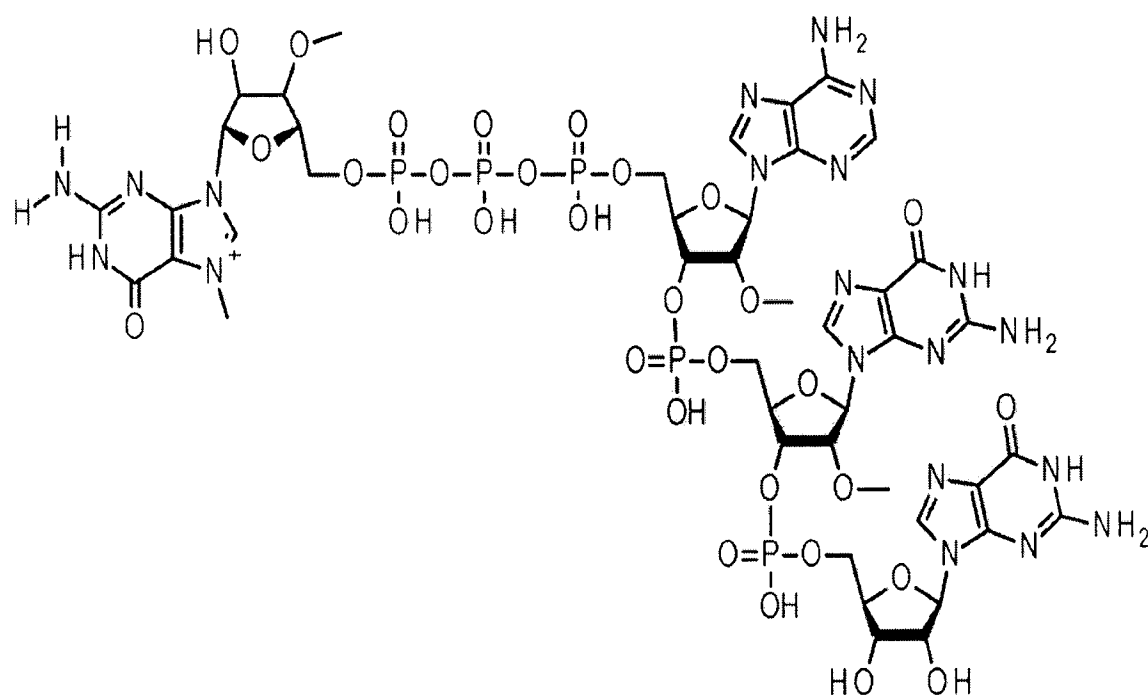


图10H

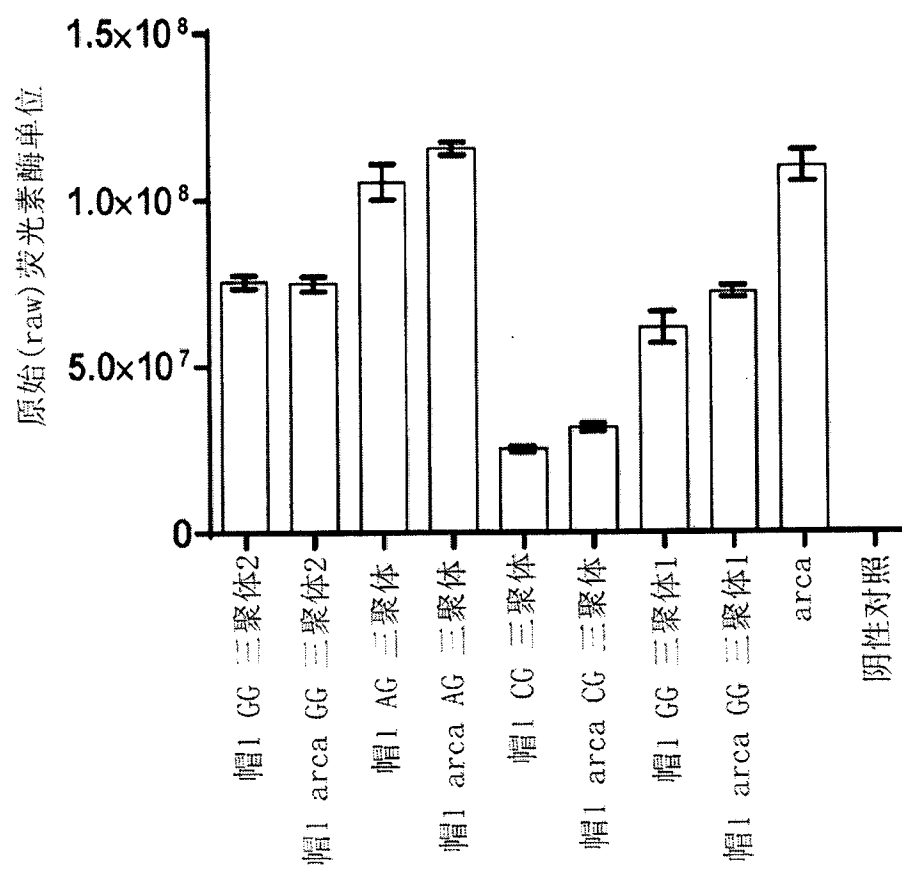


图11

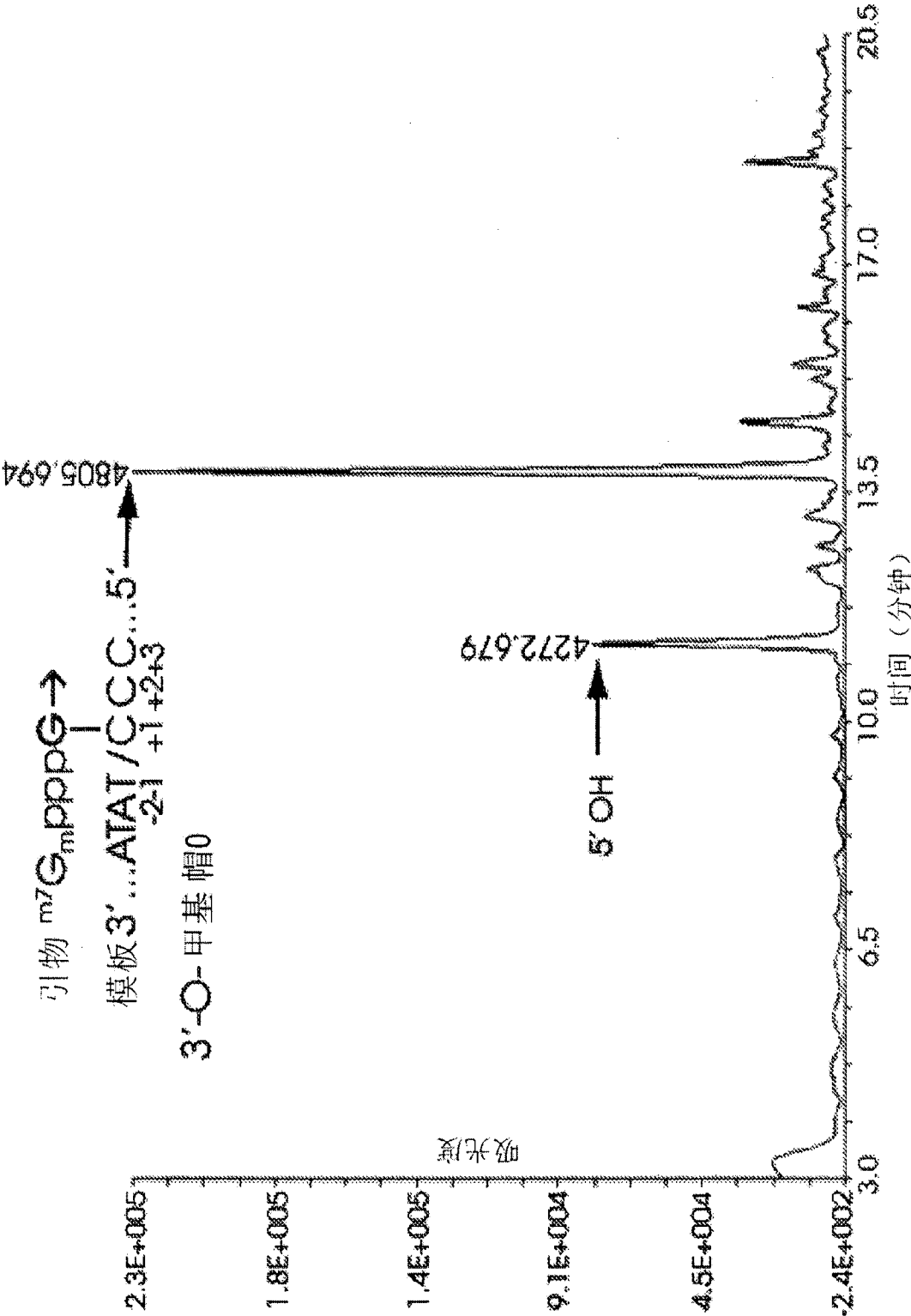


图12A

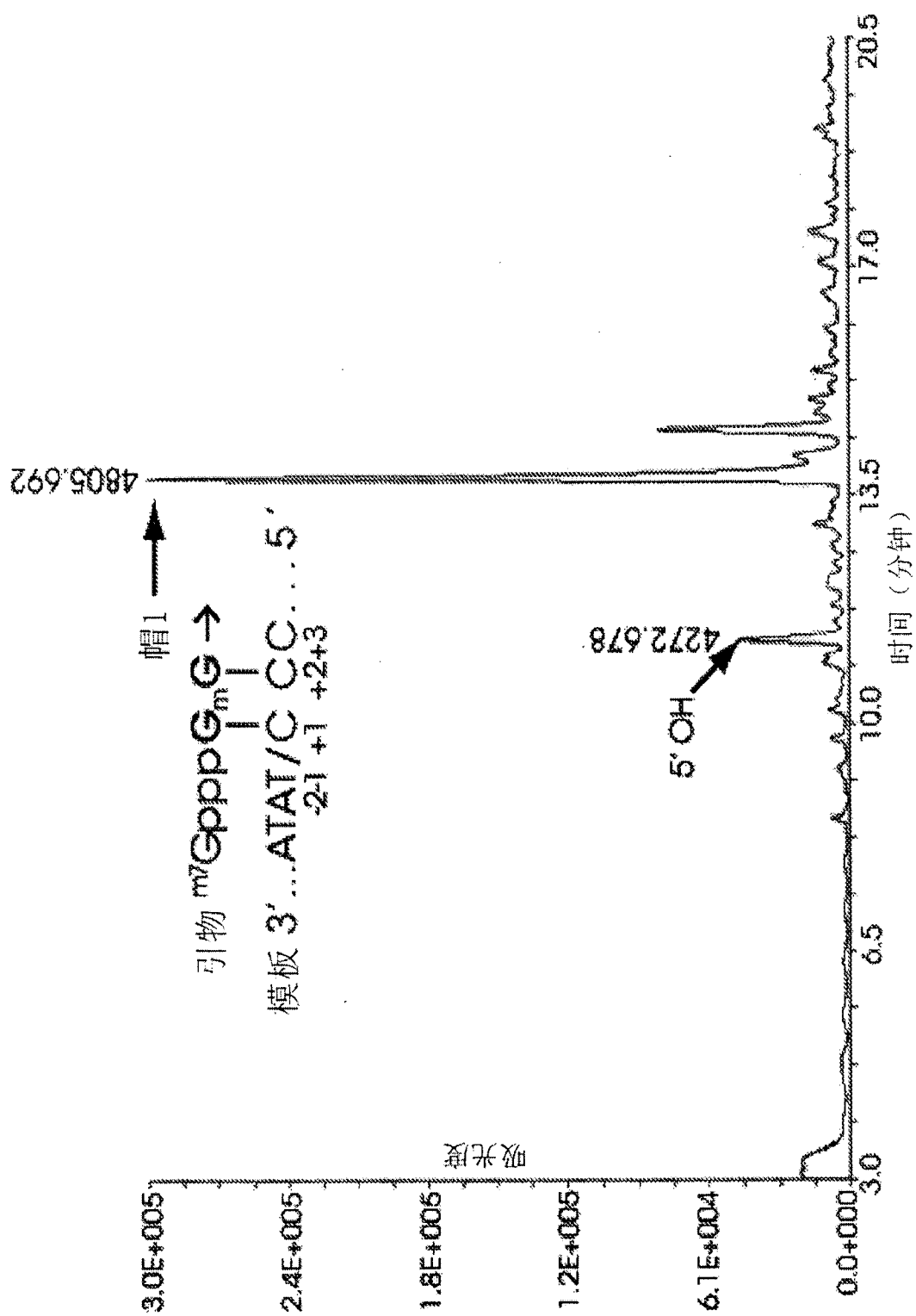


