



(51) International Patent Classification:

B01J 19/00 (2006.01) *C12Q 1/6811* (2018.01)
C12P 19/34 (2006.01) *C12Q 1/6837* (2018.01)
C12Q 1/6806 (2018.01) *C12N 9/12* (2006.01)

(21) International Application Number:

PCT/US2020/015640

(22) International Filing Date:

29 January 2020 (29.01.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

16/261,229 29 January 2019 (29.01.2019) US

(71) Applicant: **MOLECULAR ASSEMBLIES, INC.**
[US/US]; 10355 Science Center Drive, Suite 130, San Diego, CA 92121 (US).

(72) Inventors: **EFCAVITCH, J., William**; 10355 Science Center Drive, Suite 130, San Diego, CA 92121 (US).
HOLDEN, Matthew, T.; 10355 Science Center Drive, Suite 130, San Diego, CA 92121 (US).

(74) Agent: **MEYERS, Thomas, C.** et al.; Brown Rudnick LLP, One Financial Center, Boston, MA 02111 (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

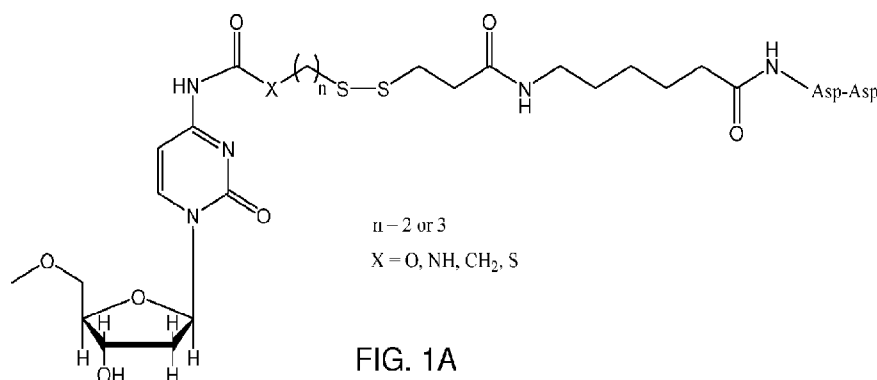
(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*):

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).

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: REUSABLE INITIATORS FOR SYNTHESIZING NUCLEIC ACIDS



(57) Abstract: The invention provides improved methods for synthesizing polynucleotides, such as DNA and RNA, using renewable initiators coupled to a solid support. Using the methods of the invention, specific sequences of polynucleotides can be synthesized de novo, base by base, in an aqueous environment, without the use of a nucleic acid template.

REUSABLE INITIATORS FOR SYNTHESIZING NUCLEIC ACIDS

RELATED APPLICATIONS

This application claims the benefit of and priority to U.S. Non-Provisional Patent Application Ser. No. 16/261,229, filed January 29, 2019, the content of which is incorporated by reference herein in its entirety.

FIELD OF THE INVENTION

The invention relates to methods and apparatus for synthesizing polynucleotides with a desired sequence and without the need for a template.

BACKGROUND

Most *de novo* nucleic acid sequencing is performed using well-established solid-phase phosphoramidite-techniques. The phosphoramidite technique involves the sequential de-protection and synthesis of sequences built from phosphoramidite reagents corresponding to natural (or non-natural) nucleic acid bases. Phosphoramidite nucleic acid synthesis is length-limited, however, in that nucleic acids greater than 200 base pairs (bp) in length experience high rates of breakage and side reactions. Additionally, phosphoramidite synthesis produces toxic by-products, and the disposal of this waste limits the availability of nucleic acid synthesizers, and increases the costs of contract oligonucleotide production. (It is estimated that the annual demand for oligonucleotide synthesis is responsible for greater than 300,000 gallons of hazardous chemical waste, including acetonitrile, trichloroacetic acid, toluene, tetrahydrofuran, and pyridine. See LeProust et al., Nucleic Acids Res., vol. 38(8), p.2522-2540, (2010), incorporated by reference herein in its entirety). Thus, there is a need for more efficient and cost-effective methods for oligonucleotide synthesis.

SUMMARY

The invention provides improved methods for nucleic acid synthesis. Methods of the invention provide faster and longer *de novo* synthesis of polynucleotides. As such, the invention dramatically reduces the overall cost of custom nucleic acid synthesis. Methods of the invention are directed to template-independent synthesis of polynucleotides by using a

nucleotidyl transferase enzyme to incorporate nucleotide analogs coupled to an inhibitor by a cleavable linker. Because of the inhibitor, synthesis pauses with the addition of each new base, whereupon the linker is cleaved, separating the inhibitor and leaving a polynucleotide that is essentially identical to a naturally occurring nucleotide (i.e., is recognized by the enzyme as a substrate for further nucleotide incorporation).

In particular, the invention provides a renewable substrate for template-independent nucleic acid synthesis. *De novo* synthesis begins with a nucleic acid initiator that is bound to a solid support. In the presence of a suitable enzyme, e.g., a polymerase, e.g., a terminal deoxynucleotidyl transferase (TdT), nucleotide analogs are added to the nucleic acid initiator in order to create an oligonucleotide. It is preferable that the nucleotide analogs include removable terminating groups that cause the enzymatic addition to stop after the addition of one nucleotide. A removable terminating group can be linked to the base portion of the nucleic acid and/or to the 3' hydroxyl of the nucleic acid. Deblocking of the terminating group and/or the 3' blocking group, creates a new active site that is a substrate for the enzyme. With subsequent addition of a new nucleotide or nucleotide analog, the oligonucleotide is extended.

In some instances, the nucleic acid initiator comprises a 3' moiety that is a substrate for the enzyme. A releasing agent is used to decouple the 3' moiety, thereby releasing the oligonucleotide. The 3' moiety, the nucleic acid initiator, and the solid support are reusable after the release of the nascent oligonucleotide.

The invention additionally includes an apparatus that utilizes methods of the invention for the production of custom polynucleotides. An apparatus of the invention includes one or more bioreactors providing aqueous conditions and a plurality of sources of nucleotide analogs. The bioreactor may be e.g., a reservoir, a flow cell, or a multi-well plate. The bioreactor may include a solid support having a nucleic acid initiator and a cleavable 3' moiety. Starting from the solid support, the polynucleotides are grown in the reactor by adding successive nucleotides via the natural activity of a nucleotidyl transferase, e.g., a terminal deoxynucleotidyl transferase (TdT) or any other enzyme that elongates DNA or RNA strands without template direction. Upon cleavage of the linker, a natural polynucleotide is released from the solid support. Once the sequence is complete, the support is cleaved away, or the 3' moiety is contacted with a releasing agent, leaving a polynucleotide essentially equivalent to that found in nature. In some embodiments, the apparatus is designed to recycle nucleotide analog solutions by recovering the solutions after nucleotide addition and reusing solutions for subsequent nucleotide

addition. Thus, less waste is produced, and the overall cost per base is reduced as compared to state-of-the-art methods. In certain embodiments, a bioreactor may include a microfluidic device and/or use inkjet printing technology.

Terminating groups may include, for example, charged moieties or steric inhibitors. In general, large macromolecules that prevent nucleotidyl transferase enzymes from achieving a functional conformation are useful used to inhibit oligonucleotide synthesis. Such macromolecules include polymers, polypeptides, polypeptoids, and nanoparticles. The macromolecules should be large enough to physically block access to the active site of the nucleotidyl transferase, not so large as to negatively alter the reaction kinetics. The macromolecules are linked to nucleotide analogs using any of a variety of linkers, as described below.

In embodiments using a 3'-O-blocked nucleotide analog, the 3'-O-blocking groups are typically small and easily removed, thus allowing use with engineered enzymes having modified active sites. For example, the 3'-O-blocking groups may comprise azidomethyl, amino, or allyl groups.

In some embodiments, oligonucleotide synthesis may include introduction of a 3' exonuclease to the one or more synthesized oligonucleotides after each nucleotide analog addition, but before cleaving the terminating group. The terminating group blocks the 3' exonuclease from acting on any oligonucleotide that to which a nucleotide analog has been added, while oligonucleotides that have not successfully added an analog containing a terminator are removed by the 3' exonuclease. In this manner, the invention allows for in-process quality control and may eliminate the need for post-synthesis purification.

Other aspects of the invention are apparent to the skilled artisan upon consideration of the following figures and detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A shows a genus of deoxycytidine triphosphate (dCTP) analogs having a cleavable terminator linked at the N-4 position;

FIG. 1B shows cleavage of the cleavable terminator from a dCTP analog of FIG. 1A to achieve a "natural" dCTP and a cyclic leaving molecule;

FIG. 2A shows a genus of deoxyadenosine triphosphate (dATP) analogs having a cleavable terminator linked at the N-6 position;

FIG. 2B shows cleavage of the cleavable terminator from a dATP analog of FIG. 2A to

achieve a “natural” dATP and a cyclic leaving molecule;

FIG. 3A shows a genus of deoxyguanosine triphosphate (dGTP) analogs having a cleavable terminator linked at the N-2 position;

FIG. 3B shows cleavage of the cleavable terminator from a dGTP analog of FIG. 3A to achieve a “natural” dGTP and a cyclic leaving molecule;

FIG. 4A shows a genus of deoxythymidine triphosphate (dTTP) analogs having a cleavable terminator linked at the N-3 position;

FIG. 4B shows cleavage of the cleavable terminator from a dTTP analog of FIG. 4A to achieve a “natural” dTTP and a cyclic leaving molecule;

FIG. 5A shows a genus of deoxyuridine triphosphate (dUTP) analogs having a cleavable terminator linked at the N-3 position;

FIG. 5B shows cleavage of the cleavable terminator from a dUTP analog of FIG. 5A to achieve a dUTP and a cyclic leaving molecule;

FIG. 6 shows an exemplary deoxycytidine triphosphate (dCTP) analog having a Staudinger linker connecting a blocking Asp-Asp molecule to the N-4 position of the deoxycytidine and subsequent cleavage of the Staudinger linker under aqueous conditions to achieve a dCTP and a leaving group;

FIG. 7A shows a genus of cytidine triphosphate (rCTP) analogs having a cleavable terminator linked at the N-4 position;

FIG. 7B shows cleavage of the cleavable terminator from a rCTP analog of FIG. 7A to achieve a “natural” rCTP and a cyclic leaving molecule;

FIG. 8A shows a genus of adenosine triphosphate (rATP) analogs having a cleavable terminator linked at the N-6 position;

FIG. 8B shows cleavage of the cleavable terminator from an rATP analog of FIG. 8A to achieve a “natural” rATP and a cyclic leaving molecule;

FIG. 9A shows a genus of guanosine triphosphate (rGTP) analogs having a cleavable terminator linked at the N-2 position;

FIG. 9B shows cleavage of the cleavable terminator from a rGTP analog of FIG. 9A to achieve a “natural” rGTP and a cyclic leaving molecule;

FIG. 10A shows a genus of thymidine triphosphate (rTTP) analogs having a cleavable terminator linked at the N-3 position;

FIG. 10B shows cleavage of the cleavable terminator from a rTTP analog of FIG. 10A to achieve a “natural” rTTP and a cyclic leaving molecule;

FIG. 11A shows a genus of uridine triphosphate (rUTP) analogs having a cleavable terminator linked at the N-3 position;

FIG. 11B shows cleavage of the cleavable terminator from a rUTP analog of FIG. 11A to achieve a rUTP and a cyclic leaving molecule;

FIG. 12 shows an exemplary cytidine triphosphate (rCTP) analog having a Staudinger linker connecting a blocking Asp-Asp molecule to the N-4 position of the cytidine and subsequent cleavage of the Staudinger linker under aqueous conditions to achieve a rCTP and a leaving group;

FIG. 13 shows an exemplary terminal deoxynucleotidyl transferase (TdT) mediated polynucleotide synthetic cycle, including: (a) incorporation of a nucleotide triphosphate analog comprising cleavable terminator, dN*TP-OH, and (b) removal of the terminating blocking group (indicated by *), thus enabling the next dN*TP-OH to be incorporated, wherein N = A, G, C, or T;

FIG. 14 shows an exemplary nucleotide analog with a cleavable linker comprising a variable number of methylene bridges;

FIG. 15 shows an exemplary nucleotide analog with a cleavable linker comprising a cysteine residue;

FIG. 16A shows an exemplary nucleotide analog with an anionic inhibitor comprising a single phosphate group;

FIG. 16B shows an exemplary nucleotide analog with an anionic inhibitor comprising two phosphate groups;

FIG. 16C shows an exemplary nucleotide analog with an anionic inhibitor comprising three phosphate groups;

FIG. 17 shows an exemplary microfluidic polynucleotide synthesis device;

FIG. 18 shows an exemplary polypeptoid inhibitor suitable for use in the invention;

FIG. 19 shows a flow-chart describing the use of a 3' exonuclease to digest oligonucleotides that are not properly terminated between oligonucleotide synthesis cycles;

FIG. 20 illustrates synthesis of a *de novo* oligonucleotide using nucleotide triphosphate analogs having a 3'-O-blocking group;

FIG. 21 shows four exemplary 3'-O-blocked nucleotide analogs that can be used for the synthesis of *de novo* oligonucleotides in conjunction with a suitable template-independent polymerase; and

FIG. 22 illustrates the incorporation of a reusable 3' moiety into a nucleic acid coupled

to a solid support, growth of a nucleic acid using a modified TdT, and release of the *de novo* oligonucleotide.

FIG. 23 shows the results of enzymatic installation of a 3'-poly-U tract with TdT, followed by USER digestion of the tract.

FIG. 24 shows the results of USER digestion of an internal poly-U tract to produce a 5'-monodisperse cleavage product.

FIG. 25 illustrates solid-phase dephosphorylation and TdT extension results.

FIG. 26 shows an exemplary resin regeneration cycle.

FIG. 27 shows an exemplary index strand regeneration process.

DETAILED DESCRIPTION

The invention provides improved methods for synthesizing polynucleotides, such as DNA and RNA, using enzymes and nucleic acid analogs. Using the disclosed methods, specific sequences of polynucleotides can be synthesized *de novo*, base by base, in an aqueous environment, and without the use of a nucleic acid template.

Nucleotide analogs may have an unmodified 3' hydroxyl, or may have a 3'-O-blocking group or may have a blocker releasably attached to a phosphate. In any case, the blocking group is designed to not leave behind substantial additional molecules, i.e., designed to leave behind "scarless" nucleotides that are recognized as "natural" nucleotides by the enzyme. Thus, at the conclusion of the synthesis, upon removal of the last blocking group, the synthesized polynucleotide is chemically and structurally equivalent to the naturally-occurring polynucleotide with the same sequence. The synthetic polynucleotide can, thus, be incorporated into living systems without concern that the synthesized polynucleotide will interfere with biochemical pathways or metabolism.

The process and analogs of the invention are used for the non-template mediated enzymatic synthesis of oligo- and oligodeoxynucleotides especially of long oligonucleotides (<5000 nt). Products can be single stranded or partially double stranded depending upon the initiator used. The synthesis of long oligonucleotides requires high efficiency incorporation and high efficiency of reversible terminator removal. The initiator bound to the solid support consists of a short, single strand DNA sequence that is either a short piece of the user defined sequence or a universal initiator from which the user defined single strand product is removed.

In one aspect, the disclosed methods employ commercially-available nucleotidyl transferase enzymes, such as terminal deoxynucleotidyl transferase (TdT), to synthesize

polynucleotides from nucleotide analogs in a step-by-step fashion. The nucleotide analogs are of the form:



wherein NTP is a nucleotide triphosphate (i.e., a dNTP or an rNTP), the linker is a cleavable linker between the pyridine or pyrimidine of the base, and the inhibitor is a group that prevents the enzyme from incorporating subsequent nucleotides. At each step, a new nucleotide analog is incorporated into the growing polynucleotide chain, whereupon the enzyme is blocked from adding an additional nucleotide by the inhibitor group. Once the enzyme has stopped, the excess nucleotide analogs are removed from the growing chain, the inhibitor can be cleaved from the NTP, and new nucleotide analogs can be introduced in order to add the next nucleotide to the chain. By repeating the steps sequentially, it is possible to quickly construct nucleotide sequences of a desired length and sequence. Advantages of using nucleotidyl transferases for polynucleotide synthesis include: 1) 3'- extension activity using single strand (ss) initiating primers in a template-independent polymerization, 2) the ability to extend primers in a highly efficient manner resulting in the addition of thousands of nucleotides, and 3) the acceptance of a wide variety of modified and substituted NTPs as efficient substrates.

In addition, the invention can make use of an initiator sequence that is a substrate for nucleotidyl transferase. The initiator is attached to a solid support and serves as a recognition site for the enzyme. The initiator is preferably a universal initiator for the enzyme, such as a homopolymer sequence, and is recyclable on the solid support, the formed oligonucleotide being cleavable from the initiator.

Methods of the invention are well-suited to a variety of applications that currently use synthetic nucleic acids, e.g., phosphoramidite-synthesized DNA oligonucleotides. For example, polynucleotides synthesized using the methods of the invention are used as primers for nucleic acid amplification, hybridization probes for detection of specific markers, and for incorporation into plasmids for genetic engineering. However, because the disclosed methods produce longer synthetic strings of nucleotides, at a faster rate, and in an aqueous environment, the disclosed methods also lend themselves to high-throughput applications, such as screening for expression of genetic variation in cellular assays, as well as synthetic biology. Furthermore, the methods of the invention will provide the functionality needed for next-generation applications, such as using DNA as synthetic read/write memory, or creating macroscopic materials synthesized completely (or partially) from DNA.

The invention and systems described herein provide for synthesis of polynucleotides,

including deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). While synthetic pathways for “natural” nucleotides, such as DNA and RNA, are described in the context of the common nucleic acid bases, e.g., adenine (A), guanine (G), cytosine (C), thymine (T), and uracil(U), it is to be understood that the methods of the invention can be applied to so-called “non-natural” nucleotides, including nucleotides incorporating universal bases such as 3-nitropyrrole 2'-deoxynucleoside and 5-nitroindole 2'-deoxynucleoside, alpha phosphorothiolate, phosphorothioate nucleotide triphosphates, or purine or pyrimidine conjugates that have other desirable properties, such as fluorescence. Other examples of purine and pyrimidine bases include pyrazolo[3,4-d]pyrimidines, 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo (e.g., 8-bromo), 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 8-azaguanine and 8-azaadenine, deazaguanine, 7-deazaguanine, 3-deazaguanine, deazaadenine, 7-deazaadenine, 3-deazaadenine, pyrazolo[3,4-d]pyrimidine, imidazo[1,5-a]1,3,5 triazinones, 9-deazapurines, imidazo[4,5-d]pyrazines, thiazolo[4,5-d]pyrimidines, pyrazin-2-ones, 1,2,4-triazine, pyridazine; and 1,3,5 triazine. In some instances, it may be useful to produce nucleotide sequences having unreactive, but approximately equivalent bases, i.e., bases that do not react with other proteins, i.e., transcriptases, thus allowing the influence of sequence information to be decoupled from the structural effects of the bases.