图12B

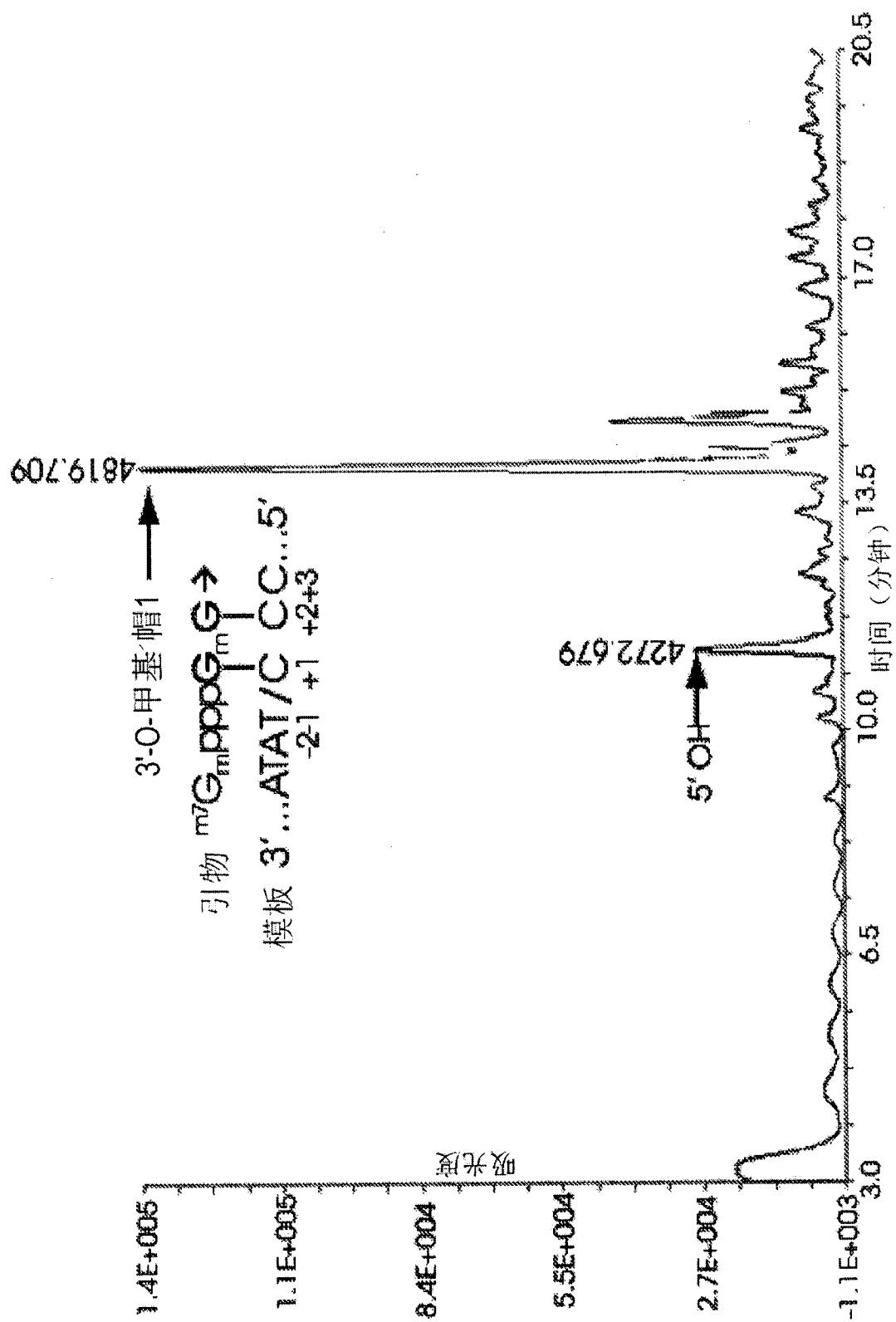


图12C

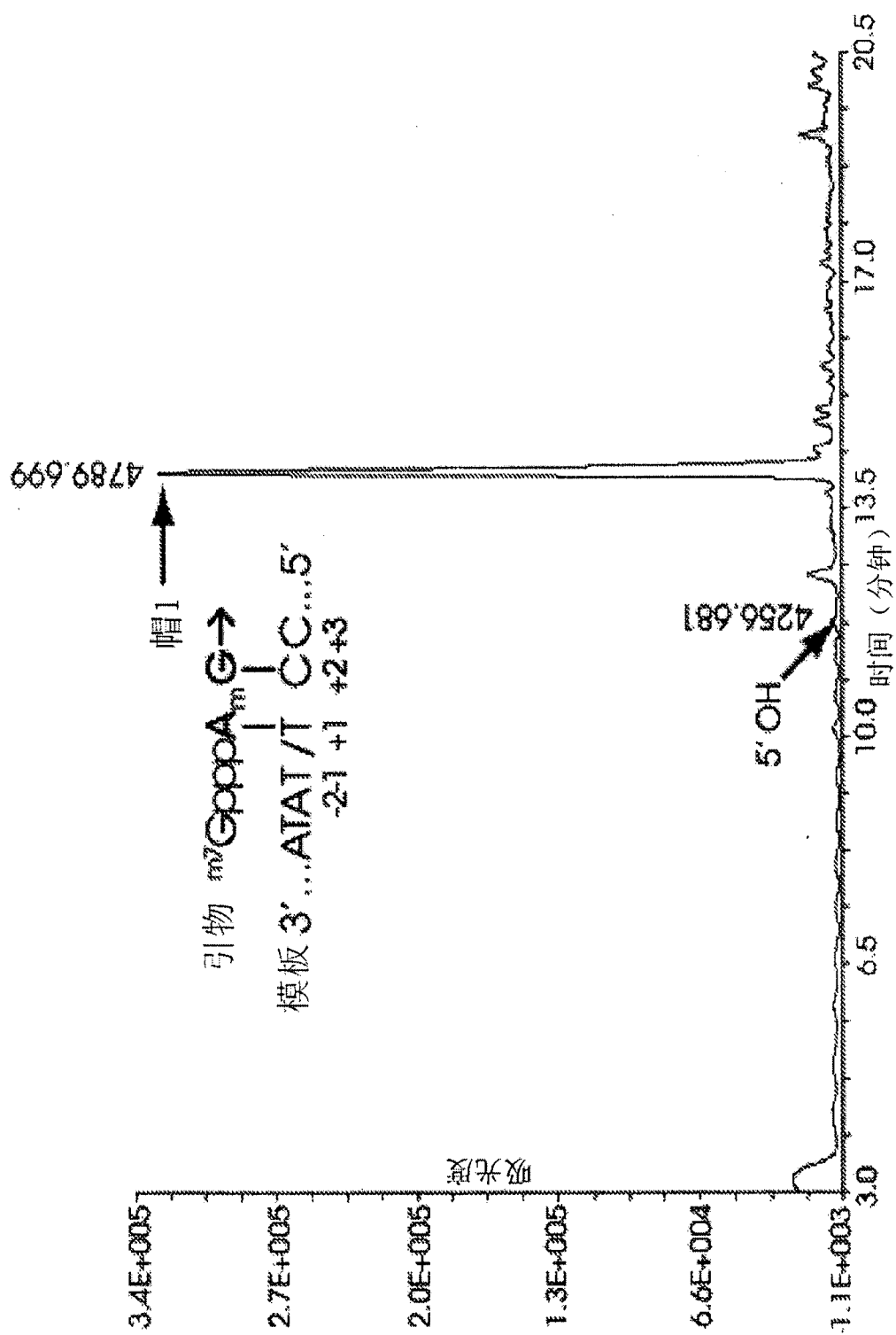


图12D

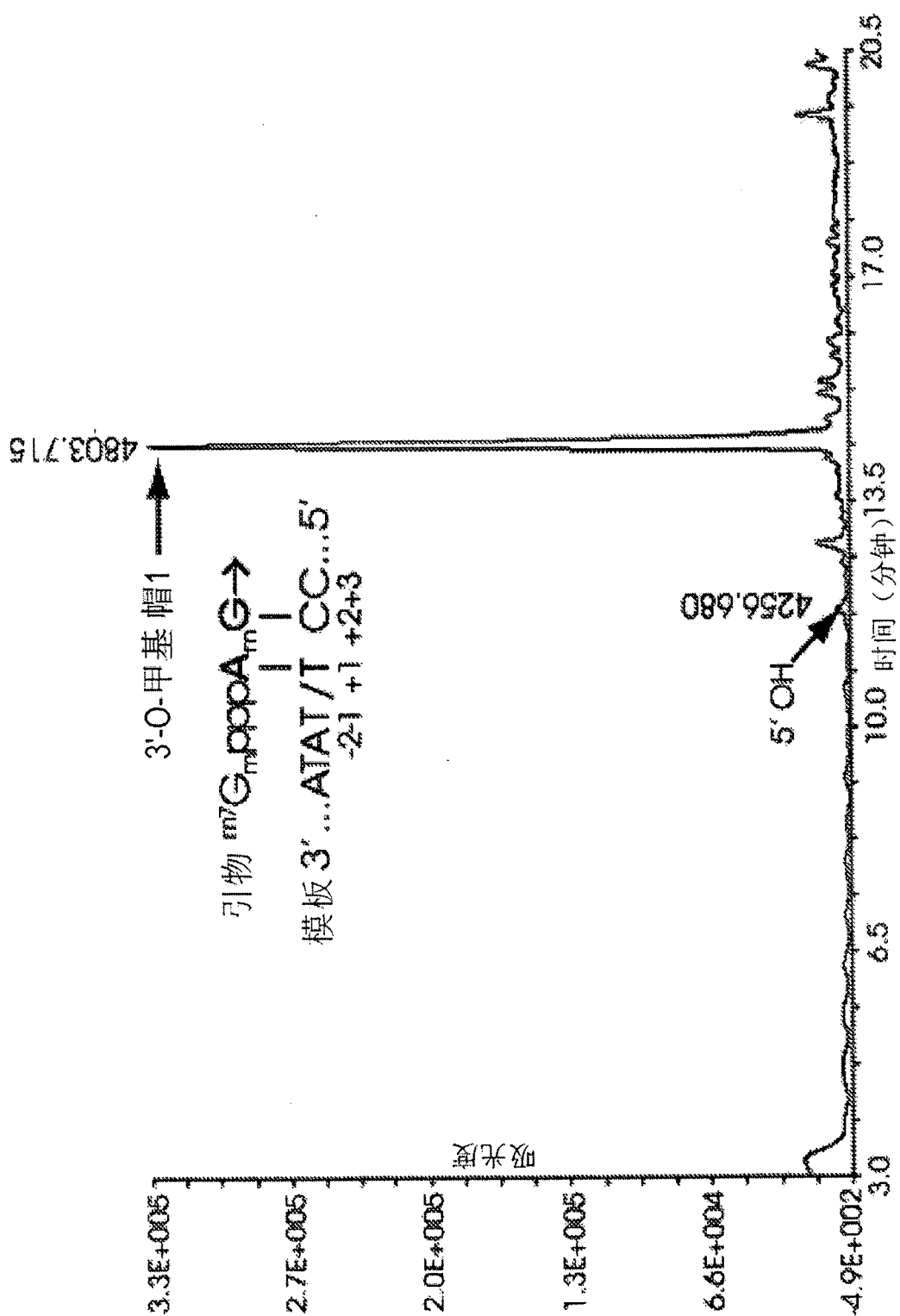


图12E