ANALOGS

The invention provides nucleotide analogs having the formula NTP–linker–inhibitor for synthesis of polynucleotides in an aqueous environment. With respect to the analogs of the form NTP–linker–inhibitor, NTP can be any nucleotide triphosphate, such as adenosine triphosphate (ATP), guanosine triphosphate (GTP), cytidine triphosphate (CTP), thymidine triphosphate (TTP), uridine triphosphate (UTP), nucleotide triphosphates, deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), deoxycytidine triphosphate (dCTP), deoxythymidine triphosphate (dTTP), or deoxyuridine triphosphate (dUTP).

The linker can be any molecular moiety that links the inhibitor to the NTP and can be

cleaved. For example, the linkers can be cleaved by adjusting the pH of the surrounding environment. The linkers may also be cleaved by an enzyme that is activated at a given temperature, but inactivated at another temperature. In some embodiments, the linkers include disulfide bonds.

Linkers may, for example, include photocleavable, nucleophilic, or electrophilic cleavage sites. Photocleavable linkers, wherein cleavage is activated by a particular wavelength of light, may include benzoin, nitroveratryl, phenacyl, pivaloyl, silyl, 2-hydroxy-cinamyl, coumarin-4-yl-methyl, or 2-nitrobenzyl based linkers.

Examples of nucleophilic cleavage sites include fluoride ion cleavable silicon-oxygen bonds or esters which may be cleaved in a basic solution. Electrophilically-cleaved linkers may include acid induced cleavage sites which may comprise trityl, tert-butyloxycarbonyl groups, acetal groups, and p-alkoxybenzyl esters and amides. In certain aspects, a cleavable linker may include a cysteine residue as shown in FIG. 15.

The linker can be attached, for example, at the N4 of cytosine, the N3 or O4 of thymine, the N2 or N3 of guanine, and the N6 of adenine, or the N3 or O4 of uracil because attachment at a carbon results in the presence of a residual scar after removal of the polymerase-inhibiting group. The linker is typically on the order of at least about 10 Angstroms long, e.g., at least about 20 Angstroms long, e.g., at least about 25 Angstroms long, thus allowing the inhibitor to be far enough from the pyridine or pyrimidine to allow the enzyme to bind the NTP to the polynucleotide chain via the attached sugar backbone. In some embodiments, the cleavable linkers are self-cyclizing in that they form a ring molecule that is particularly non-reactive toward the growing nucleotide chain.

In certain aspects, a cleavable linker may include a variable number of methylene bridges on the NTP or the inhibitor side of a disulfide bond, including, for example, 1, 2, 3, or 4 methylene bridges as shown in FIGS. 14 and 16A-C. These methylene bridges may be used to increase the space between the NTP and the inhibitor. As noted above, the length of the cleavable linker may be selected in order to prevent the inhibitor from interfering with coupling of the NTP to the synthesized polynucleotide. In some embodiments of the invention, the distance of the charged group to the NTP plays an important role in the effectiveness of inhibiting a subsequent nucleotide incorporation.

For example, in some embodiments using a charged moiety as an inhibitor, the charged moiety may be from about 5 to about 60 bonds away from the NTP. In some other embodiments, the charged moiety of the inhibitor may be from about 10 to about 40 bonds

away from the NTP. In some other embodiments, the charged moiety of the inhibitor can be from about 10 to about 35 bonds away from the NTP. In some other embodiments, the charged moiety of the inhibitor may be from about 10 to about 30 bonds away from the NTP. In some other embodiments, the charged moiety of the inhibitor is from about 10 to about 20 bonds away from the NTP. The number of bonds between the charged moiety and the NTP may be increased by including additional methylene bridges.

The nucleotide analogs can include any moiety linked to the NTP that inhibits the coupling of subsequent nucleotides by the enzyme. The inhibitory group can be a charged group, such as a charged amino acid, or the inhibitory group can be a group that becomes charged depending upon the ambient conditions. In some embodiments, the inhibitor may include a moiety that is negatively charged or capable of becoming a negatively charged. For example, an inhibitor may include a chain of phosphate groups (e.g., 1, 2, or 3, phosphates) as shown in FIGS. 16A–C, wherein additional phosphates increase the overall anionic charge of the inhibitor. In other embodiments, the inhibitor group is positively charged or capable of becoming positively charged. In some other embodiments, the inhibitor is an amino acid or an amino acid analog. The inhibitor may be a peptide of 2 to 20 units of amino acids or analogs, a peptide of 2 to 10 units of amino acids or analogs, a peptide of 3 to 7 units of amino acids or analogs, a peptide of 3 to 5 units of amino acids or analogs. In some embodiments, the inhibitor includes a group selected from the group consisting of Glu, Asp, Arg, His, and Lys, and a combination thereof (e.g., Arg, Arg-Arg, Asp, Asp-Asp, Asp, Glu, Glu-Glu, Asp-Glu-Asp, Asp-Asp-Glu or AspAspAspAsp, etc.). Peptides or groups may be combinations of the same or different amino acids or analogs. In certain embodiments, a peptide inhibitor may be acetylated to discourage errant bonding of free amino groups. The inhibitory group may also include a group that reacts with residues in the active site of the enzyme thus interfering with the coupling of subsequent nucleotides by the enzyme. The inhibitor may have a charged group selected from the group consisting of $-\text{COO}$, $-\text{NO}_2$, $-\text{PO}_4$, $-\text{PO}_3$, $-\text{SO}_2$, or $-\text{NR}_3$ where each R may be H or an alkyl group. In other embodiments, the inhibitor moiety does not comprise a $-\text{PO}_4$ group.

In certain aspects, a terminator or inhibitor may include a steric inhibitor group. Such a steric inhibitor group may allow for the NTP-linker-inhibitor (i.e., nucleotide analog) to be incorporated onto the unblocked 3' OH of an oligonucleotide, said incorporation being catalyzed by nucleotidyl transferase. The steric inhibitor group may physically block the incorporation of nucleotides or additional nucleotide analogs onto the unblocked 3' OH of the

incorporated nucleotide analog. Steric inhibitors may also block 3' endonucleases from acting on a nucleotide analog and, accordingly, on oligonucleotides to which an un-cleaved nucleotide analog has been incorporated.

Steric inhibitors can include, for example, chemical polymers, nanoparticles, poly-N-substituted glycines (peptoids), or proteins. A steric inhibitor of the invention may be a variety of sizes including, e.g., greater than 20 Å, greater than 30 Å, greater than 40 Å, greater than 50 Å, greater than 60 Å, greater than 70 Å, greater than 80 Å, greater than 90 Å, greater than 100 Å, greater than 110 Å, greater than 120 Å, greater than 130 Å, greater than 140 Å, or greater than 150 Å. In preferred embodiments, a steric inhibitor may be monodisperse or substantially monodisperse. Steric inhibitors may be water soluble and conformationally-constrained (i.e., of a rigid or semi-rigid form). In certain aspects, a steric inhibitor will physically block access to the active site of the relevant nucleotidyl transferase enzyme because of the size or the conformation of the inhibitor. In preferred embodiments, the steric inhibitor may comprise a non-natural bio-inspired polymer such as a polypeptoid or a non-natural polypeptide.

In certain aspects, a self-assembling polypeptoid sequence may be used as a steric inhibitor. Peptoid monomers are often based on an N-substituted glycine backbone. Because the backbone is devoid of hydrogen bond donors, polypeptoids are readily processed while still being able to form secondary structures such as helices. They also provide the beneficial properties of allowing polarities and side chains similar to peptides, while being generally chemically and thermally stable. Self-assembling polypeptoid steric inhibitors according to the invention may self-assemble single peptoid helices to form microspheres in the micrometer range of diameters including, .3 μm, .4 μm, .5 μm, .6 μm, .7 μm, .8 μm, .9 μm, 1.0 μm, 1.5 μm, 2 μm, 2.5 μm, 3 μm, or 3.5 μm, among others. In certain aspects, steric inhibitors may include peptoids with C-α-branched side chains, N-Aryl side chains, N-1-naphthylethyl side chains, or other formations capable of forming stable helical structures. An example of a peptoid steric inhibitor is shown in FIG. 18. FIG. 18 illustrates a branched poly-N-methoxyethyl glycine which may be used as a steric inhibitor according to the invention. In certain embodiments, a steric inhibitor may include a reactive group that is easily joined to a linker group, e.g., a cleavable linking group as described herein.

In other embodiments, a steric inhibitor may comprise a polymer, such as a biocompatible polymer. The polymer may comprise blocks of different polymers, such that the blocks form a desired macroscopic structure, e.g., a sphere when exposed to an aqueous environment. For example, the copolymer may comprise blocks of hydrophilic and

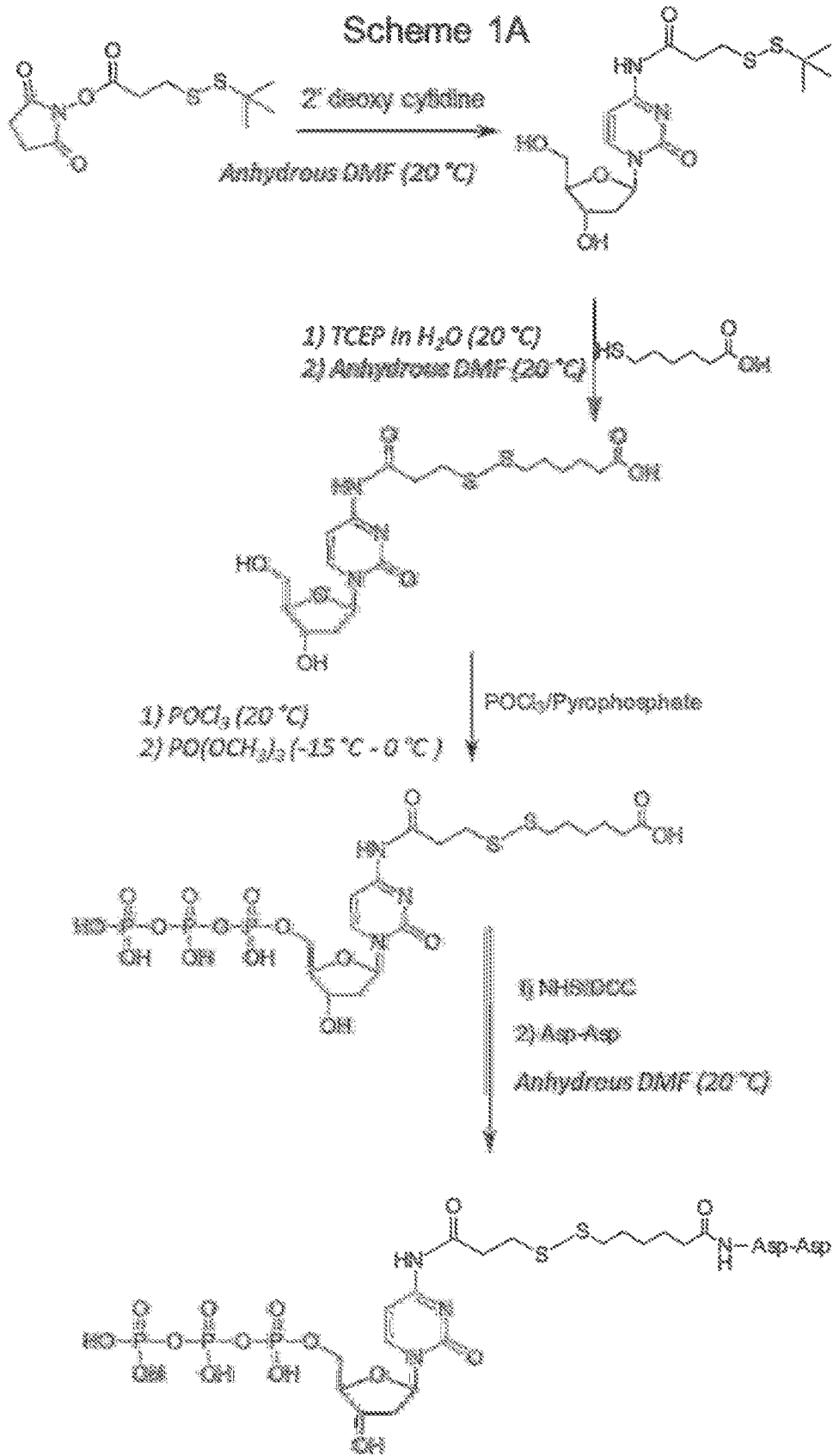
hydrophobic blocks so that the polymer self-assembles into a micellar structure upon addition to water. In some embodiments, the hydrophobic blocks may be selected from polycaprolactones (PCL), polydimethylsiloxanes (PDMS), polymethylmethacrylate (PMMA), or polylactides (PLA). The hydrophilic blocks may include polyethylene glycol (PEG) or other polyalcohol.

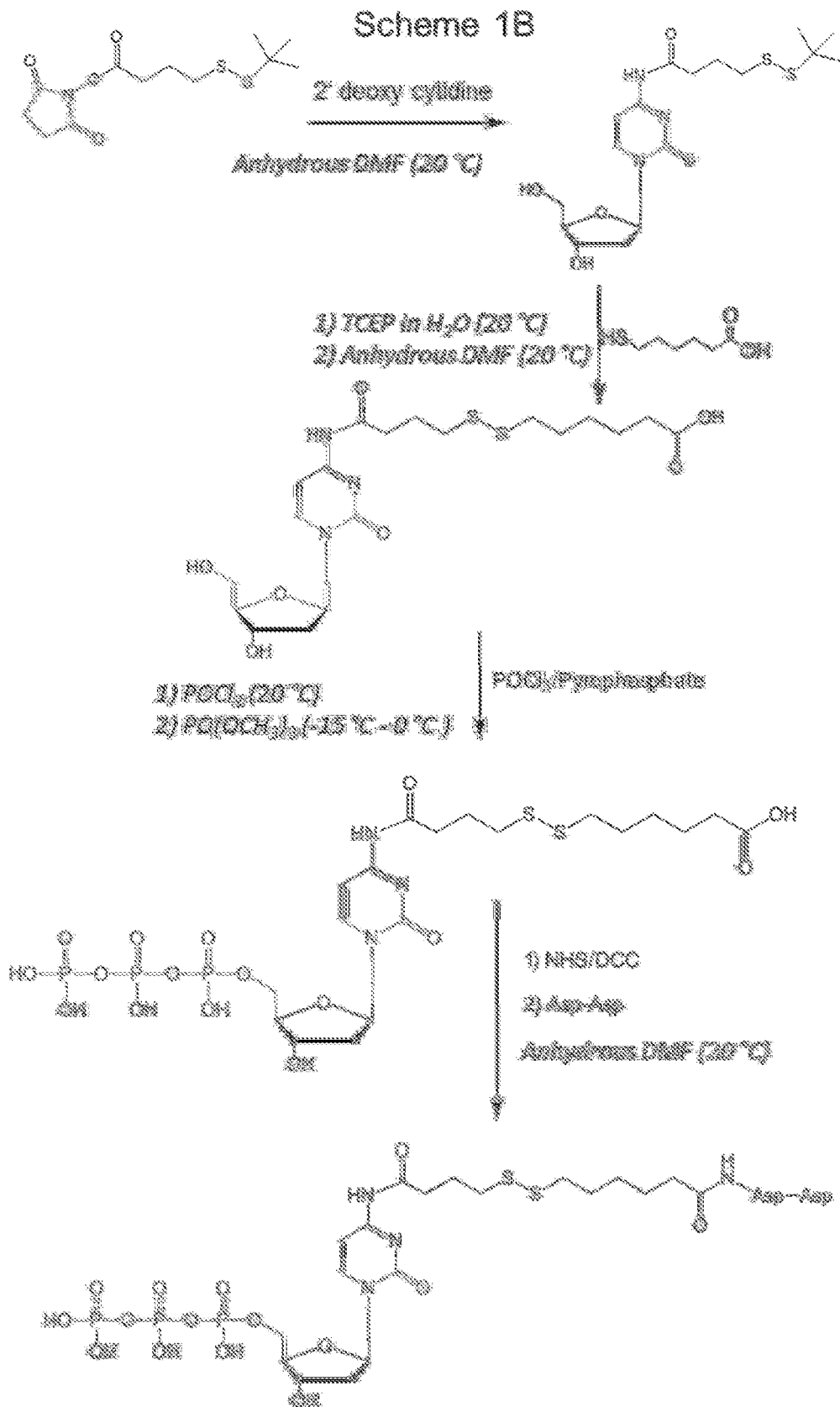
In other embodiments, the inhibitor may comprise a nanoparticle of sufficient size to block the activity of a nucleotidyl transferase. Such nanoparticles may comprise, e.g., gold, silver, silicon, cerium oxide, iron oxide, titanium dioxide, silicon nitride, silicon boride, or silica, e.g., mesoporous silica. In other embodiments, the nanoparticles may comprise highly-ordered molecular structures, such as fullerenes, e.g., buckyballs and nanotubes, comprising carbon, or semiconductors.

Steric inhibitors may have no charge or may be positively or negatively charged to provide compatibility with the nucleotide to which it is linked and with the nucleotidyl transferase enzyme so that the inhibitor does not interfere with the incorporation reaction on the 5' end of the NTP analog. Steric inhibitors may incorporate a variety of amino acid residues in order to provide a desired conformation, charge, or attachment site.

An example of a nucleotide analog of the type NTP-linker-inhibitor is shown in FIG. 1A. The analog in FIG. 1A includes an inhibitory (-Asp-Asp-) group linked to the N⁴ position of dCTP through a disulfide (-S-S-) bond while providing an unblocked, unmodified 3'-OH on the sugar ring. The linker is constructed such that all linker atoms (including the 2nd incorporation-inhibiting moiety) can be removed, thereby allowing the nascent DNA strand to revert to natural nucleotides. As shown in FIG. 1B, an aqueous reducing agent, such as tris(2-carboxyethyl) phosphine (TCEP) or dithiothreitol (DTT), can be used to cleave the -S-S- bond, resulting in the loss of the inhibitor function (deblocking). As shown in FIG. 1B, a self-cyclizing linker can be incorporated, resulting in a cyclic oxidized tetrahydrothiophene leaving group that is easily removed from the reagent solution at the conclusion of nucleotide synthesis.

Exemplary schemes for synthesizing dCTP analogs of FIG. 1A are shown below in Schemes 1A and 1B:

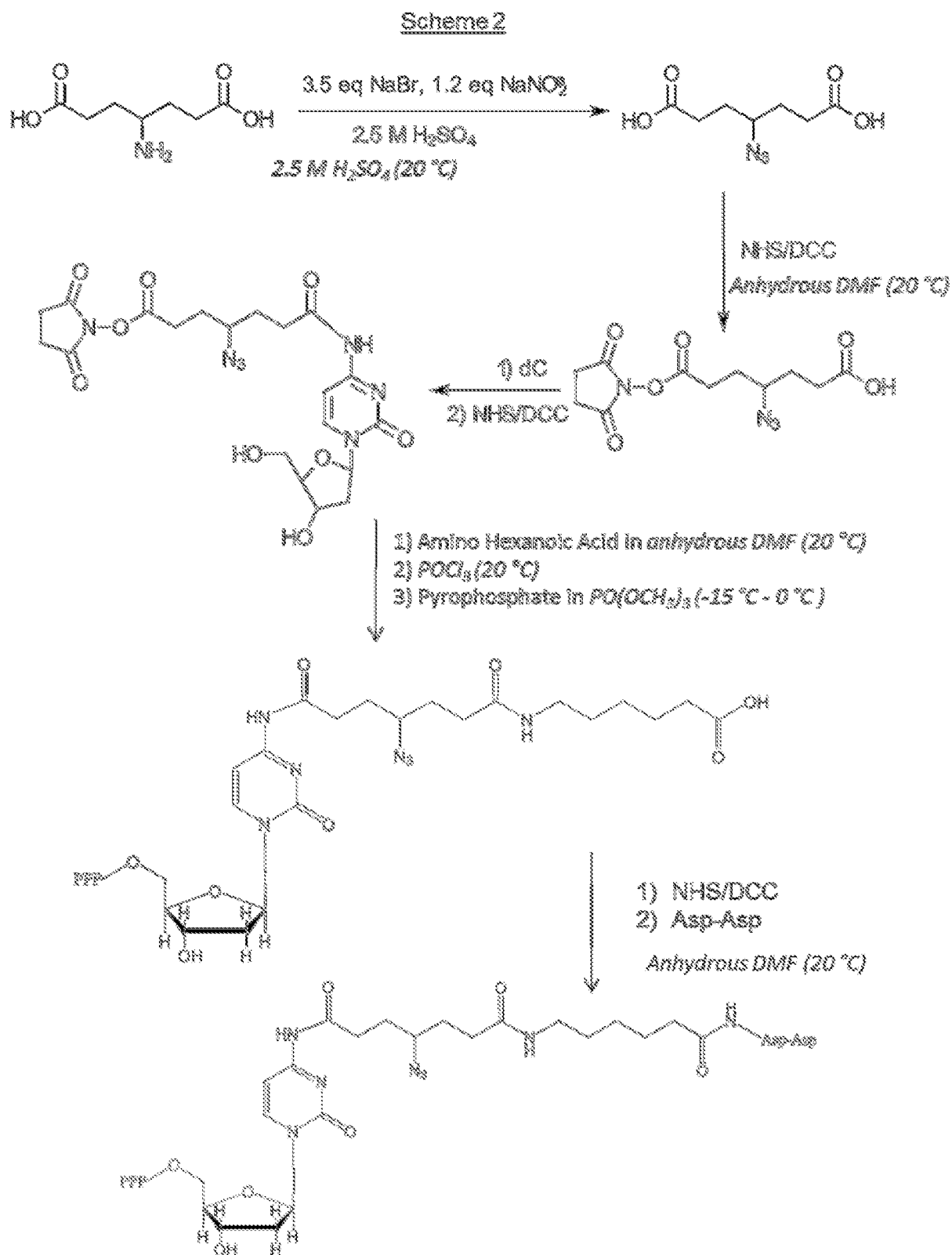




In a fashion analogous to Schemes 1A and 1B, nucleotide analogs of the type NTP–linker–inhibitor can also be formed by attaching the linker–inhibitor moiety to the N6 of adenine (FIG. 2), the N2 of guanine (FIG. 3), the N3 of thymine (FIG. 4), or the N3 of uracil (FIG. 5), thereby providing analogs of the “naturally-occurring” dNTPs, as well as a deoxyuracil nucleotide (dUTP). While it is unlikely that there will be wide use of a dUTP, the synthesis is straightforward based upon the chemistry.

The invention is not limited to the linking chemistry of Schemes 1A and 1B, however, as carbamate, amide, or other self-eliminating linkages could also be employed. For example, nucleotides can also be prepared with Staudinger linkers, as shown in Scheme 2:

Scheme 2



A deoxycytidine triphosphate (dCTP) analog created with a Staudinger linker (Scheme 2) to an Asp-Asp blocking group is shown in FIG. 6. As shown in FIG. 6, the Staudinger dCTP analog undergoes cleavage under aqueous conditions with the addition of azide and triphenylphosphine. The Staudinger analog shown in FIG. 6 is also suitable for nucleotide extension using nucleotidyl transferases, such as TdT, as described above and exemplified in FIGS. 1-5. While not shown explicitly in the FIGS., one of skill in the art can use Scheme 2 in conjunction with the suitable reactant to produce other nucleotide analogs having Staudinger linkers as needed for complete *de novo* nucleotide synthesis. In a fashion analogous to FIG. 6, nucleotide analogs of Scheme 2 can be formed by attaching the Staudinger moiety to the N6 of adenine, the N2 of guanine, the N3 of thymine, or the N3 of uracil, thereby providing analogs of the “naturally-occurring” dNTPs, as well as a deoxyuracil nucleotide (dUTP).

The methodologies of Scheme 1A can be used to produce corresponding ribonucleotide analogs, e.g., as shown in FIGS. 7-10, by starting with the appropriate ribonucleotide reactants. Ribonucleotide analogs comprising the Staudinger linker can also be created using Scheme 2 in order to form the needed ribonucleotide analogs, including, e.g., CTP analogs, as shown in FIG. 12. Furthermore, all of the ribonucleotide analogs, i.e., C, A, T, G, U, can be formed using a reaction similar to Scheme 2.

In other embodiments, a 3'-O-blocked nucleotide analog can be used along with a modified enzyme capable of incorporating 3'-O-blocked nucleotide analogs into an oligonucleotide. Such modified enzymes will allow 3'-O-blocked dNTP analogs to be used in a step-by-step method to extend an initiating nucleic acid into a user defined sequence (see FIG. 20). Furthermore, after each nucleotide extension step, the reactants can be recovered and recycled from the solid support back to the original reagent reservoir. Once that step is complete, the 3'-O-blocking group will be removed, allowing the cycle to start anew. At the conclusion of n cycles of extension-recover-deblock-wash, the full length, single strand polydeoxynucleotide will be cleaved from the solid support and isolated for subsequent use. A variety of 3'-O-blocked deoxynucleotides, may be used, but the choice of specific 3'-O-blocking groups is dictated by: 1) the smallest possible bulk to maximize substrate utilization by TdT and 2) removal of the blocking group with the mildest and preferably aqueous conditions in the shortest period of time.

A variety of 3'-O-modified dNTPs and NTPs may be used with the disclosed proteins for *de novo* synthesis. In some embodiments, the preferred removable 3'-O-blocking group is a 3'-O-amino, a 3'-O-allyl or a 3'-O-azidomethyl. In other embodiments, the removable 3'-O-blocking

moiety is selected from the group consisting of O-phenoxyacetyl; O-methoxyacetyl; O-acetyl; O-(p-toluene)-sulfonate; O-phosphate; O-nitrate; O-[4-methoxy]-tetrahydrothiopyranlyl; O-tetrahydrothiopyranlyl; O-[5-methyl]-tetrahydrofuranlyl; O-[2-methyl,4-methoxy]-tetrahydropyranlyl; O-[5-methyl]-tetrahydropyranlyl; and O-tetrahydrothiofuranlyl (see US 8,133,669). In other embodiments the removable blocking moiety is selected from the group consisting of esters, ethers, carbonitriles, phosphates, carbonates, carbamates, hydroxylamine, borates, nitrates, sugars, phosphoramidate, phosphoramidates, phenylsulfenates, sulfates, sulfones and amino acids (see Metzker ML et al. Nuc Acids Res. 1994;22(20):4259-67, U.S.P.N. 6,232,465; U.S.P.N. 7,414,116; and U.S.P.N. 7,279,563, all of which are incorporated by reference in their entireties).

FIG. 21 shows four exemplary 3'-O-blocked dNTP analogs, namely 3'-O-azidomethyl-dATP, 3'-O-azidomethyl-dCTP, 3'-O-azidomethyl-dGTP, and 3'-O-azidomethyl-dTTP. The 3'-O-blocked dNTP analogs can be purchased from specialty suppliers, such as Azco Biotech, Oceanside, CA. Corresponding 3'-O-blocked ribonucleotides can also be obtained commercially, thus enabling the creation of custom RNA oligonucleotides.

ENZYMES

The methods of the invention employ nucleotidyl transferases to assemble the nucleotide analogs into polynucleotides. Nucleotidyl transferases include several families of related transferase and polymerase enzymes. Some nucleotidyl transferases polymerize deoxyribonucleotides more efficiently than ribonucleotides, some nucleotidyl transferases polymerize ribonucleotides more efficiently than deoxyribonucleotides, and some nucleotidyl transferases polymerize ribonucleotides and deoxyribonucleotides at approximately the same rate.

Of particular import to the invention, transferases having polymerase activity, such as terminal deoxynucleotidyl transferase (TdT), are capable of catalyzing the addition of deoxyribonucleotides to the 3' end of a nucleotide chain, thereby increasing chain length in DNA nucleotides. TdT will only catalyze the addition of 1-2 ribonucleotides to the growing end of a DNA strand which could be useful in the construction of site specific DNA-RNA chimeric polynucleotides. In particular, calf thymus TdT, sourced from engineered *E. coli*, is suitable for use with the invention and available from commercial sources such as Thermo Scientific (Pittsburgh, PA). The amino acid sequence corresponding to calf TdT is listed in Table 1 as SEQ ID NO. 1.

Table 1. Amino Acid Sequence of Bovine TdT

SEQ ID NO. 1: MAQQRQHQL PMDPLCTASS GPRKKRPRQV GASMASPPHD
 IKFQNLVLFILKLEKMGTTTRNFLMELARRKGFRVENELSDSVTHIVAENN
 SGSEVLEWLQVQNIRASSQL ELLDVSWLIESMGAGKPVEITGKHQLVVRT
 DYSATPNPGFQKTPPLAVKKISQYACQRKTLNNYNHIFTDAFEILAENS EFKENEVSIV
 TFMRAASVLKSLPFTIISMK DTEGIPCLGD KVKCIIIEEII EDGESSEVKA VLNDERYQSF
 KLFTSVFGVGLKTSEKWFRMGFRSLKIMSDKTLKFTKMQ KAGFLYYEDL
 VSCVTRAEAEAVGVLVKEAVWAFLPDAFVTMTGGFRRGKIGHDVDFLIT
 SPGSAEDEEQLLPKVINLWEKKGLLLYYDLVESTFEKFKLPSRQVDTLDHFQKCFLLKL
 HHQRVDSSKSNQEGKTWKAIRVDLVMCPYENRAFALLGW TGSRQFERDI
 RRYATHERKMMLDNHALYDKTKRVFLKAES EEIFAHLGLDYIEPWERNA

The nucleotide sequence corresponding to calf TdT is listed in Table 2 as SEQ ID NO. 2.

Table 2. Nucleic Acid Sequence of Bovine TdT

SEQ ID NO. 2: ctctctgga gataccactt gatggcacag cagaggcagc atcagcgtct tccatggat ccgctgtgca
 cagcctcctc aggcctcctg aagaagagac ccaggcaggt gggcgcctca atggcctccc ctctcatga catcaagttt
 caaaatttgg tctcttcat tttggagaag aaaatgggaa ccaccgcag aaactctc atggagctgg ctggaaggaa agtttcagg
 gttgaaaatg agctcagtga ttctgtcacc cacattgtag cagaaaacaa ctctgggtca gaggttctcg agtggctca ggtacagaac
 ataagagcca gctcgcagct agaactcctt gatgtctcct ggctgatcga aagtatggga gcaggaaaac cagtggagat
 tacaggaaaa caccagcttg ttgtgagaac agactattca gctaccccaa acccaggctt ccagaagact ccaccacttg
 ctgtaaaaa gatctccag tacgcgtgac aaagaaaac cactttgaac aactataacc acatattcac ggatgccttt gagatactgg
 ctgaaaatc tgagttaaa gaaaatgaag tctcttatgt gacattatg agagcagct ctgtactaa atctctgcca tcacaatca
 tcagtatgaa ggatacagaa ggaattcctt gctggggga caaggtgaag tgtatcatag aggaaattat tgaagatgga
 gaaagtctg aagttaaagc tgtgttaaat gatgaacgat atcagtcctt caaactctt acttctgttt ttggagtggg actgaagaca
 tctgagaaat gggtcaggat ggggttcaga tctctgagta aaataatgac agacaaaacc ctgaaattca caaaaatgca
 gaaagcagga ttctctatt atgaagacct tgcagctgc gtgaccagg ccgaagcaga ggcgggtggc gtgctggta
 aagaggctgt gtgggcattt ctgccgatg cctttgtcac catgacagga ggattccgca ggggtaagaa gattgggcat
 gatgtagatt ttttaattac cagcccagga tcagcagagg atgaagagca actttgcct aaagtataa acttatggga aaaaaaggga
 ttactttat attatgacct tgtggagtca acattgaaa agttcaagtt gccaaagcagg caggtggata cttagatca tttcaaaaa
 tgctttctga ttttaaaatt gcaccatcag agagtagaca gtgcaagtc caaccagcag gaaggaaaga cctggaaggc
 catccgtgtg gacctggta tgtgcccta cgagaacctt gcctttgcc tgctaggtg gactggctcc cggcagttg
 agagagacat ccggcgtat gccacacac agcgggaagat gatgctggat aaccagctt tatatgacaa gaccaagagg
 gtatttctca aagcggaaag tgaagaagaa atctttgcac atctgggatt ggactacatt gaacctggg aaagaaatgc ttaggagaaa

gctgtcaact tttttctttt ctgttctttt ttcagggtta gacaaattat gcttcatatt ataataaaag atgccttagt caagtttggg
 attctttaca ttttaccag atgtagattg ctctagaaa taagtagttt tggaaacgtg atcaggcacc ccttgggtta tgctctggca
 agccatttgc aggactgatg tgtagaacgc gcaatgcatt ttccatagaa acagtgttgg aattggtggc tcatttccag ggaagtcat
 caaagcccac tttgccaca gtgtagctga aatactgtat acttgccaat aaaaatagga aac

While commercially-available TdT is suitable for use with the methods of the invention, modified TdT, e.g., having an amino acid sequence at least 95% in common with SEQ ID NO. 1, e.g., having an amino acid sequence at least 98% in common with SEQ ID NO. 1, e.g., having an amino acid sequence at least 99% in common with SEQ ID NO. 1, may be used with the methods of the invention. An organism that expresses a suitable nucleotidyl transferase may comprise a nucleic acid sequence at least 95% in common with SEQ ID NO. 2, e.g., at least 98% in common with SEQ ID NO. 2, e.g., at least 99% in common with SEQ ID NO. 2. In some instances, a modified TdT will result in more efficient generation of polynucleotides, or allow better control of chain length. Other modifications to the TdT may change the release characteristics of the enzyme, thereby reducing the need for aqueous reducing agents such as TCEP or DTT.

For the synthesis of RNA polynucleotides, a nucleotidyl transferase like *E. coli* poly(A) polymerase can be used to catalyze the addition of ribonucleotides to the 3' end of a ribonucleotide initiator. In other embodiments, *E. coli* poly(U) polymerase may be more suitable for use with the methods of the invention. Both *E. coli* poly(A) polymerase and *E. coli* poly(U) polymerase are available from New England Biolabs (Ipswich, MA). These enzymes may be used with 3'unblocked reversible terminator ribonucleotide triphosphates (rNTPs) to synthesize RNA. In certain embodiments, RNA may be synthesized using 3'blocked, 2'blocked, or 2'- 3'blocked rNTPs and poly(U) polymerase or poly(A) polymerase. The amino acid and nucleotide sequences for *E. coli* Poly(A) polymerase and *E. coli* Poly(U) polymerase are reproduced below. Modified *E. coli* Poly(A) polymerase or *E. coli* Poly(U) polymerase may be suitable for use with the methods of the invention. For example, an enzyme, having an amino acid sequence at least 95% in common with SEQ ID NO. 3, e.g., having an amino acid sequence at least 98% in common with SEQ ID NO. 3, e.g., having an amino acid sequence at least 99% in common with SEQ ID NO. 3, may be used with the methods of the invention. An organism that expresses a suitable enzyme may comprise a nucleic acid sequence at least 95% in common with SEQ ID NO. 4, e.g., at least 98% in common with SEQ ID NO. 4, e.g., at least 99% in common with SEQ ID NO. 4. Alternatively, an enzyme having an amino acid sequence at least 95% in common with SEQ ID NO. 5, e.g., having an amino acid sequence at least 98% in common with SEQ ID NO. 5, e.g., having an amino acid

sequence at least 99% in common with SEQ ID NO. 5, may be used with the methods of the invention. An organism that expresses a suitable enzyme may comprise a nucleic acid sequence at least 95% in common with SEQ ID NO. 6, e.g., at least 98% in common with SEQ ID NO. 6, e.g., at least 99% in common with SEQ ID NO. 6.

Table 3: Amino Acid Sequence of *E. coli* Poly(A) polymerase

SEQ ID NO. 3: MFTRVANFCR KVLSREESEA EQAVARPQVT VIPREQHAIS
 RKDISENALK VMYRLNKAGY EAWLVGGGVR DLLLGKKPKD FDVTTNATPE
 QVRKLFNRNC LVGRRFRLAH VMFGPEIIEV ATFRGHHEGN VSDRTTSQRG
 QNGMLLRDNI FGSIEEDAQR RDFTINSLYY SVADFTVRDY VGGMKDLKDG
 VIRLIGNPET RYREDPVRML RAVRFAAKLG MRISPETAEP IPRLATLLND IPPARLFEEES
 LKLLQAGYGY ETYKLLCEYH LFQPLFPTIT RYFTENGDSP MERIIEQVLK NTDTRIHNDM
 RVNPAFLFAA MFWYPLLETA QKIAQESGLT YHDAFALAMN DVLDEACRSL
 AIPKRLTTLT RDIWQLQLRM SRRQGKRAWK LLEHPKFRAA YDLLALRAEV
 ERNAELQRLV KWWGEFQVSA PPDQKGMLE LDEEPSRRR TRRPRKRAPR REGTA

The nucleotide sequence corresponding to *E. coli* poly(A) polymerase is listed in Table 4 as SEQ ID NO. 4.