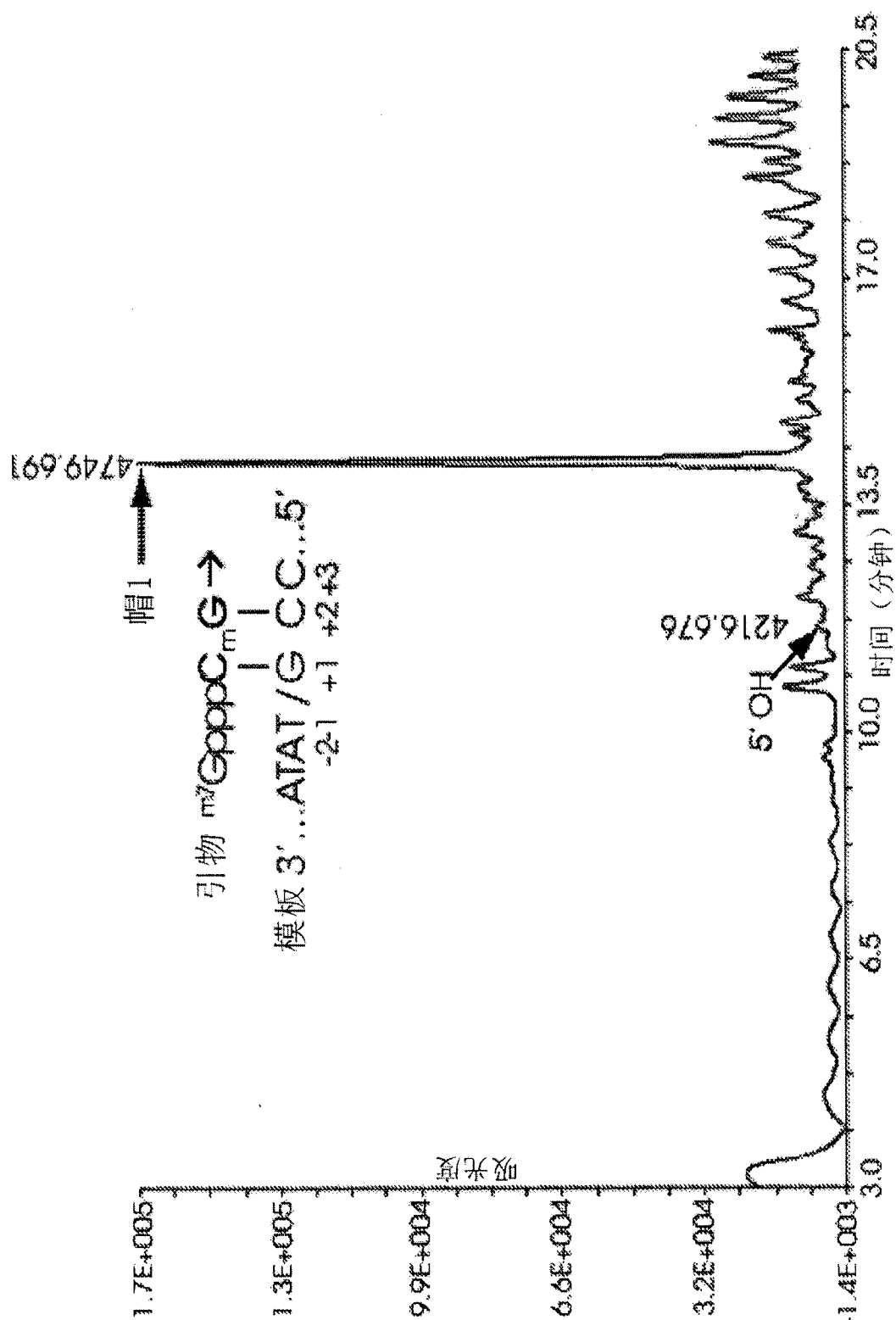


图12F

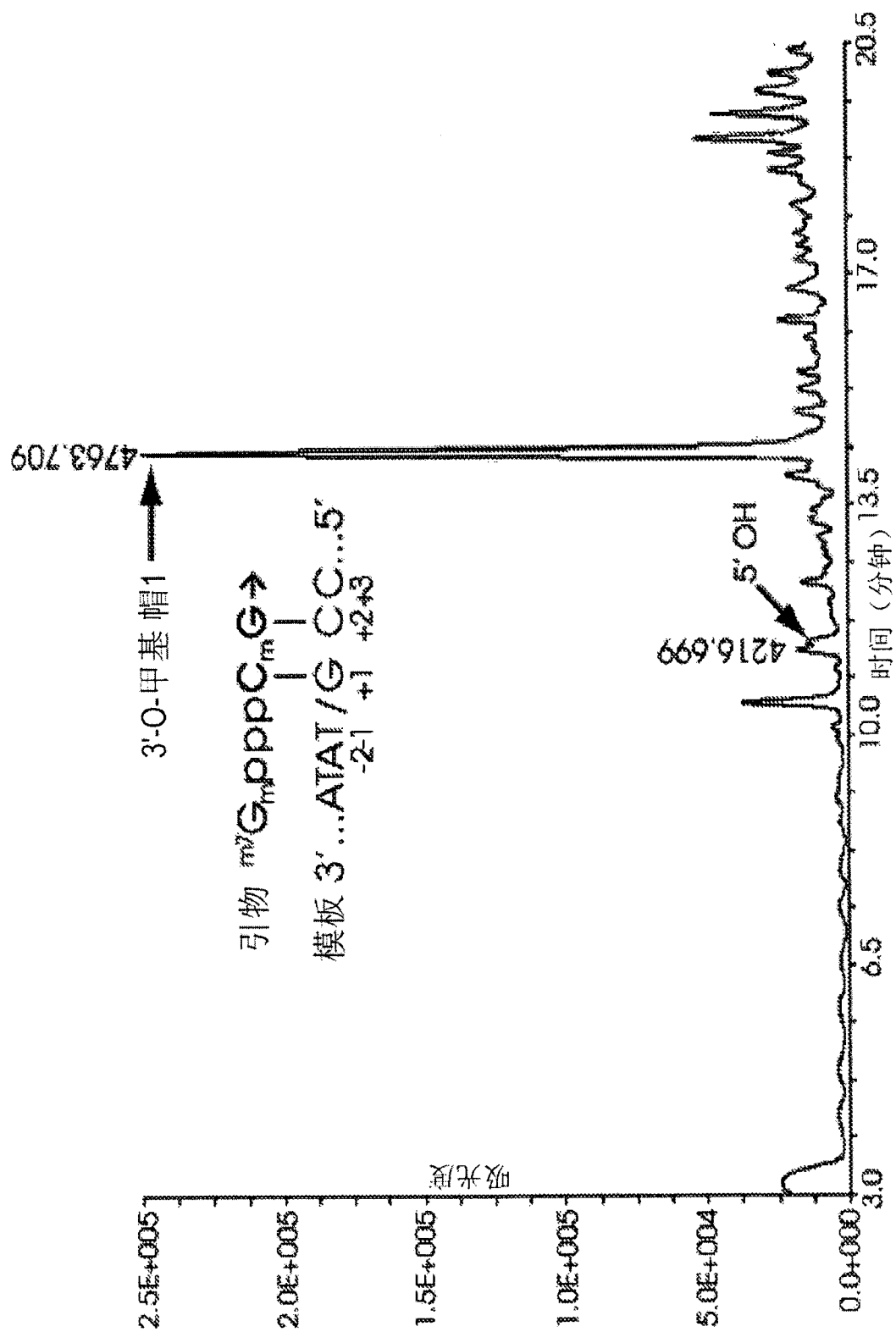


图12G

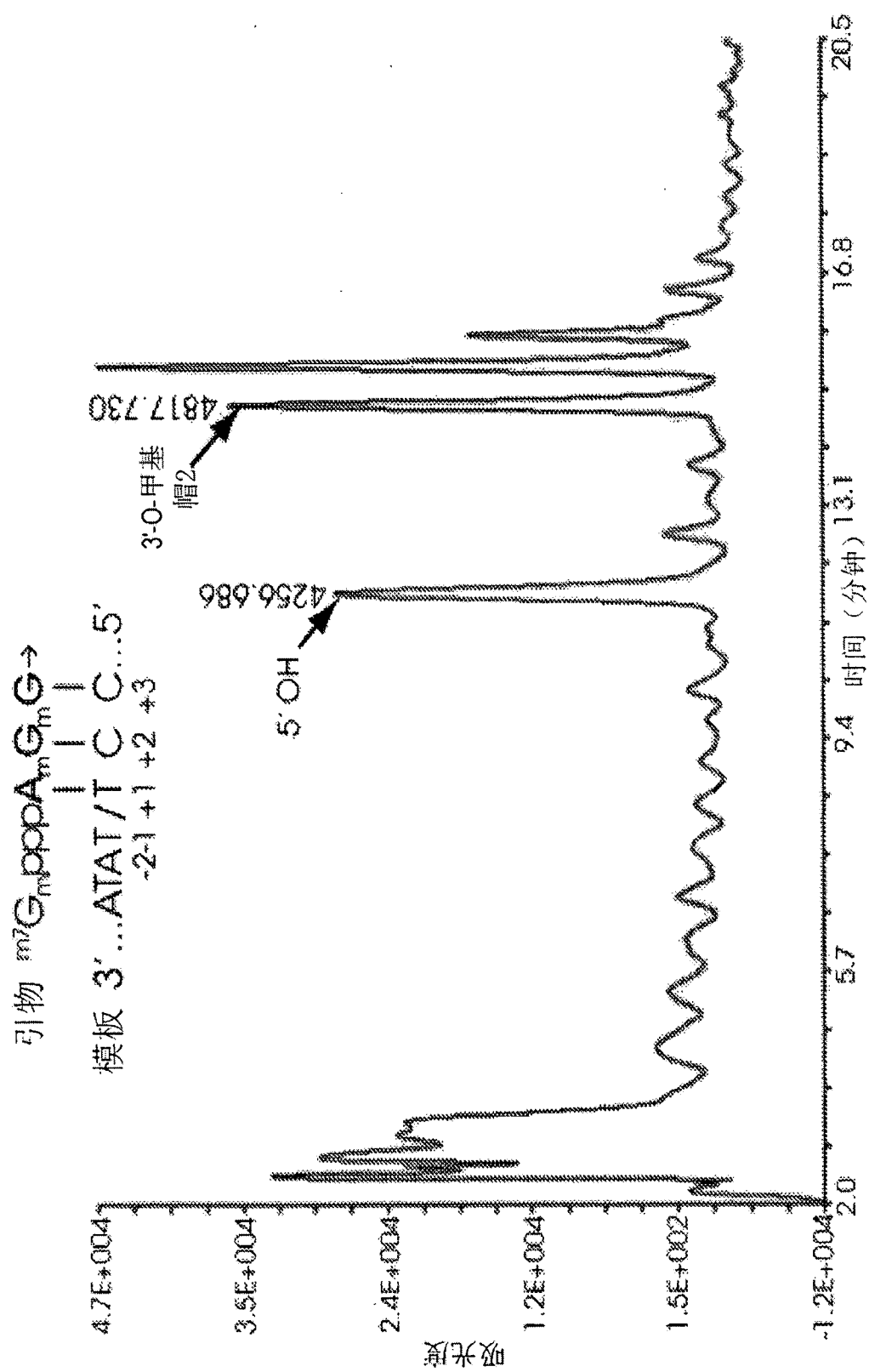


图12H

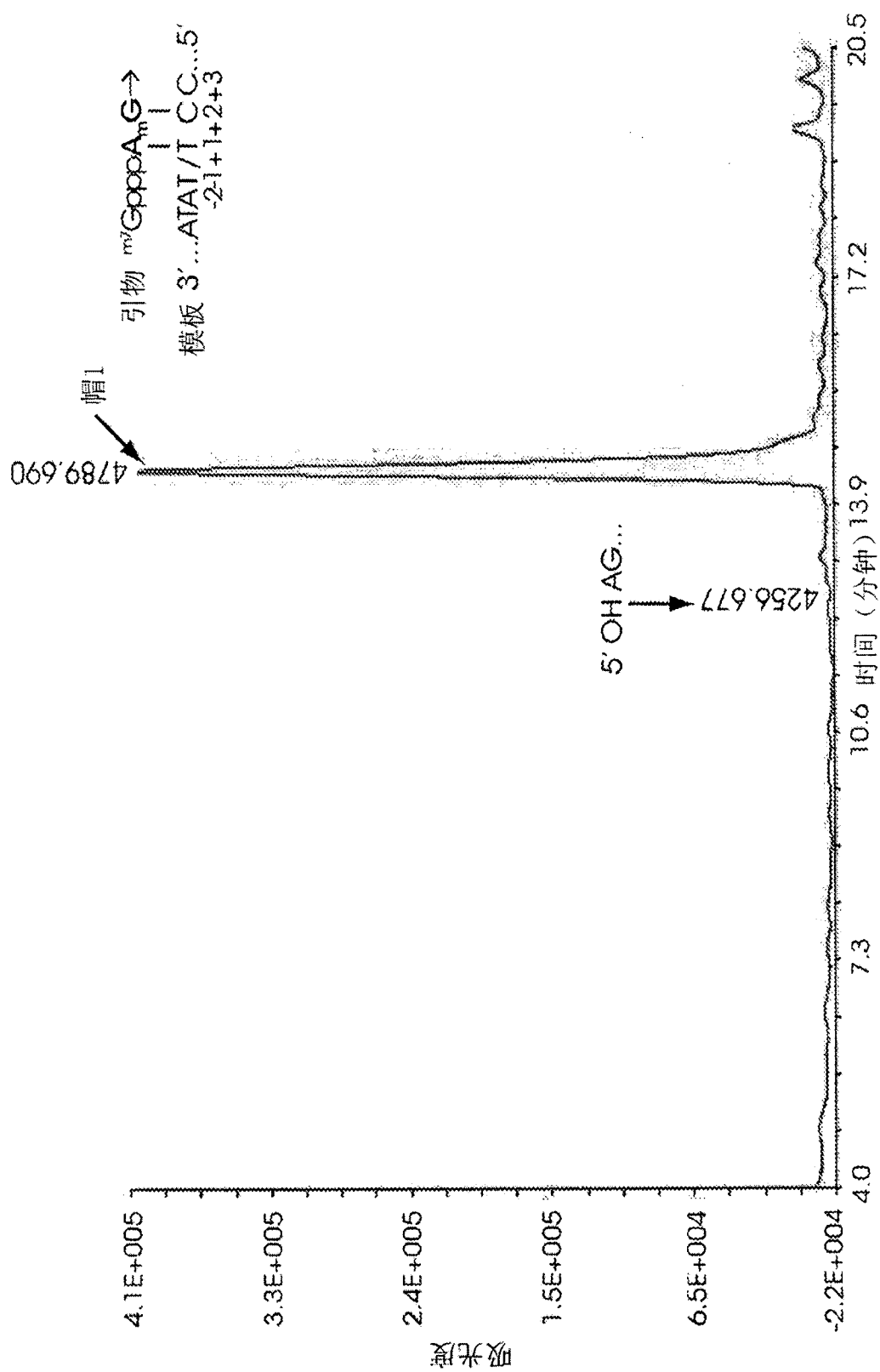


图13A