Table 4: Nucleotide Sequence of *E. coli* Poly(A) polymerase

SEQ ID NO. 4: atttttacc gagtcgctaa tttttgccgc aaggtgctaa gccgcgagga aagcgagget gaacaggcag
 tcgcccgtcc acaggtgacg gtgatcccgc gtgacgagca tgctatttcc cgcaaagata tcagtaaaa tgcctgaag
 gtaatgtaca ggctcaataa agcgggatac gaagcctggc tggttggcgg cggcgtgcgc gacctgttac ttggcaaaaa
 gccgaaagat tttgacgtaa cactaacgc cacgcctgag caggtgcgca aactgttccg taactgccgc ctggtgggtc
 gccgttccg tctggetcat gtaatgtttg gcccgagat tategaagt ggcacctec gtggacacca cgaagtaac gtcagcgacc
 gcacgacctc ccaacgcggg caaacggca tgttgctgcg cgacaacatt ttcggctcca tcgaagaaga cgcccagegc
 cgcgatttca ctatcaacag cctgtattac agcgtagcgg atttaccgt ccgtgattac gttggcggca tgaaggatct gaaggacggc
 gttatccgtc tgattggtaa cccggaaacg cgctaccgtg aagatccggt acgtatgctg cgcgeggtag gttttgccgc caaattgggt
 atgcgcatca gcccgaaac cgcagaaccg atccctgcc tcgctaccct gctgaacgat atcccaccg cacgcctgtt
 tgaagaatcg cttaaactgc tacaagcggg ctacggttac gaaacctata agctgtgtg tgaatatcat ctgtccage cgctgttccc
 gaccattacc cgctacttca cggaaaatgg cgacagcccc atggagcgga tcattgaaca ggtgctgaag aataccgata
 cgcgtatcca taacgatatg cgcgtgaacc cgcgcttct gtttccgcgc atgttctggt acccaactgct ggagacggca
 cagaagatcg cccaggaaag cggcctgacc taccacgacg ctftcgcgct ggcgatgaac gacgtgctgg acgaagcctg
 ccgttcactg gcaatcccga aacgtctgac gacattaacc cgcgatatct ggcagttgca gttgcgtatg tcccgtcgtc agggtaaacc

cgcatgaaa ctgctggagc atcctaagtt ccgtgcggct tatgacctgt tggccttgcg agctgaagtt gagcgtaac
 ctgaactgca ggcgtctggg aaatgggtggg gtgagttcca ggtttccgcg ccaccagacc aaaaagggat gctcaacgag
 ctggatgaag aaccgtcacc ggcgtcgtct actcgtcgtc cacgcaaagc cgcaccacgt cgtgagggta ccgcatga

Table 5: Amino Acid Sequence of *E. coli* Poly(U) polymerase

SEQ ID NO. 5: GSHMSYQKVP NSHKEFTKFC YEVYNEIKIS DKEFKEKRAA
 LDTLRLCLKR ISPDAELVAF GSLESGLALK NSDMDLCVLM DSRVQSDTIA LQFYEELIAE
 GFEGKFLQRA RIPIIKLTSD TKNFGFASQ CDIGFNNRLA IHNTLLLSSY TKLDARLKPM
 VLLVKHWAKR KQINSPYFGT LSSYGYVLMV LYYLIHVIKP PVFPNLLLSP
 LKQEKIVDGF DVGFDKLED IPPSQNYSSL GSLLHGFFRF YAYKFEPREK
 VVTFRRPDGY LTKQEKGWTS ATEHTGSADQ IIKDRYILAI EDPFEISHNV GRTVSSSGLY
 RIRGEFMAAS RLLNSRSYPI PYDSLFEAA

The nucleotide sequence corresponding to *E. coli* poly(U) polymerase is listed in Table 6 as SEQ ID NO. 6.

Table 6: Nucleotide Sequence of *E. coli* Poly(A) polymerase

SEQ ID NO. 6: ggcagccata tgagctatca gaaagtgccg aacagccata aagaatttac caaattttgc tatgaaggt
 ataacgaaat taaaattagc gataaagaat taaagaaaa acgcgcggcg ctggataccc tgcgcctgtg cctgaaacgc
 attagcccgg atgcggaact ggtggcggtt ggcagcctgg aaagcggcct ggcgctgaaa aacagcgata tggatctgtg
 cgtgctgatg gatagccgcg tgcagagcga taccattgcg ctgcagtttt atgaagaact gattgcggaa ggctttgaag gcaaatttct
 gcagcgcg cgcatccga ttattaaact gaccagcgt accaaaaacg gctttggcgc gagctttcag tgcgatattg gctttaacaa
 ccgcctggcg atcataaca cctgctgct gagcagctat accaaactgg atgcgcgctt gaaaccgatg gtgctgctgg
 tgaaacattg ggcgaaacgc aaacagatta acagcccgta ttttggcacc ctgagcagct atggctatgt gctgatggtg ctgtattatc
 tgattcatgt gattaaaccg cgggtgttc cgaacctgct gctgagcccc ctgaaacagg aaaaaattgt ggatggcttt gatgtgggct
 ttgatgataa actggaagat attccgccga gccagaacta tagcagcctg ggcagcctgc tgcattgctt tttcgttt tatgcgtata
 aattgaacc gcgcgaaaaa gtggtgacct ttcgccgcc gtaggctat ctgaccaaac aggaaaaagg ctggaccagc
 gcgaccgaac ataccggcag cgcggatcag attattaaag atcgctatat tctggcgatt gaagatccgt ttgaaattag ccataacgtg
 ggccgcaccg tgagcagcag cggcctgtat cgcattcgcg gcgaatttat ggcggcgagc cgctgctga acagccgcag
 ctatccgatt ccglatgata gcctgtttga agaagcg

As discussed above, the inhibitor coupled to the nucleotide analog will cause the transferase, e.g., TdT, to not release from the polynucleotide or prevent other analogs from being incorporated into the growing chain. A charged moiety results in better inhibition, however,

research suggests that the specific chemical nature of the inhibitor is not particularly important. For example, both phosphates and acidic peptides can be used to inhibit enzymatic activity.

See, e.g., Bowers et al., Nature Methods, vol. 6, (2009) p. 593-95, and U.S. Pat. No. 8,071,755, both of which are incorporated herein by reference in their entireties. In some embodiments, the inhibitor will include single amino acids or dipeptides, like $-(\text{Asp})_2$, however the size and charge on the moiety can be adjusted, as needed, based upon experimentally determined rates of first nucleotide incorporation and second nucleotide incorporation. That is, other embodiments may use more or different charged amino acids or other biocompatible charged molecule.

Other methods of nucleotide synthesis may be used to build *de novo* oligonucleotides in a template independent fashion using nucleotidyl transferases or modified nucleotidyl transferases. In one embodiment, the polymerase/transferase enzymes can be modified so that they cease nucleotide addition when they encounter a modification to the phosphate of a 3'-unmodified dNTP analog. This scheme would require a deblocking reagent/reaction that modifies the phosphate end of the nucleotide analog, which frees up the nascent strand for subsequent nucleotide incorporation. Preferred embodiments of this approach would use nucleotide analogs modified only at the phosphates (alpha, beta or gamma) although modifications of the purine/pyrimidine base of the nucleotide are allowed.

In some embodiments, it may be advantageous to use a 3' exonuclease to remove oligonucleotides that have not been properly terminated with an inhibitor prior to subsequent nucleotide analog addition. In particular, the inhibitor of the nucleotide analog can be chosen to inhibit the activity of nucleotidyl transferase and 3' exonucleases, such that only properly terminated oligonucleotides would be built up. Using this quality control technique, the purity of the resulting oligonucleotide sequences would be improved. In some embodiments, use of such quality control measures can negate the need for post-synthesis purification. This technique is represented schematically in FIG. 19, where a 3' exonuclease is introduced after a wash step to remove excess nucleotide analogs and prior to linker cleavage. Such a cleaning step, as shown in FIG. 19, will reduce the number of oligonucleotides that are of an undesired length and/or sequence.

Another embodiment for using non-template dependent polymerase/transferase enzymes would be to using protein engineering or protein evolution to modify the enzyme to remain tightly bound and inactive to the nascent strand after each single nucleotide incorporation, thus preventing any subsequent incorporation until such time as the polymerase/transferase is released from the strand by use of a releasing reagent/condition. Such modifications would be selected to allow the

use of natural unmodified dNTPs instead of reversible terminator dNTPs. Releasing reagents could be high salt buffers, denaturants, etc. Releasing conditions could be high temperature, agitation, etc. For instance, mutations to the Loop1 and SD1 regions of TdT have been shown to dramatically alter the activity from a template-independent activity to more of a template dependent activity. Specific mutations of interest include but are not limited to $\Delta 3384/391/392$, del loop1 (386-398), L398A, D339A, F401A, and Q402K403C404-E402R403S404. Other means of accomplishing the goal of a post-incorporation tight binding (i.e., single turnover) TdT enzyme could include mutations to the residues responsible for binding the three phosphates of the initiator strand including but not limited to K261, R432, and R454.

Another embodiment for using non-template dependent polymerase/transferase enzymes is to use protein engineering or protein evolution to modify the enzyme to accept 3'-blocked reversible terminators with high efficiency. Naturally occurring polymerase/transferase enzymes will not incorporate 3'-blocked reversible terminators due to steric constraints in the active site of the enzyme. Modifying either single or several amino acids in the active site of the enzyme can allow the highly-efficient incorporation of 3'-blocked reversible terminators into a support bound initiator in a process completely analogous to that described above. After incorporation, the 3'-reversible terminator is removed with a deblocking reagent/condition thus generating a completely natural (scarless) single strand molecule ready for subsequent controlled extension reactions. The enzyme contains amino acids close to the 3'-OH of the incoming dNTP which explains the propensity of TdT for incorporating ribonucleotide triphosphates as readily as deoxyribonucleotide triphosphates; amino acids including but not limited to those between $\beta 1$ and $\beta 2$ especially R334, Loop1, and those between $\alpha 13$ and $\alpha 14$, especially R454, are likely targets for mutagenesis to accommodate the bulk of 3'-reversible terminator groups and allow their efficient incorporation. In certain embodiments additional amino acid changes may be required to compensate for alterations made in order to accommodate a 3'-reversible terminator. Another embodiment for using template-dependent polymerases is to use the either 3'-blocked or 3'-unblocked dNTP analogs with a plurality of primer-template pairs attached to a solid support where the template is a nucleic acid analog that supports polymerase mediated primer extension of any of the four bases as specified by the user.

In some embodiments, an engineered TdT is used to achieve stepwise synthesis with 3'-O-blocked nucleotide analogs. It is possible to model the active site of the TdT protein using AutoDock (Molecular Graphics Laboratory, Scripps Research Institute, La Jolla, CA). Based upon this calculation, it is predicted that modified TdTs, having changes at Arg336 and Arg454 may

have enzymatic activity against 3'-O-blocked nucleotide analogs. It is thought that Gly452 and Ser453 exist in a cis-peptide bond conformation (see Delarue et al., EMBO J., 2002; 21(3):427-39, incorporated herein by reference in its entirety) and that the guanidinium group of Arg336 assists in the stabilization of this conformation. The stability provided by Arg336 may help explain why substitutions at this position have a negative impact on the reactivity of modified TdT proteins. In some instances, the instability created by modifying position 336 may be overcome by using proline to stabilize cis-peptide bond conformation. However, if Arg336 is substituted, e.g., with alanine or glycine, the entire TGSR motif (positions 451, 452, 435, 454) may also have to be modified to compensate for this change. For example, the TGSR motif may be modified to TPRS or TGPR. In another embodiment, substitutions at Arg454 to accommodate the steric bulk of a 3'-O-blocking group may require additional modifications to the α 14 region to compensate for substitutions of glycine or alanine at Arg454. In other embodiments, substitutions for other amino acids in the α 11 region may be required to compensate for substitution to Arg336 either instead of, or in addition to, modification of the GSR motif.

While modification to Arg336 and Arg454 may change the binding interactions of 3'-O-modified dNTPs, it may also be necessary to explore substitutions that would result in improved steric interactions of 3'-O-modified dNTPs with TdT. Such steric modifications can also be explored computationally. Residues Gly332, Gly333, Gly452, Thr451 and Ser453 are also potential targets for substitution to allow the extra steric bulk of a 3'-blocking group like 3'-O-azidomethyl or 3'-O-allyl. Residues that are within 1.2 nm of the 3'-OH such as Glu457, Ala510, Asp509, Arg508, Lys199, Ser196 Met192, or Leu161 may also potentially interfere with the substrate utilization of a 3'-O-blocked dNTP and are thus targets for substitution in addition to or in combination with Arg336 and Arg454. In addition to amino acid substitutions at positions 508, 509 and 510, it may be necessary to delete amino acids in order to remove interference with a 3'-O-blocking group. Since those amino acids are located near the C-terminus of the protein, and exist in a relatively unstructured region, they may be deleted singly or altogether, either instead of or in combination with the modifications described above.

Another embodiment for using non-template dependent polymerase/transferase enzymes can use protein engineering or protein evolution to modify the enzyme to optimize the use of each of the four different nucleotides or even different modified nucleotide analogs in an analog specific manner. Nucleotide specific or nucleotide analog specific enzyme variants could be engineered to possess desirable biochemical attributes like reduced K_M or enhanced addition rate which would further reduce the cost of the synthesis of desired polynucleotides.

SOLID STATE SYNTHESIS

The methods of the invention can be practiced under a variety of reaction conditions, however the orderly construction and recovery of desired polynucleotides will, in most cases, require a solid support on which the polynucleotides can be extended. When used in conjunction with the NTP, linker, and inhibitor analogs discussed above, it is possible to construct specific polynucleotide sequences of DNA, as well as RNA, by using, for example, TdT or poly(A) polymerase in an aqueous environment. As shown in FIG. 13, the TdT can be used to effect the stepwise construction of custom polynucleotides by extending the polynucleotide sequence in a stepwise fashion. As discussed previously, the inhibitor group of each NTP analog causes the enzyme to stop with the addition of a nucleotide. After each nucleotide extension step, the reactants are washed away from the solid support prior to the removal of the inhibitor by cleaving the linker, and then new reactants are added, allowing the cycle to start anew.

In certain embodiments, an additional quality control step may be incorporated in which the oligonucleotide or polynucleotide is exposed to 3' exonuclease after the nucleotidyl transferase-mediated nucleotide analog extension step and before inhibitor cleavage. A 3' exonuclease degrades oligonucleotide or polynucleotide strands with an unblocked 3' OH. An uncleaved inhibitor (e.g., a steric inhibitor) may physically block the 3' exonuclease from degrading a strand to which an uncleaved nucleotide analog has been successfully incorporated. Such a quality control step degrades only the oligonucleotides or polynucleotides that have unsuccessfully incorporated the desired nucleotide analog in the prior addition step, thereby eliminating any errors in the finished synthesized sequence. After 3' exonuclease exposure, the enzyme may be washed away before carrying on with the inhibitor cleavage step.

The 3' exonuclease acts by shortening or completely degrading strands that have not successfully added the desired nucleotide analog. Strands that fail to be enzymatically extended at a given cycle will not have a terminal macromolecule-dNMP conjugate prior to the linker cleavage step. If a 3'-exonuclease is introduced at this stage, the full length strand could be protected from degradation while "failure" strands will be shortened in length or potentially degraded completely to mononucleotide phosphates. The yield of long (>500 bases) synthetic DNA is dependent on highly efficient reactions occurring at each and every cycle; both the enzymatic extension and the deblocking/self-elimination steps must occur at near quantitative yields. The introduction of a 3'-exonuclease after the enzymatic extension step but before the macromolecule terminator cleavage step has a positive impact on the nascent strand purity if the extension efficiency is low (i.e., there

are strands that are not extended and therefore possess a natural unmodified terminal nucleotide).

Conversely, a 3'-exonuclease step would have no impact on the quality of the synthesis if the deblocking/elimination step is less than quantitative because those strands would still be protected by the macromolecule terminator and fail to extend during the next extension step. Thus the actual improvement of the quality of the synthesis with the addition of a 3'-exonuclease step can only be experimentally determined and then an assessment made if it is worth the additional cost and cycle time.

At the conclusion of n cycles of extension-remove-deblocking-wash, the finished full-length, single-strand polynucleotide is complete and is cleaved from the solid support and recovered for subsequent use in applications such as DNA sequencing or PCR. Alternatively, the finished, full-length, single-strand polynucleotide can remain attached to the solid support for subsequent use in applications such as hybridization analysis, protein or DNA affinity capture. In other embodiments, partially double-stranded DNA can be used as an initiator, resulting in the synthesis of double-stranded polynucleotides.

In some embodiments, a nucleic acid initiator will include a 3' moiety that will release the synthesized oligonucleotide when in the presence of a releasing agent. This feature is illustrated, generally, in FIG. 22, where a nucleic acid initiator (5'-initiator-) is shown coupled to a solid state support (open circle) and a releasable 3' moiety (open star). In some embodiments the initiator is a single-stranded oligonucleotide, such as a dimer, trimer, tetramer, pentamer, hexamer, septamer, or octomer. Because the 3' moiety attached to the initiator is a substrate for the enzyme, e.g., a TdT, e.g., a modified TdT, the enzyme can add additional nucleotides or nucleotide analogs in a stepwise fashion. With each addition, the length of the synthesized oligonucleotide increases. Once the oligonucleotide synthesis is complete, a releasing agent can be introduced to cause the 3' moiety to decouple from the nucleic acid initiator. In some embodiments, the 3' moiety is a ribonucleotide, such as an A, C, G, or U ribonucleotide. In other embodiments, the 3' moiety is an abasic deoxyribose. In other embodiments, the 3' moiety is an abasic ribose. In other embodiments, the 3' moiety is a non-nucleoside 5'-monophosphate. The releasing agent may include a basic solution or a metal ion. For example, the releasing agent can be a concentrated NH₄OH solution, having a pH greater than 8, i.e., greater than pH 8.5, i.e., greater than pH 9.0, i.e., greater than pH 9.5. In some embodiments, the releasing agent will be an enzyme, such as a type II restriction nuclease. In some embodiments, the enzyme will uniquely interact with the nucleic acid sequence of the initiator, and lyse the synthesized oligonucleotide from the initiator, leaving behind the initiator.

In an embodiment, the initiator is a nucleic acid hexamer and the 3' moiety is a ribonucleotide, such as adenosine. Once the oligonucleotide synthesis is complete, e.g., using nucleotides comprising a cleavable terminator linked at the N-4 position, or nucleotides having a 3'-O-blocked position, the oligonucleotide can be released by exposing the bound oligonucleotide to an ammonium hydroxide solution of a pH around 8. The basic solution containing the synthesized oligonucleotides can then be separated from the solid substrate comprising the hexamer initiator. The solid substrate is then washed and/or neutralized to prepare the initiator and 3' moiety for fabrication of a new oligonucleotide. In some embodiments, the terminal ribonucleotide is regenerated prior to oligonucleotide synthesis with the use of a phosphatase or the 3' phosphatase activity of T4 polynucleotide kinase.

In some embodiments, the solid support and nucleic acid initiator including the 3' moiety will be reusable, thereby allowing the initiator coupled to the solid support to be used again and again for the rapid synthesis of oligonucleotides. Solid supports suitable for use with the methods of the invention may include glass and silica supports, including beads, slides, pegs, or wells. In some embodiments, the support may be tethered to another structure, such as a polymer well plate or pipette tip. In some embodiments, the solid support may have additional magnetic properties, thus allowing the support to be manipulated or removed from a location using magnets. In other embodiments, the solid support may be a silica coated polymer, thereby allowing the formation of a variety of structural shapes that lend themselves to automated processing.

The selection of substrate material and covalent linkage chemistry between initiator and substrate is limited only by the ability of the construct to withstand the synthesis conditions without loss of initiator. Preferred embodiments utilize substrates and linkers of greater chemical stability than the initiator so that the overall construct stability is that of the attached oligonucleotide and not dependent on the substrate. In some embodiments, initiators may be synthesized in a 5' to 3'-direction from a material presenting surface hydroxyl groups, though in preferred embodiments the initiator is instead grafted to the substrate so that density and initiator quality can be precisely controlled.

The covalent linkage between the initiator and the substrate may be any bond which does not compromise the stability of the construct. Preferred embodiments may utilize couplings between oligonucleotides containing either 5'-amine, 5'-hydroxyl, 5'-phosphate, 5'-sulfhydryl, or 5'-benzaldehyde groups and surfaces or resins containing formyl, chloromethyl, epoxide, amine, thiol, alkene, or terminal C-F bonds on the substrate.

In some embodiments, the initiator may contain elements for sequence-specific cleavage

strategies, such as those utilizing restriction enzymes, uracil specific excision reagent (USER), or any variety of sequence-specific nuclease. In other embodiments, these elements may instead be enzymatically synthesized or added to the initiator after it has been coupled to a resin. Such a scenario may be preferred when, for example, the initiator is comprised of entirely thymidine in order to minimize potential side reactions with surface functional groups during the coupling process.

FIG. 23 shows a solution-phase example where a short poly-uracil (U) tract is added to the 3'-end of an initiator using TdT. The resultant initiator may then be used for further extension reactions as described to generate the desired sequence. Upon completion of the synthesis, the newly generated sequence may be cleaved at the internal poly-U tract to separate the new sequence from the original initiator. The 30 nucleotide initiator used is of the sequence 5'-TTATTATTATTATTAAAAAGGCCAAAAAA. The gel shown in FIG. 23 includes the initiator run in lane 1 while the initiator after one or more additions of dUTP (i.e., the poly-uracil tract) was run in lane 2. The generated sequence was then subjected to USER digestion for 10, 30, and 60 minutes and run in lanes 3, 4, and 5 respectively.

FIG. 24 shows that the cleavage of a sequence containing an internal poly-U tract produces a single product of homogenous length bearing a 5'-phosphate. The cleavage process leaves a 3'-phosphate on the resin-bound initiator, which may be removed with treatment with T4 polynucleotide kinase, alkaline phosphatase or other dephosphorylation process (See FIG. 25). The internal poly-U tract sequence cleaved in the example depicted in FIG. 24 was of sequence 5'-TTAATTAATUUUUGTGAGCTTAATGTCCTTATGT which, after USER digestion, resulted in a product of sequence 5'-phos-GTGAGCTTAATGTCCTTATGT. The resulting product was run after 0, 10, 30, and 60 minutes of USER digestion in lanes, 1, 2, 3, and 4 as shown in FIG. 24, indicating the successful cleavage and homogenous length of the products. In lane 5, a control oligonucleotide of sequence GTGAGCTTAATGTCCTTATGT was run.

FIG. 25 shows an exemplary solid-phase dephosphorylation and TdT extension process. The chart depicts the absorbance of the TdT reaction mixture in a solid-phase extension. As dNTPs are added to the resin-bound initiator, they are depleted from solution, reducing absorbance of the supernatant. Trace A shows the negligible rate of dNTP incorporation onto a resin where the 3'-end of the initiator is blocked with a terminal phosphate. Trace B shows the rate of incorporation after such an initiator is treated with shrimp alkaline phosphatase for 15-minutes. Trace C shows the rate of dNTP incorporation under the same conditions using an unblocked initiator.

After cleavage of the synthesized sequence, the resin may then be used in another cycle of

cleavage site installation, synthesis, product removal, and regeneration. FIG. 26 provides an overview of an exemplary regeneration cycle.

In other embodiments, the cleavage site installation may be used to homogenize the enzymatic accessibility of the initiator oligonucleotides. Each enzyme used during the enzymatic synthesis cycle has its own steric footprint with potentially distinct optimal loadings and spacing from the surface. This can produce unexpected behavior in regard to the kinetics of stepwise addition and yield from the enzymatic cleavage process. In some embodiments, repeated cycles of cleavage site installation, cleavage, and regeneration may be conducted prior to the oligonucleotide synthesis so that the cleavage enzyme and the template-independent polymerase are accessing the same population of surface oligonucleotides.

Some embodiments may install multiple different cleavage sites throughout a strand during synthesis. Upon digestion, a complex library of the strands located between the cleavage sites may be released from the resin. Such sequences can then be further amplified and used for enzymatic assembly processes by a skilled artisan. This approach may be uniquely suited to parallel synthesis schemes in order to produce greater varieties of sequence fragments with relatively few distinct locations on a surface, or to avoid the synthesis of contiguous strands which risk secondary structure formation during synthesis.

In other embodiments, cleavage site installation may be used immediately after each cycle of enzymatic extension to assist in the removal of unreacted initiator sequences, analogous to the acetylation used in phosphoramidite-based oligonucleotide synthesis. Strands which are unextended during the addition cycle act as substrates for a poly-U tract, while extended initiators bearing a chain terminator do not. At the end of the synthesis cycle, a single USER digestion is conducted so that the oligonucleotides are released from the support, and sequences containing failed additions are simultaneously digested to a shorter length than the full-sized product. Failure sequences are also alike in that they are now phosphorylated at their 3'-ends, rendering them unreactive to further enzymatic extension. The full-length sequences may then undergo extension to append any element which will enable selective capture, isolation, or enrichment. Such elements may either be additional homopolymer tracts, such as further poly-U tracts, which can be isolated by hybridization-based approaches and subsequently digested, or biotinylated elements for non-covalent capture. This approach compensates for the lack of suitable chromatographic techniques suitable for long (150 nt+ oligonucleotides), low sample quantities, or complex mixtures of sequences of varied lengths.

Further embodiments may employ cleavage site installations to assist in the recycling of

highly-indexed DNA-based data recording media. An example is shown in FIG. 27. In such cases, 5'-surface-immobilized sequences are terminated with a short poly-A tract. Extension to produce a terminal 3'-poly-U site of sufficient length will allow a hairpin to fold under appropriate conditions, so that the elements of the initiator preceding the poly-A stretch can be replicated using a template-dependent polymerase. The sequence can then be extended with a new homopolymer tract to leave a free 3'-terminus which can be used in subsequent data-writing operations. Upon completion of the write steps, the hairpin linker can then be digested with the USER enzyme to release the data strand, while leaving the template initiator to be regenerated with a 3'-dephosphorylation step, poly-U addition, and recopying of the template strand. The use of homopolymer tracts in DNA-based data recording is described in co-owned U.S. Pat. App. Ser. No. 15/994,335, incorporated herein by reference.

SYNTHESIZERS

To capitalize on the efficiency of the disclosed methods, an aqueous phase DNA synthesizer can be constructed to produce desired polynucleotides in substantial quantities. In one embodiment, a synthesizer will include four wells of the described NTP analog reagents, i.e., dCTP, dATP, dGTP, and dTTP, as well as TdT at concentrations sufficient to effect polynucleotide growth. A plurality of initiating sequences can be attached to a solid support that is designed to be repeatedly dipped into each of the four wells, e.g., using a laboratory robot.

The robot could be additionally programmed to rinse the solid support in wash buffer between nucleotide additions, cleave the linking group by exposing the support to a deblocking agent, and wash the solid support a second time prior to moving the solid support to the well of the next desired nucleotide. With simple programming, it is possible to create useful amounts of desired nucleotide sequences in a matter of hours, and with substantial reductions hazardous waste. Ongoing synthesis under carefully controlled conditions will allow the synthesis of polynucleotides with lengths in the thousands of base pairs. Upon completion, the extension products are released from the solid support, whereupon they can be used as finished nucleotide sequences.

A highly parallel embodiment could consist of a series of initiator-solid supports on pegs in either 96 or 384 well formats that could be individually retracted or lowered so that the pegs can be indexed to contact the liquids in the wells in a controlled fashion. The synthesizer could thus consist of the randomly addressable peg device, four enzyme-dNTP analog reservoirs in the same format as the peg device (96 or 384 spacing), additional reagent reservoirs (washing, deblocking, etc.) in the same format as the peg device (96 or 384 spacing), and a transport mechanism (e.g., a

laboratory robot) for moving the peg device from one reservoir to another in a user programmable controlled but random access fashion. Care must be taken to avoid contaminating each of the four enzyme-dNTP reservoirs since the contents are reused throughout the entire synthesis process to reduce the cost of each polynucleotide synthesis.

In alternative embodiments, the reagents (e.g., nucleotide analogs, enzymes, buffers) will be moved between solid supports, allowing the reagents to be recycled. For example a system of reservoirs and pumps can move four different nucleotide analog solutions, wash buffers, and/or reducing agent solutions between one or more reactors in which the oligonucleotides will be formed. The reactors and pumps can be conventional, or the devices may be constructed using microfluidics. Because of the non-anhydrous (aqueous) nature of the process, no special care needs to be taken in the design of the hardware used to eliminate exposure to water. The synthesis process can take place with only precautions to control evaporative loss. A highly parallel embodiment could consist of a monolithic series of initiator-solid supports on pegs in either 96 or 384 well format that can be interfaced to a series of wells in the same matching format. Each well would actually be a reaction chamber that is fed by four enzyme-dNTP analog reservoirs and additional reagent reservoirs (washing, deblocking, etc.) with appropriate valves. Provisions would be made in the fluidics logic to recover the enzyme-dNTP reactants in a pristine fashion after each extension reaction since they are reused throughout the entire synthesis process to reduce the cost of each polynucleotide synthesis. In other embodiments, a system of pipetting tips could be used to add and remove reagents.

In certain aspects, polynucleotides may be synthesized using microfluidic devices and/or inkjet printing technology. An exemplary microfluidic polynucleotide synthesis device is shown in FIG. 17 for illustrative purposes and not to scale. Microfluidic channels 255, including regulators 257, couple reservoirs 253 to a reaction chamber 251 and an outlet channel 259, including a regulator 257 can evacuate waste from the reaction chamber 251. Microfluidic devices for polynucleotide synthesis may include, for example, channels 255, reservoirs 253, and/or regulators 257. Polynucleotide synthesis may occur in a microfluidic reaction chamber 251 which may include a number of anchored synthesized nucleotide initiators which may include beads or other substrates anchored or bound to an interior surface of the reaction chamber and capable of releasably bonding a NTP analog or polynucleotide initiator. The reaction chamber 251 may include at least one intake and one outlet channel 259 so that reagents may be added and removed to the reaction chamber 251. The microfluidic device may include a reservoir 253 for each respective NTP analog. Each of these NTP analog reservoirs 253 may also include an appropriate

amount of TdT or any other enzyme which elongates DNA or RNA strands without template direction. Additional reservoirs 253 may contain reagents for linker/inhibitor cleavage and washing. These reservoirs 253 can be coupled to the reaction chamber 251 via separate channels 255 and reagent flow through each channel 255 into the reaction chamber 251 may be individually regulated through the use of gates, valves, pressure regulators, or other means. Flow out of the reaction chamber 251, through the outlet channel 259, may be similarly regulated.

In certain instances, reagents may be recycled, particularly the NTP analog-enzyme reagents. Reagents may be drawn back into their respective reservoirs 253 from the reaction chamber 251 via the same channels 255 through which they entered by inducing reverse flow using gates, valves, pressure regulators or other means. Alternatively, reagents may be returned from the reaction chamber 251 to their respective reservoirs 253 via independent return channels. The microfluidic device may include a controller capable of operating the gates, valves, pressure, or other regulators 257 described above.

An exemplary microfluidic polynucleotide synthesis reaction may include flowing a desired enzyme-NTP analog reagent into the reaction chamber 251; after a set amount of time, removing the enzyme-NTP analog reagent from the reaction chamber 251 via an outlet channel 259 or a return channel; flowing a wash reagent into the reaction chamber 251; removing the wash reagent from the reaction chamber 251 through an outlet channel 259; flowing a de-blocking or cleavage reagent into the reaction chamber 251; removing the de-blocking or cleavage reagent from the reaction chamber 251 via an outlet channel 259 or a return channel; flowing a wash reagent into the reaction chamber 251; removing the wash reagent from the reaction chamber 251 through an outlet channel 259; flowing the enzyme-NTP analog reagent including the next NTP in the desired sequence to be synthesized into the reaction chamber 251; and repeating until the desired polynucleotide has been synthesized. After the desired polynucleotide has been synthesized, it may be released from the reaction chamber anchor or substrate and collected via an outlet channel 259 or other means.

In certain aspects, reagents and compounds, including NTP analogs, TdT and/or other enzymes, and reagents for linker/inhibitor cleavage and/or washing may be deposited into a reaction chamber using inkjet printing technology or piezoelectric drop-on-demand (DOD) inkjet printing technology. Inkjet printing technology can be used to form droplets, which can be deposited, through the air, into a reaction chamber. Reagent droplets may have volumes in the picoliter to nanoliter scale. Droplets may be introduced using inkjet printing technology at a variety of frequencies including 1 Hz, 10 Hz, 100 Hz, 1 kHz, 2 kHz, and 2.5 kHz. Various reagents may

be stored in separate reservoirs within the inkjet printing device and the inkjet printing device may deliver droplets of various reagents to various discrete locations including, for example, different reaction chambers or wells within a chip. In certain embodiments, inkjet and microfluidic technologies may be combined wherein certain reagents and compounds are delivered to the reaction chamber via inkjet printing technology while others are delivered via microfluidic channels or tubes. An inkjet printing device may be controlled by a computing device comprising at least a non-transitory, tangible memory coupled to a processor. The computing device may be operable to receive input from an input device including, for example, a touch screen, mouse, or keyboard and to control when and where the inkjet printing device deposits a droplet of reagent, the reagent it deposits, and/or the amount of reagent deposited.

In certain instances, a desired polynucleotide sequence may be entered into the computing device through an input device wherein the computing device is operable to perform the necessary reactions to produce the desired polynucleotide sequence by sequentially depositing the appropriate NTP analog, enzyme, cleavage reagent, and washing reagent, in the appropriate order as described above.

After synthesis, the released extension products can be analyzed by high resolution PAGE to determine if the initiators have been extended by the anticipated number of bases compared to controls. A portion of the recovered synthetic DNA may also be sequenced to determine if the synthesized polynucleotides are of the anticipated sequence.

Because the synthesizers are relatively simple and do not require the toxic components needed for phosphoramidite synthesis, synthesizers of the invention will be widely accessible for research institutions, biotechnology companies, and hospitals. Additionally, the ability to reuse/recycle reagents will reduce the waste produced and help reduce the costs of consumables. The inventors anticipate that the methods and systems will be useful in a number of applications, such as DNA sequencing, PCR, and synthetic biology.

Incorporation by Reference

References and citations to other documents, such as patents, patent applications, patent publications, journals, books, papers, web contents, have been made throughout this disclosure. All such documents are hereby incorporated herein by reference in their entirety for all purposes.

Equivalents

Various modifications of the invention and many further embodiments thereof, in addition

to those shown and described herein, will become apparent to those skilled in the art from the full contents of this document, including references to the scientific and patent literature cited herein. The subject matter herein contains important information, exemplification and guidance that can be adapted to the practice of this invention in its various embodiments and equivalents thereof.

CLAIMS

1. A composition comprising a solid support and an oligonucleotide, wherein the 5' end of the oligonucleotide is coupled to the solid support, and wherein the oligonucleotide comprises a sequence-specific cleavage element.
2. The composition of claim 1 wherein the oligonucleotide comprises a 5'-amine, 5'-hydroxyl, 5'-phosphate, 5'-sulfhydryl, or 5'-benzaldehyde group coupled to the solid support comprising formyl, chloromethyl, epoxide, amine, thiol, alkene, or terminal C-F bonds.
3. The composition of claim 1 wherein the oligonucleotide is irreversibly coupled to the solid support.
4. The composition of claim 1 wherein the sequence-specific cleavage element is incorporated into the oligonucleotide after the 5' end has been irreversibly coupled to the solid support.
5. The composition of claim 4, wherein the sequence-specific cleavage element is incorporated enzymatically.
6. The composition of claim 1, wherein the sequence-specific cleavage element comprises a poly-U tract.
7. The composition of claim 1, wherein the support-coupled oligonucleotide comprises an index element to act as a template during oligonucleotide strand synthesis from a 3' end of the support-coupled oligonucleotide.
8. A method for synthesizing oligonucleotides on a reusable solid support, the method comprising:
 - exposing a nucleic acid initiator attached to a solid support to nucleotide analogs in the presence of a polymerase to create a first oligonucleotide, wherein the nucleic acid initiator comprises a sequence-specific cleavage element;
 - contacting the nucleic acid initiator with a releasing agent to cleave the sequence-specific cleavage element and release the first oligonucleotide from the nucleic acid initiator;

dephosphorylating a 3'-end of the solid support-attached nucleic acid initiator; and enzymatically renewing the sequence-specific cleavage element.

9. The method of claim 8 wherein the nucleic acid initiator comprises a 5'-amine, 5'-hydroxyl, 5'-phosphate, 5'-sulfhydryl, or 5'-benzaldehyde group attached to the solid support comprising formyl, chloromethyl, epoxide, amine, thiol, alkene, or terminal C-F bonds.

10. The method of claim 8 wherein the nucleic acid initiator is irreversibly attached to the solid support.

11. The method of claim 8 wherein the sequence-specific cleavage element is incorporated into the nucleic acid initiator after a 5' end of the nucleic acid initiator has been irreversibly coupled to the solid support.

12. The method of claim 11 wherein the sequence-specific cleavage element is incorporated enzymatically.

13. The method of claim 8 wherein the sequence-specific cleavage element comprises a poly-U tract.

14. The method of claim 8 wherein the nucleic acid initiator comprises an index element, the method further comprising:

extending a 3' end of the nucleic acid initiator with complementary nucleotide analogs to a 3' sequence of the nucleic acid initiator to synthesize the first oligonucleotide under appropriate conditions to permit hairpin formation;

extending the first oligonucleotide after hairpin formation using the index element as a template.

15. The method of claim 14 wherein the 3' sequence of the nucleic acid initiator is complimentary to the sequence-specific cleavage element.

16. The method of claim 14, further comprising

incorporating spacer nucleotides into the first oligonucleotide sufficient to create a single

stranded 3' end thereof subsequent to the hairpin.