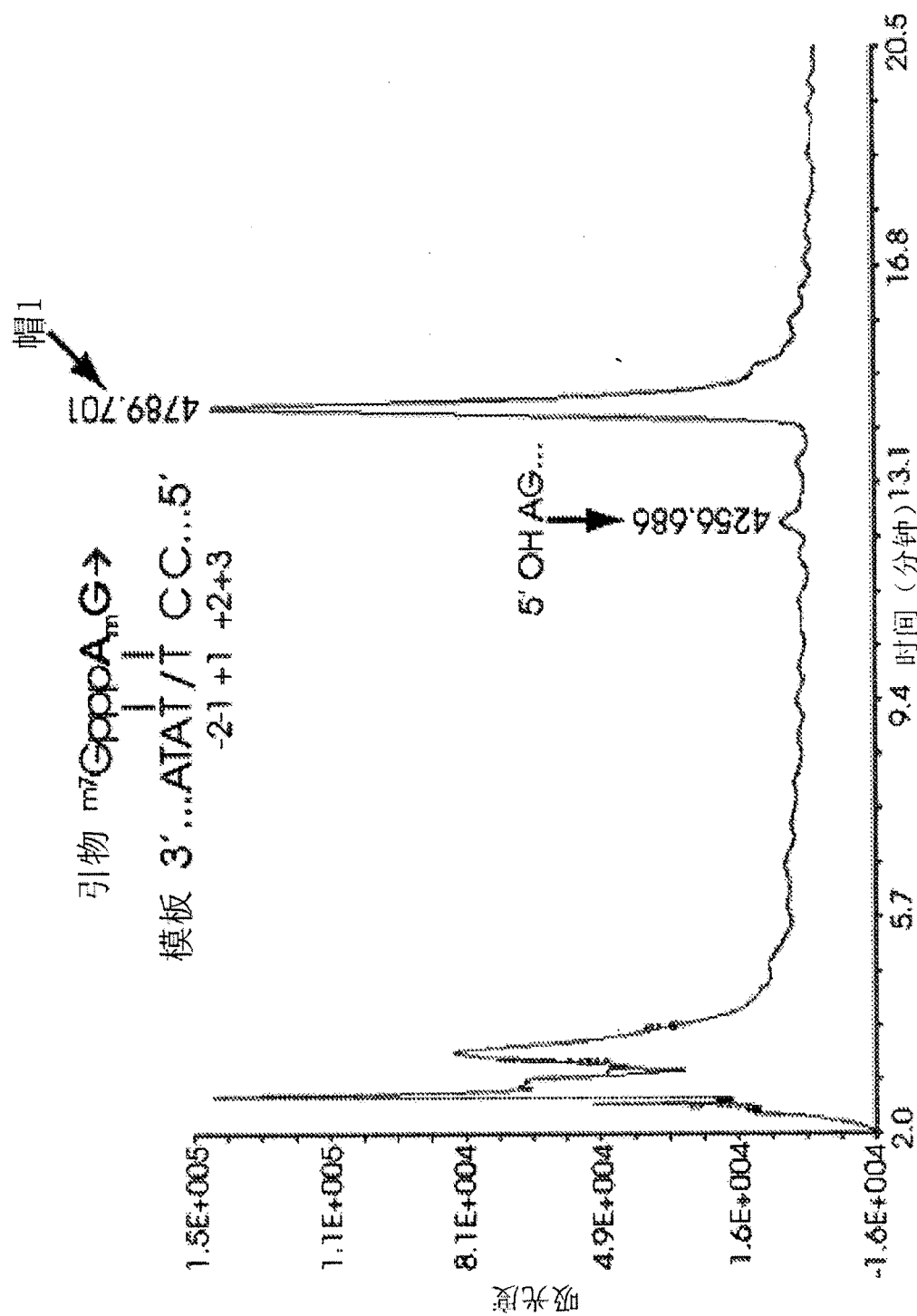


图13B

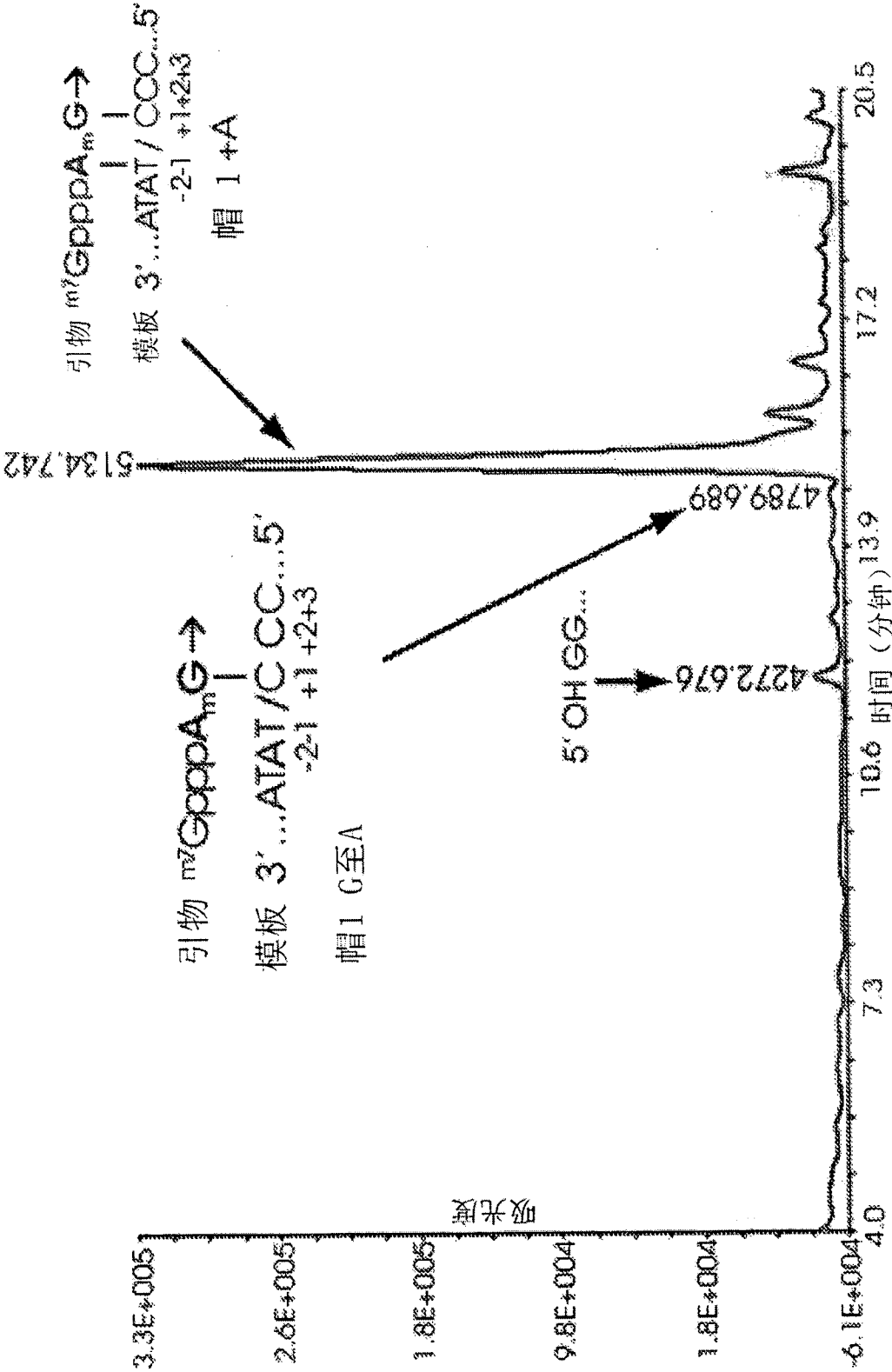


图13C

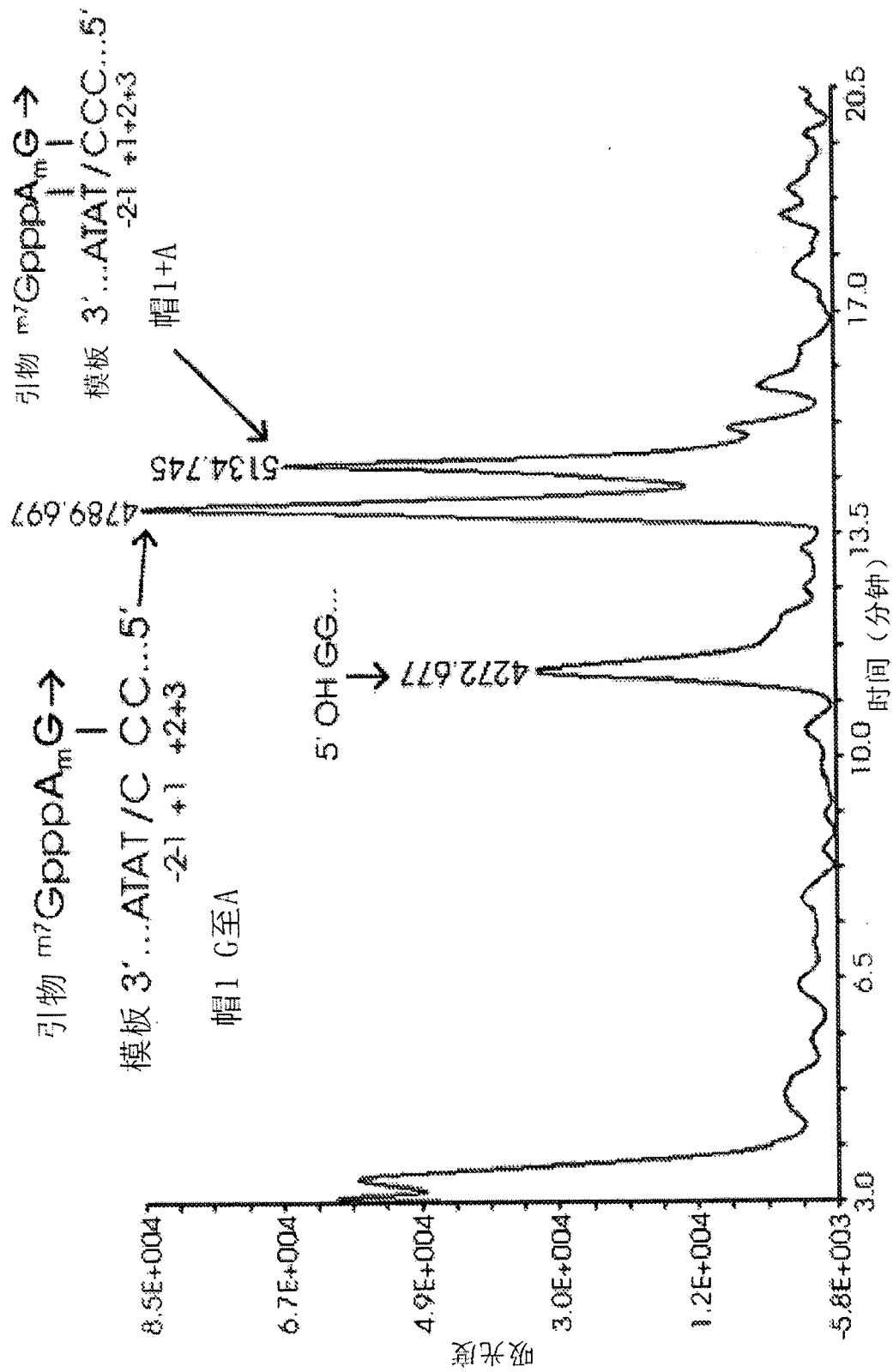


图13D

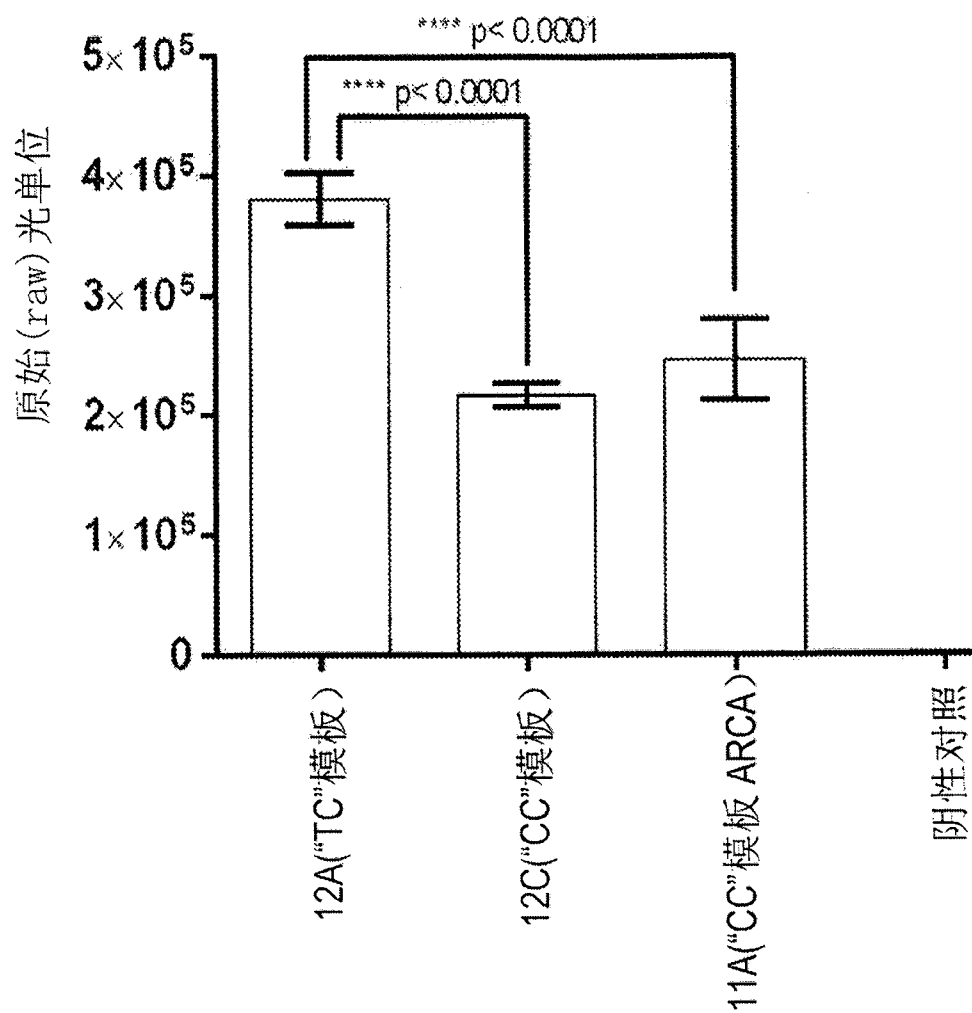