17. The method of claim 8, further comprising:

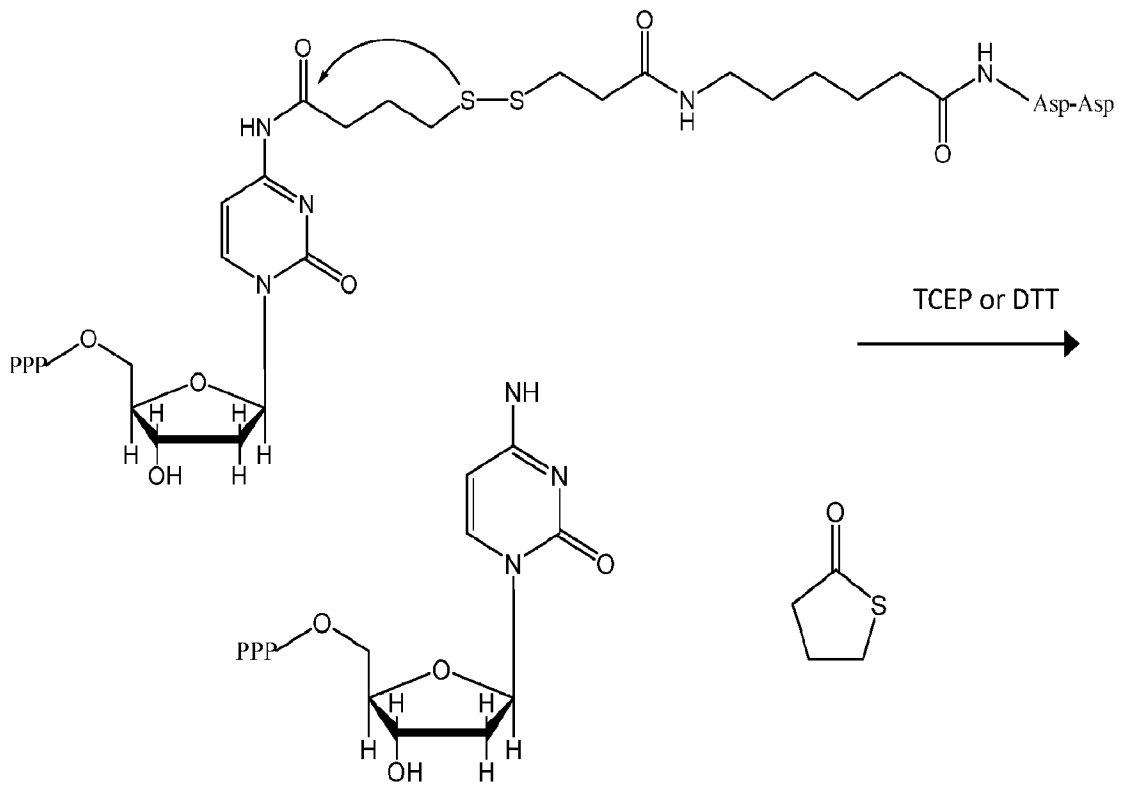
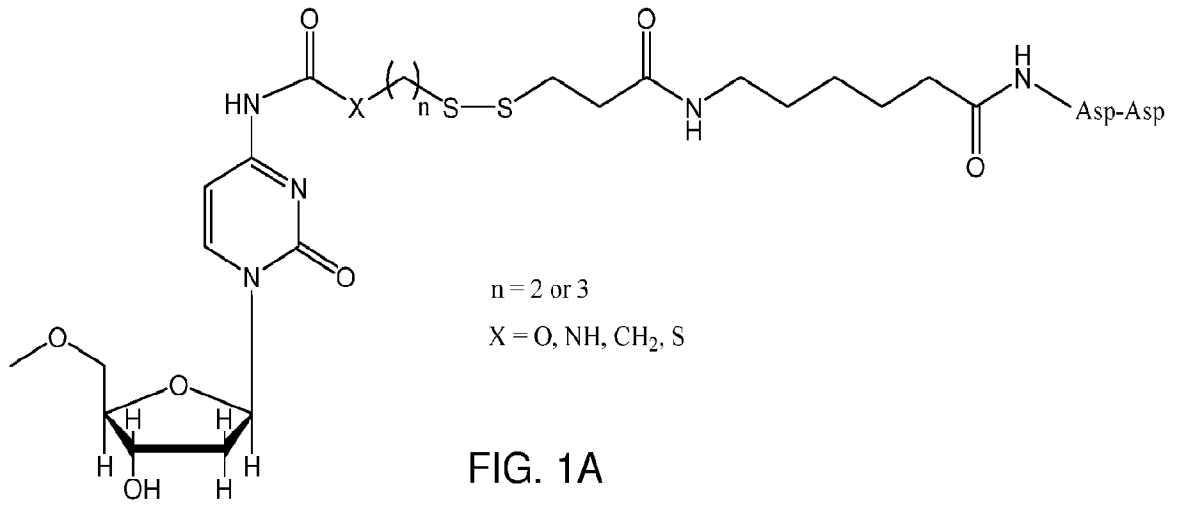
before contacting the nucleic acid initiator with a releasing agent, incorporating a second sequence-specific cleavage element on a 3' end of the first oligonucleotide to create a second nucleic acid initiator and exposing the second nucleic acid initiator, attached to the first oligonucleotide, to nucleotide analogs in the presence of a polymerase to create a second oligonucleotide,

wherein contacting the nucleic acid initiator and the second nucleic acid initiator with the releasing agent cleaves the sequence-specific cleavage element and the second sequence-specific cleavage element and release the first oligonucleotide and the second oligonucleotide from the nucleic acid initiator and the second nucleic acid initiator.

18. The method of claim 8, wherein the polymerase is a nucleotidyl transferase or a modified nucleotidyl transferase.

19. The method of claim 8, wherein the polymerase is terminal deoxynucleotidyl transferase (TdT) or a modified TdT.

20. The method of claim 8, wherein the nucleotide analogs have an unmodified 3'-OH and a cleavable terminating group, and wherein the cleavable terminating group blocks subsequent nucleotidyl transferase activity but results in a nucleotide substrate for nucleotidyl transferase upon cleavage of the terminating group.



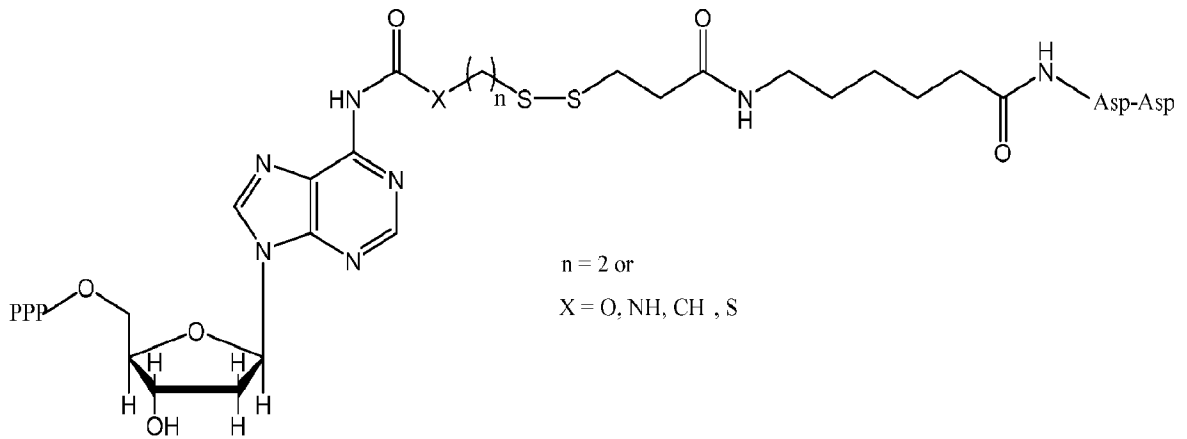


FIG. 2A

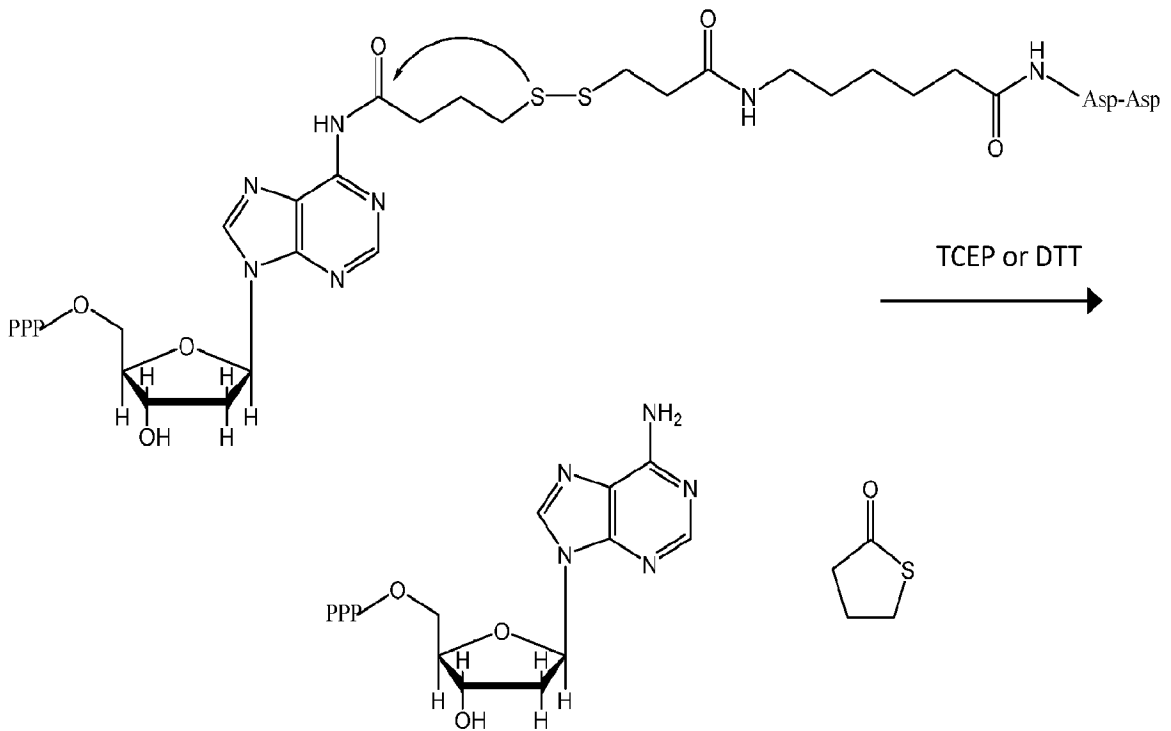


FIG. 2B

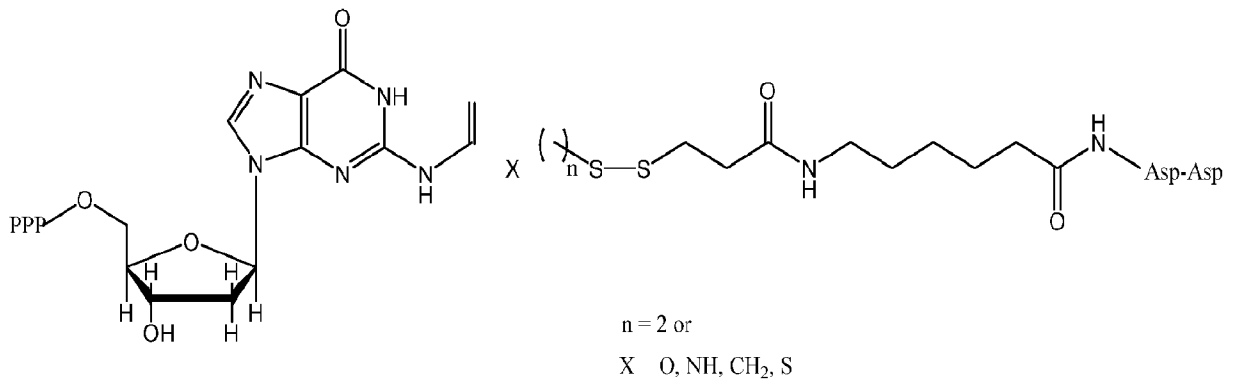


FIG. 3A

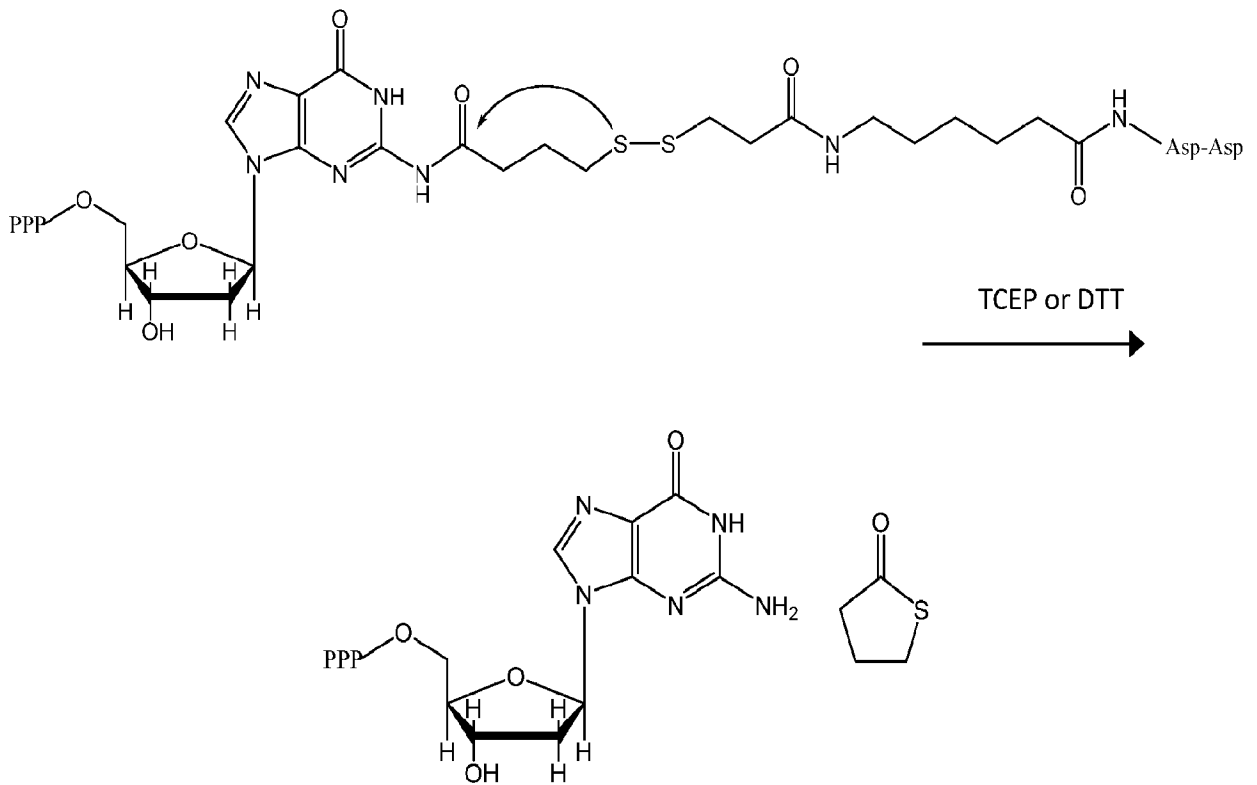


FIG. 3B

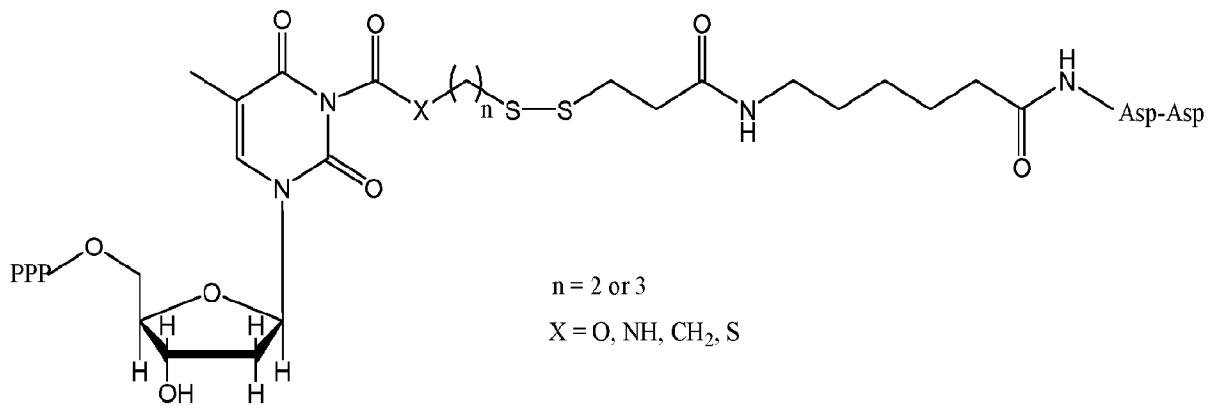


FIG. 4A

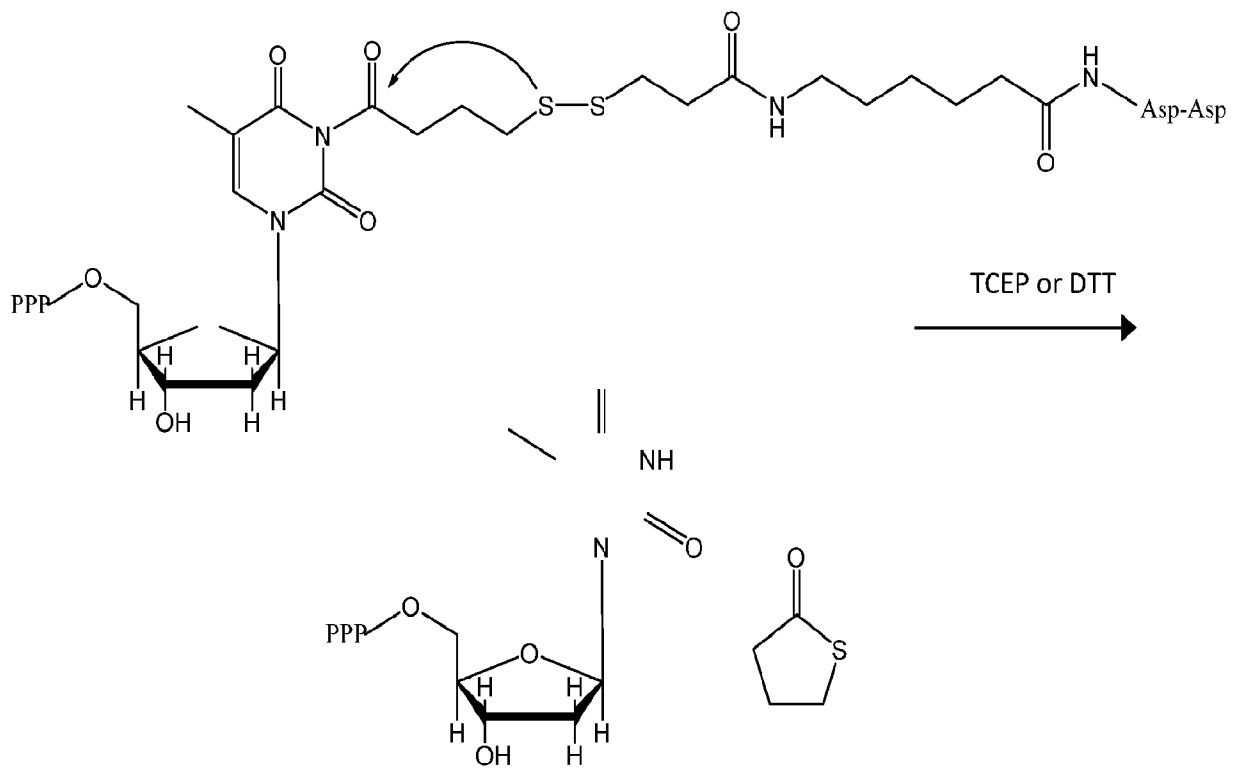


FIG. 4B

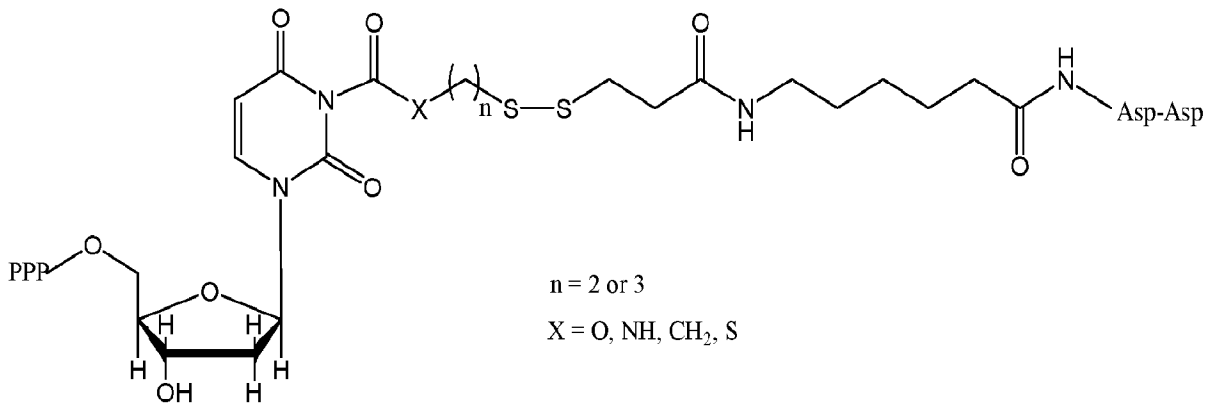


FIG. 5A

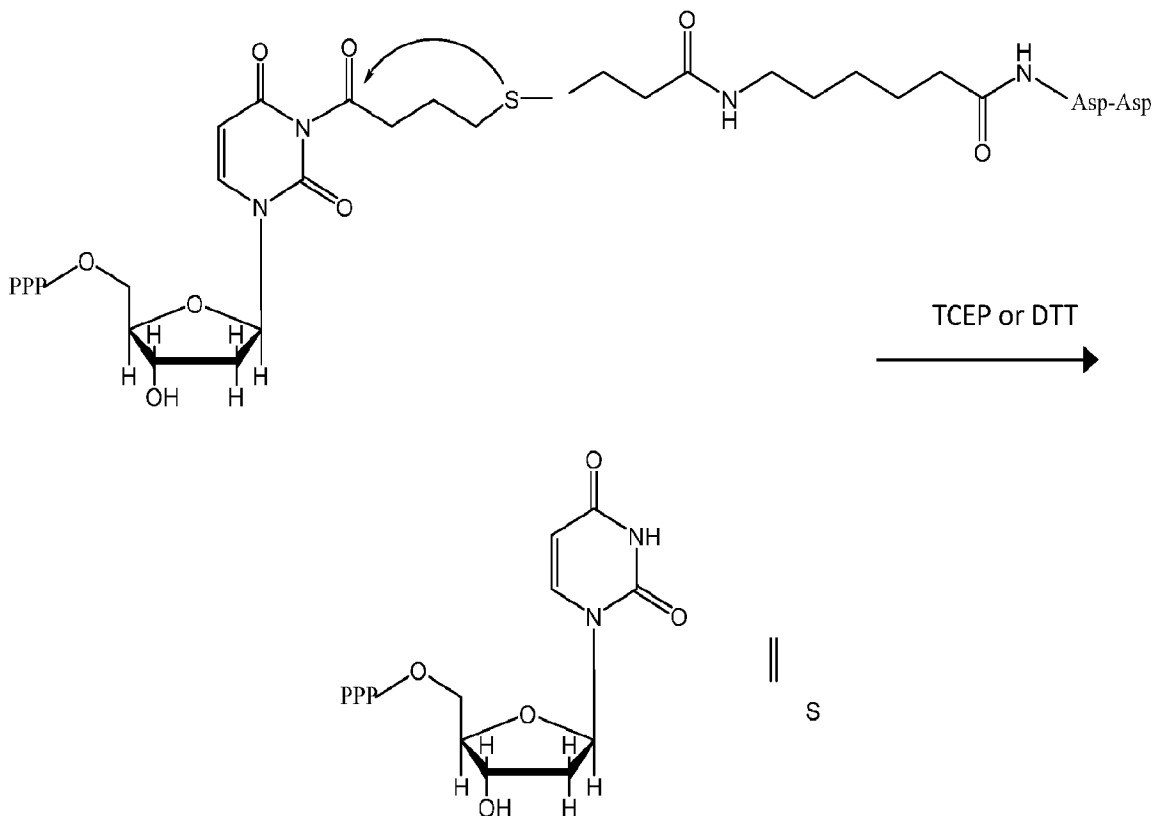
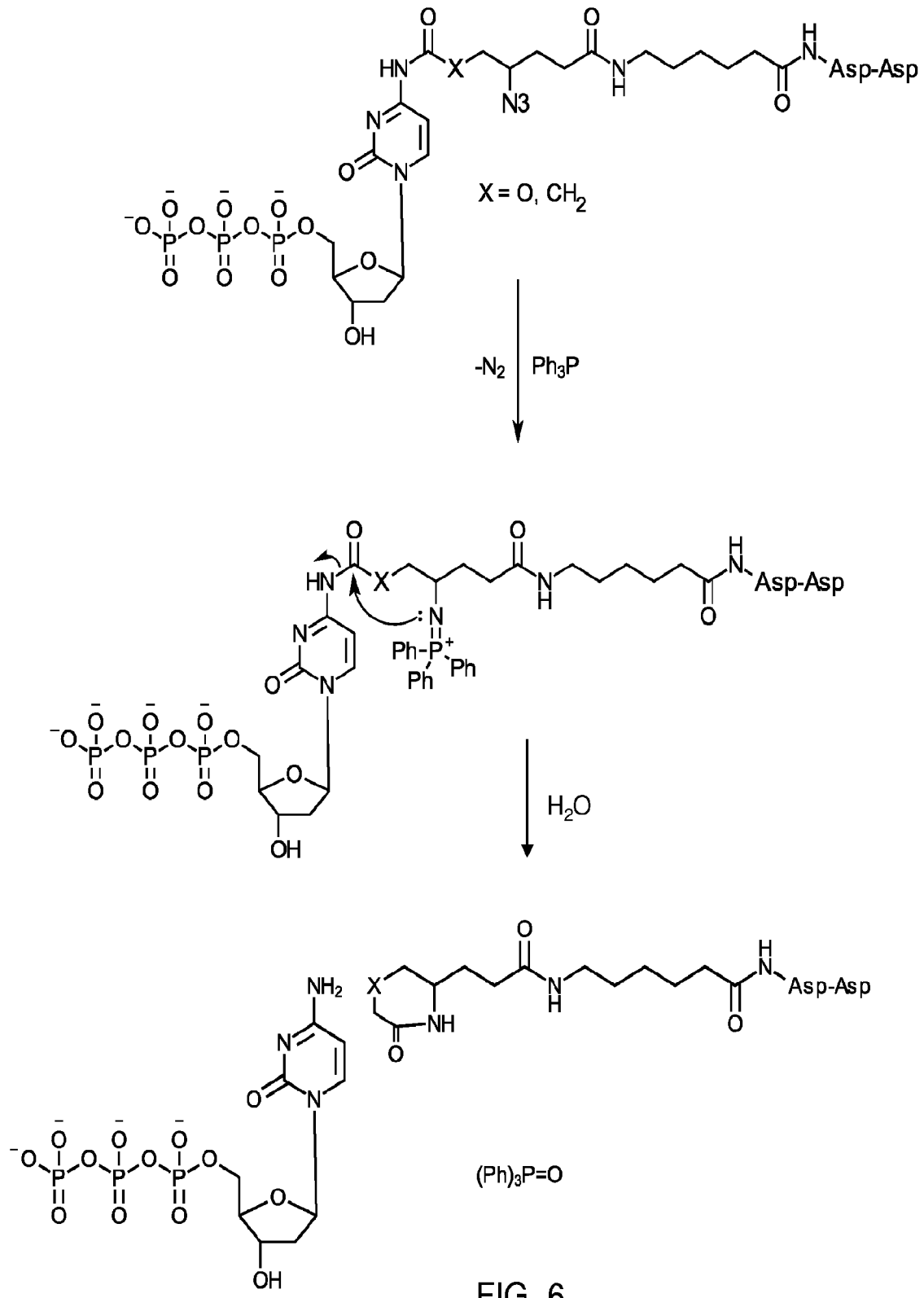