图14A

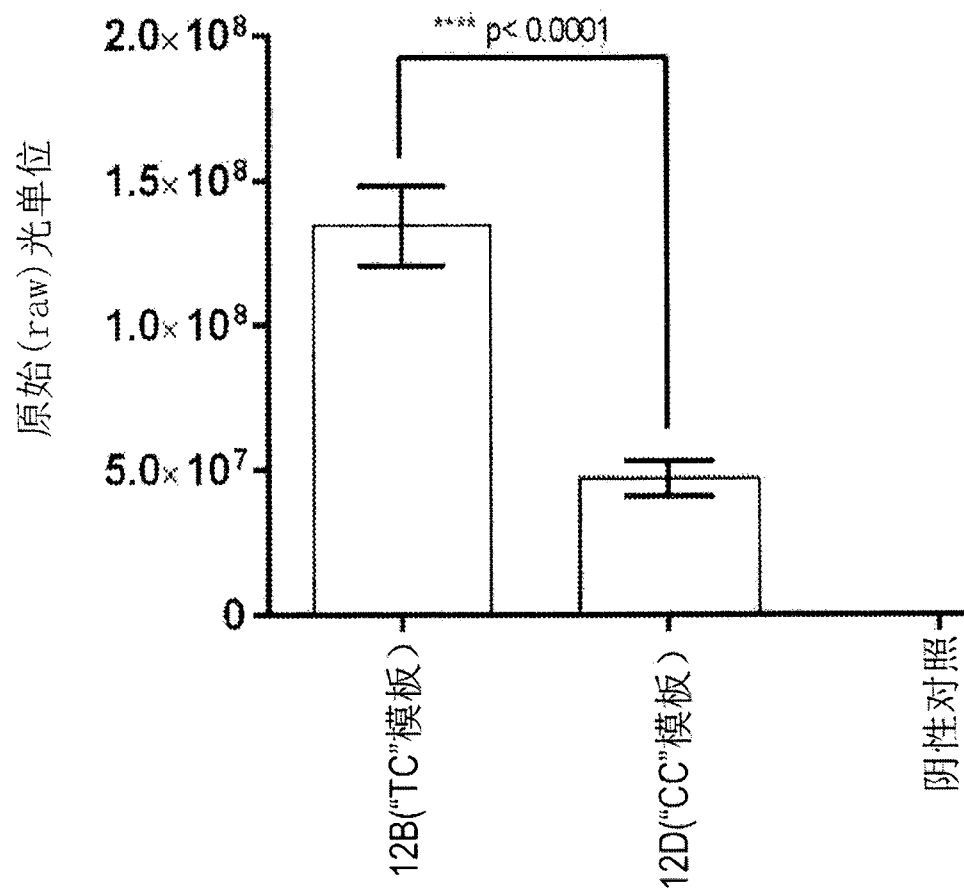


图14B

Abstract

COMPOSITIONS AND METHODS OF SYNTHESIZING 5'-CAPPED RNAS

Provided herein are methods and compositions for synthesizing 5'-Capped RNAs wherein the initiating capped oligonucleotide primers have the general form ${}^m\text{Gppp}[\text{N}_{2'-\text{Ome}}]_n[\text{N}]_m$ wherein ${}^m\text{G}$ is N7-methylated guanosine or any guanosine analog, N is any natural, modified or unnatural nucleoside, "n" can be any integer from 0 to 4 and "m" can be an integer from 1 to 9.

摘要

用于合成 5'-加帽 RNA 的组合物和方法

本文提供用于合成5'加帽RNA的方法和组合物，其中起始加帽寡核苷酸引物具有通式形式 $m^7Gppp[N_{2'-Ome}]_n[N]_m$ ，其中 m^7G 为N7-甲基化的鸟苷或任何鸟苷类似物，N为任何天然的、修饰的或非天然的核苷，“n”可以是0至4的任何整数且“m”可以是1至9的整数。