FIG. 5B



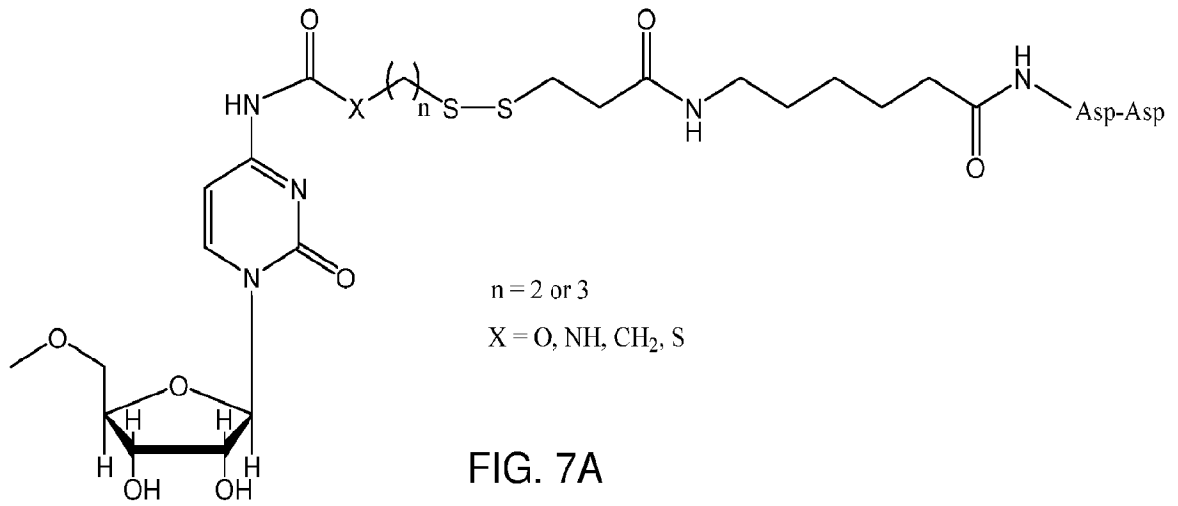


FIG. 7A

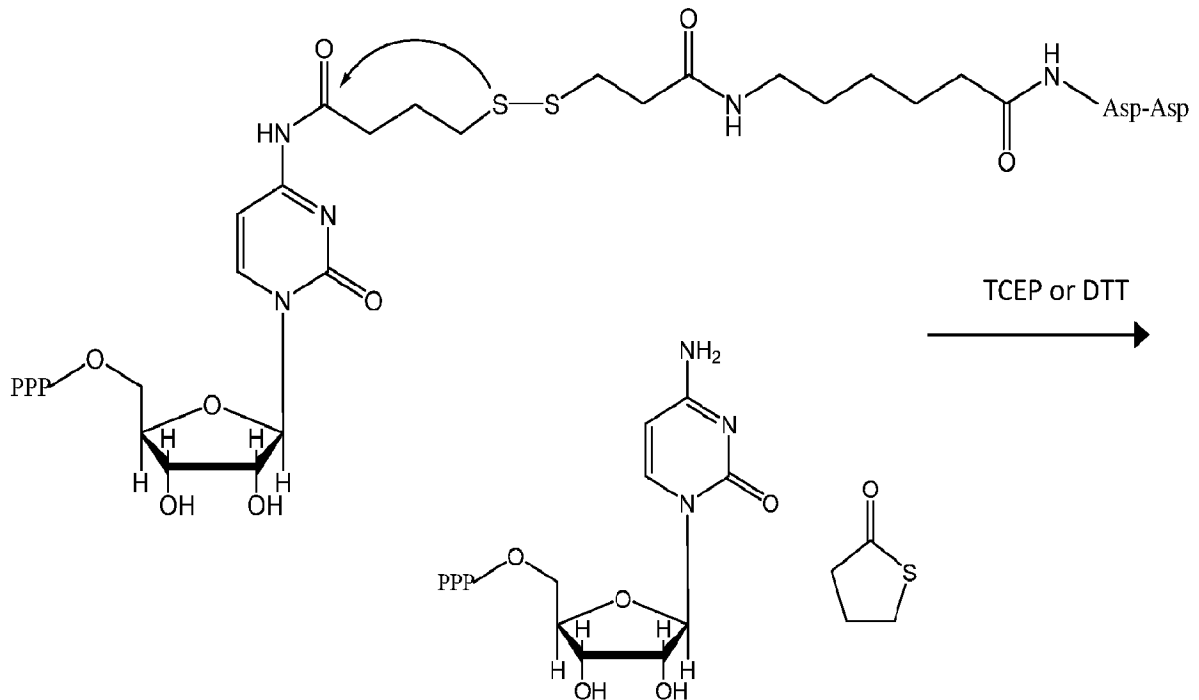


FIG. 7B

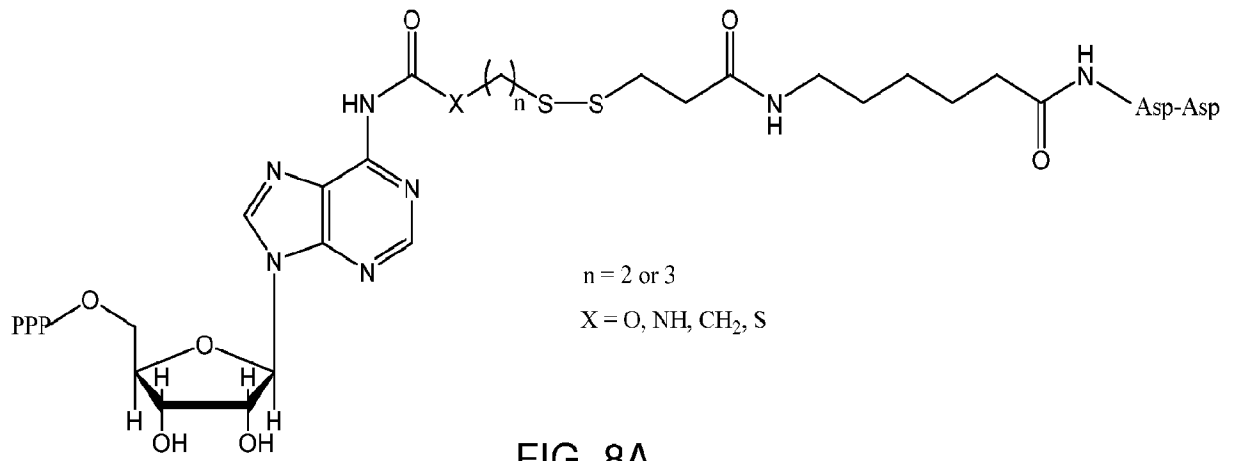


FIG. 8A

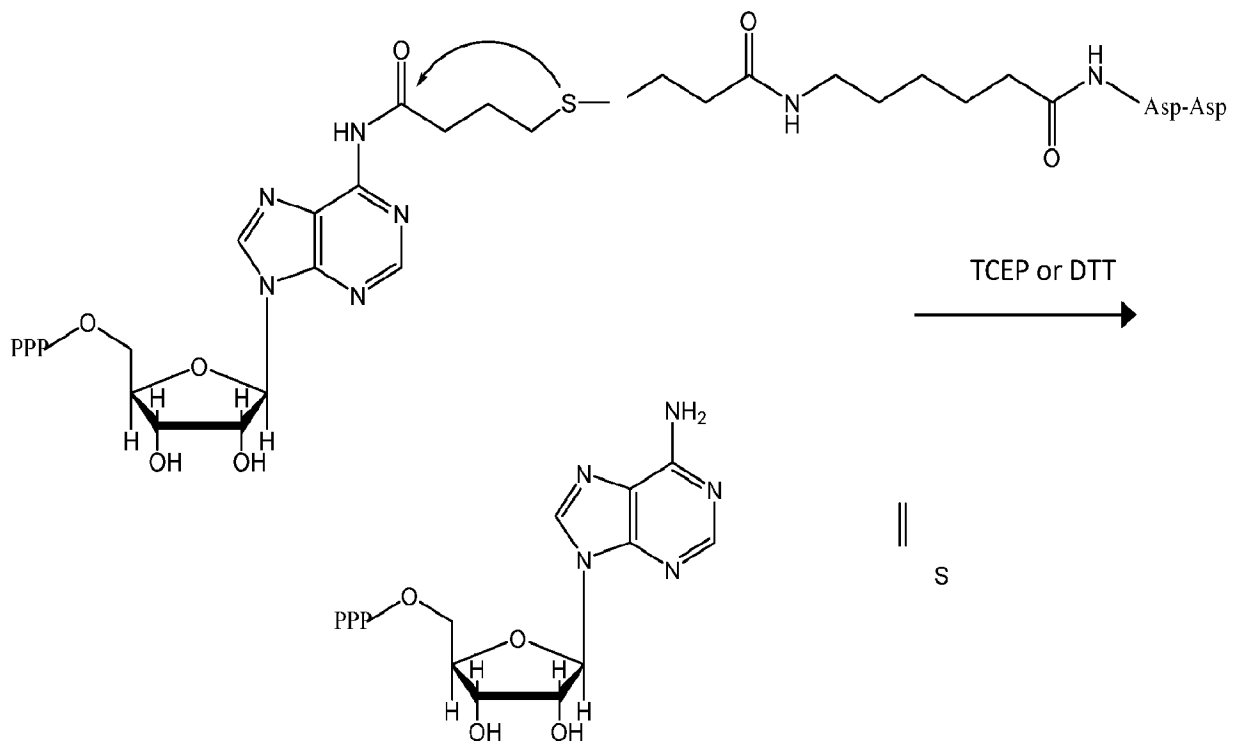


FIG. 8B

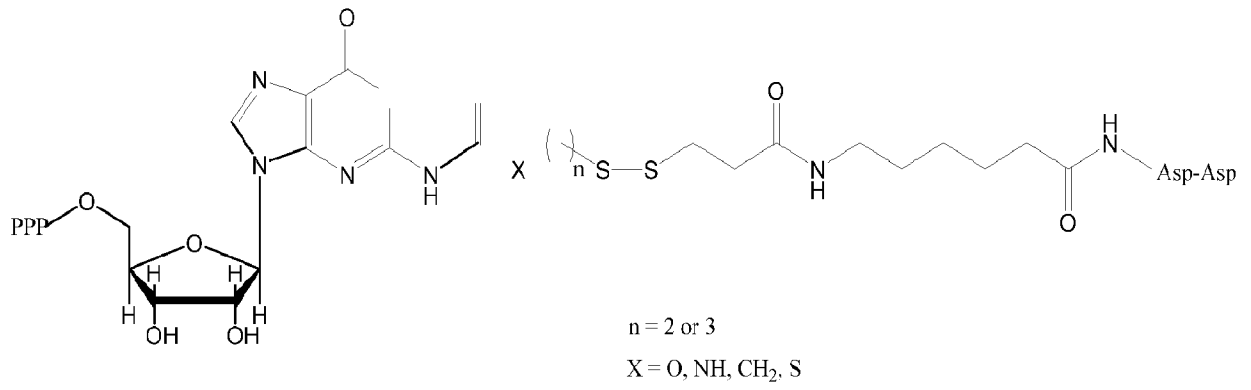


FIG. 9A

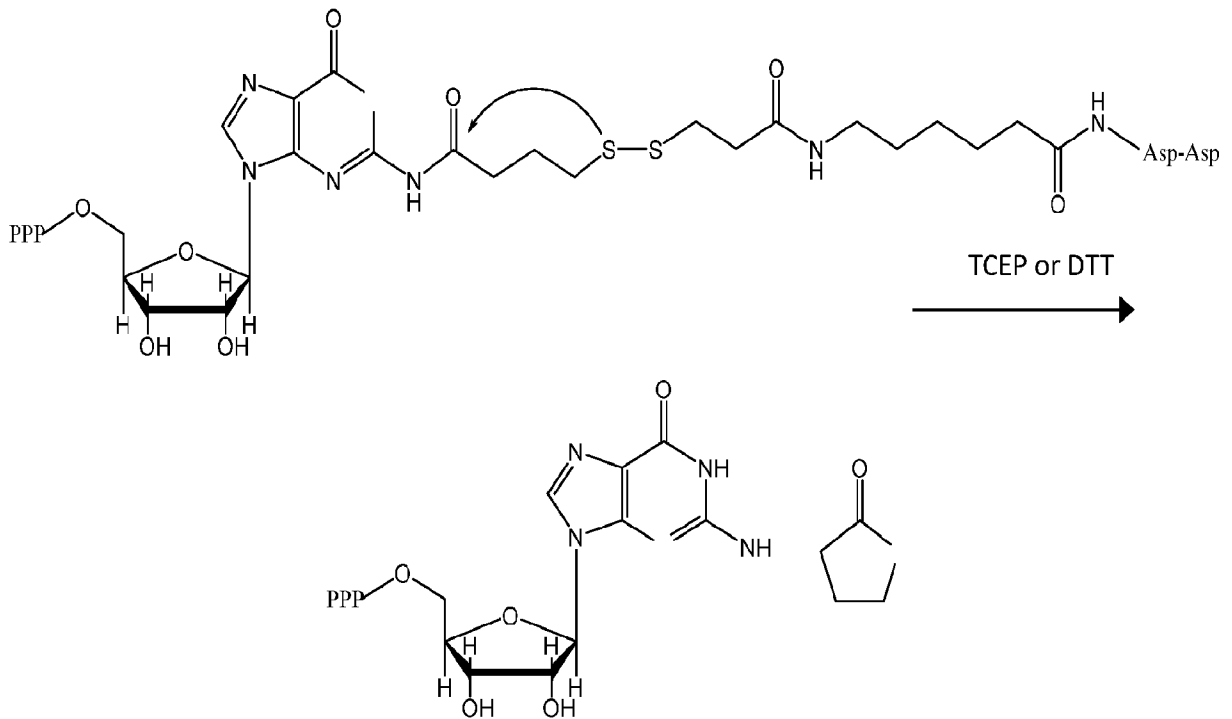
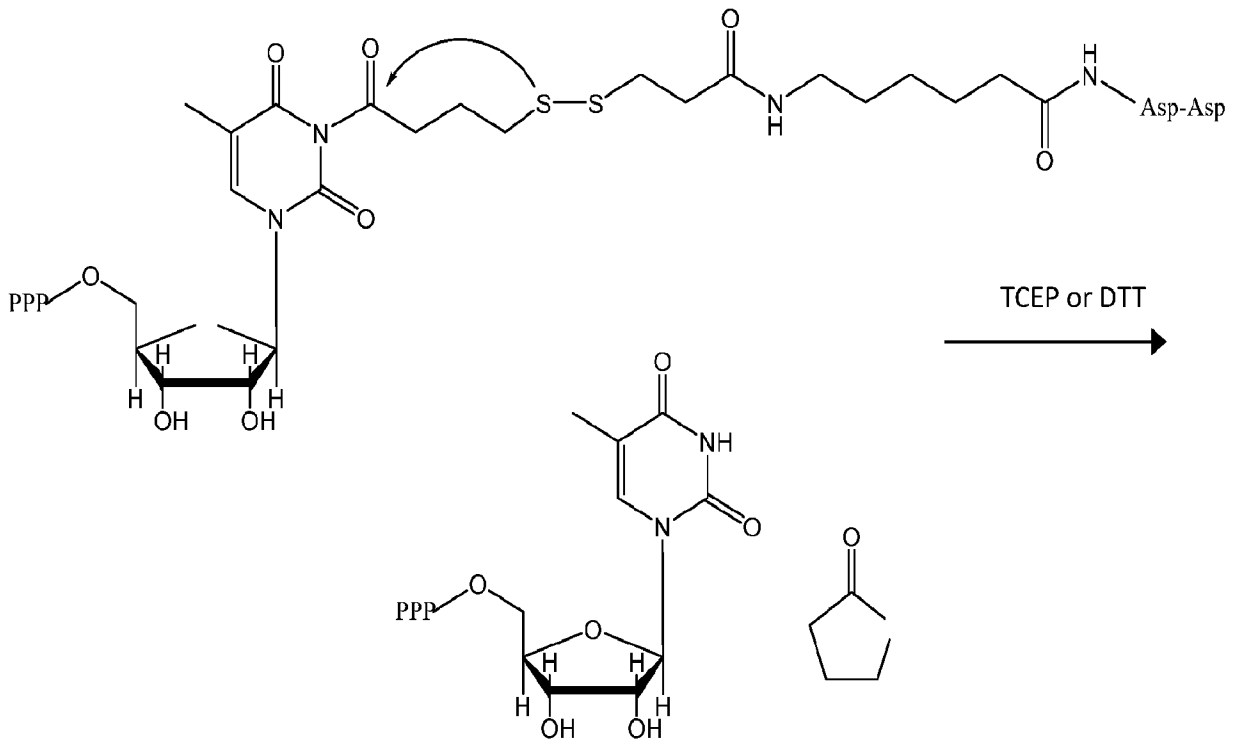
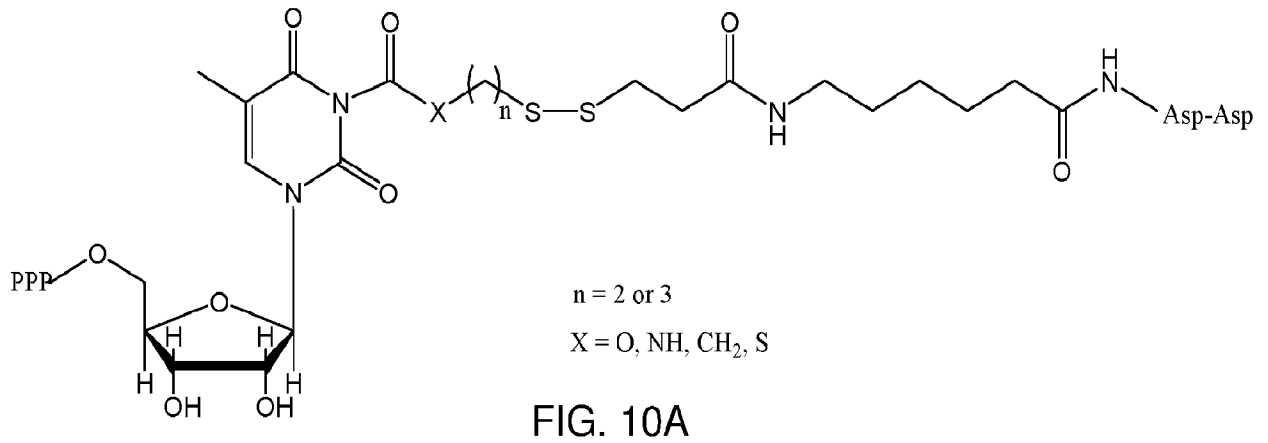


FIG. 9B



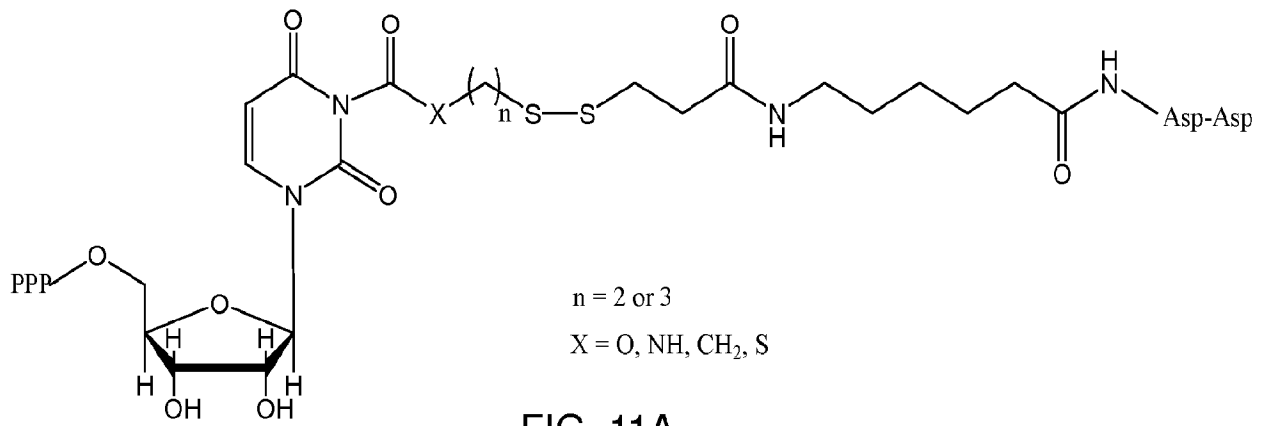


FIG. 11A

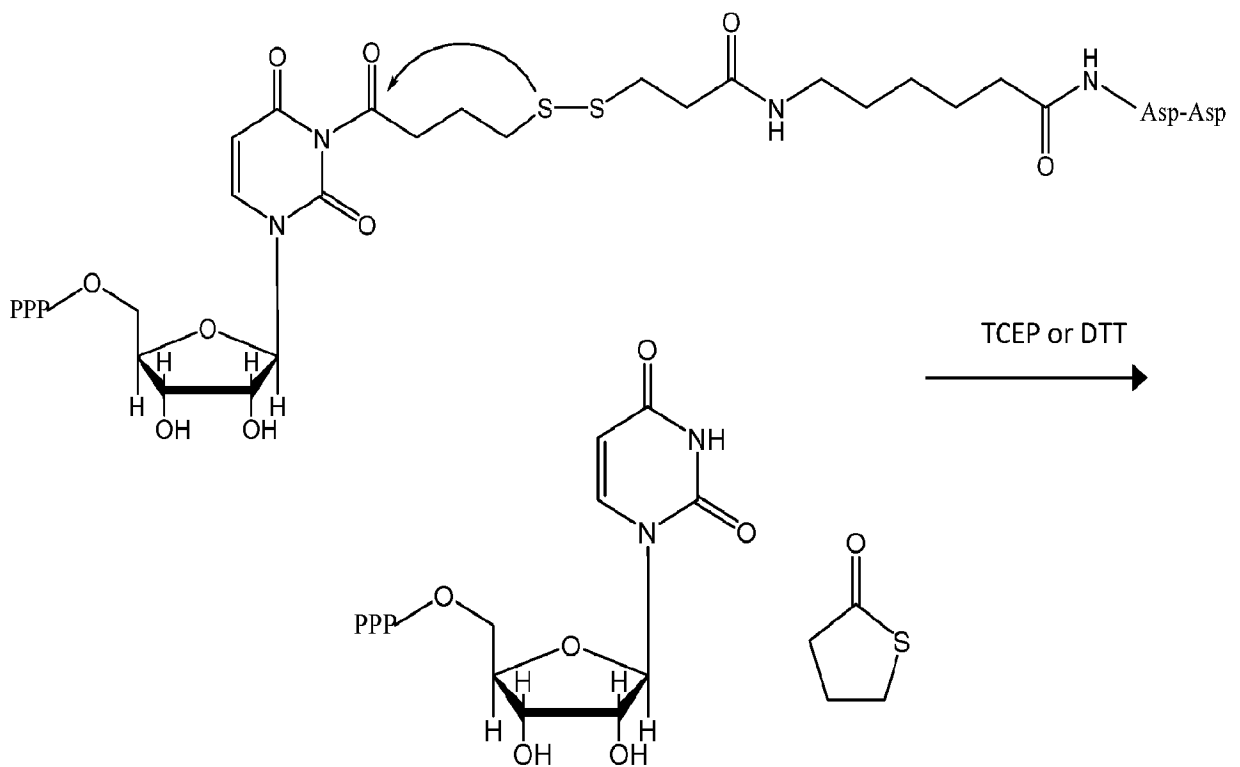


FIG. 11B

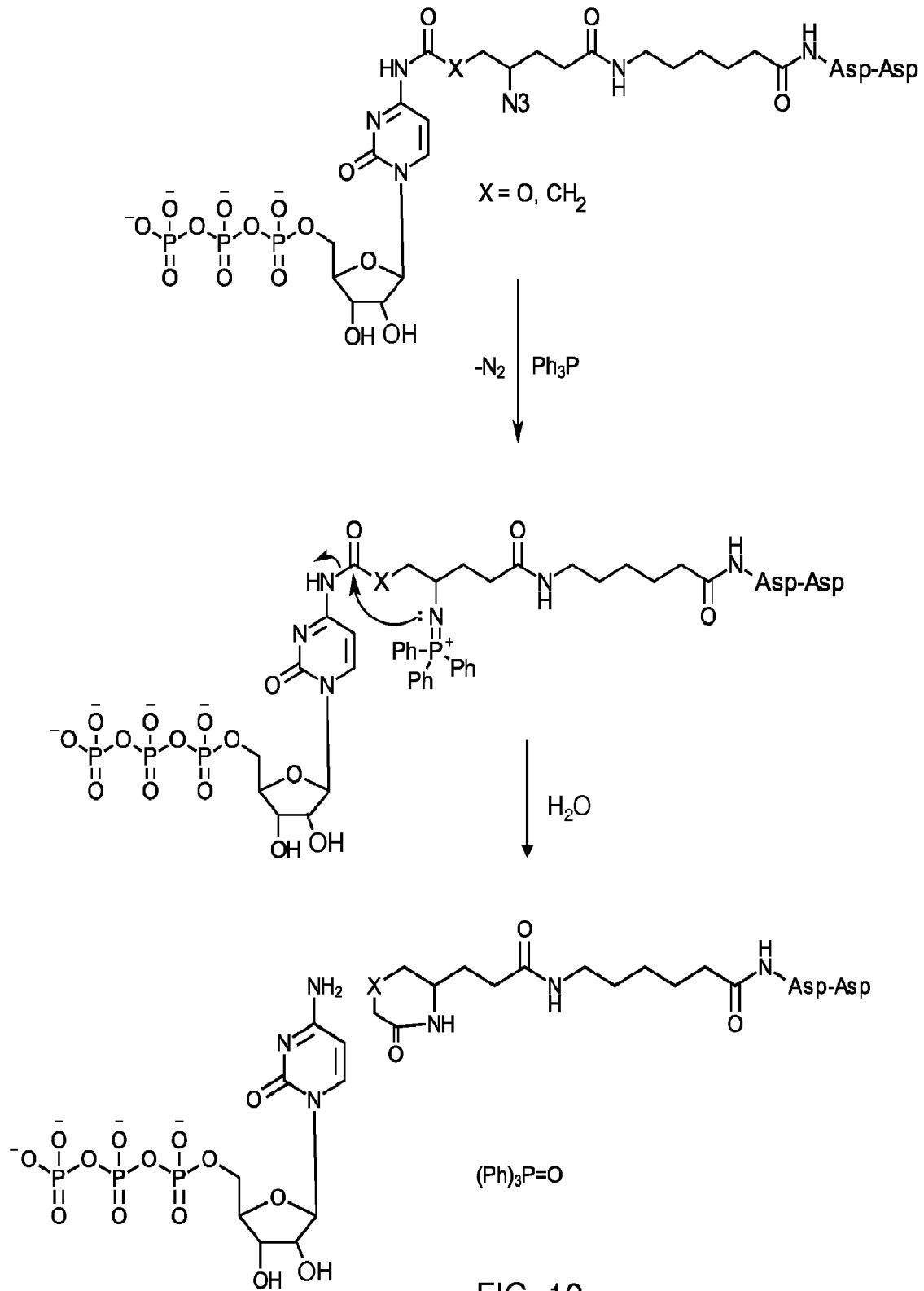


FIG. 12

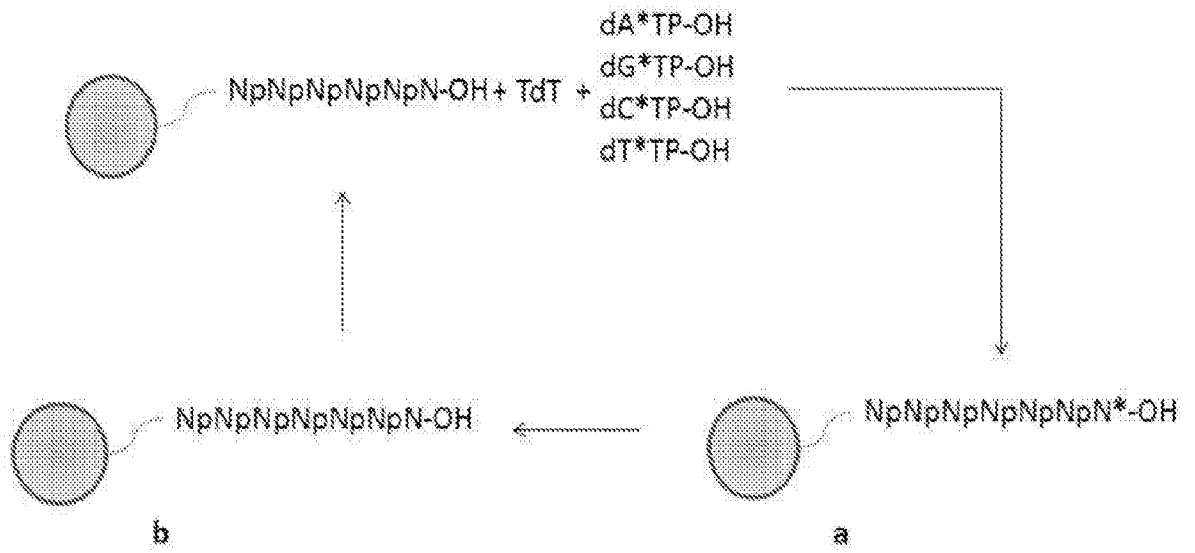


FIG. 13

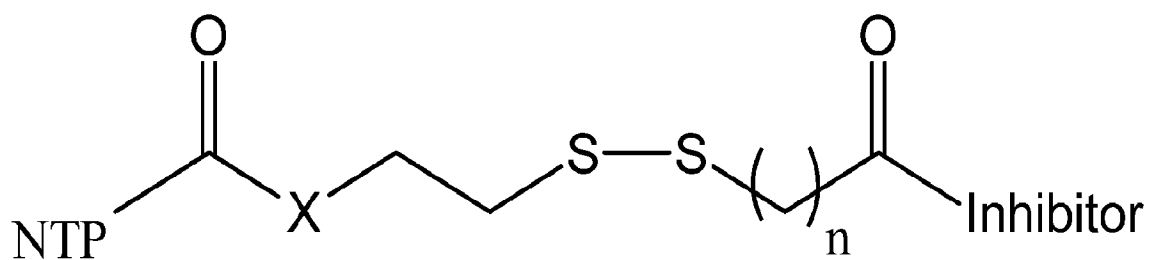
 $n = 1 - 4$ $X = O, S, NH, CH_2$

FIG. 14

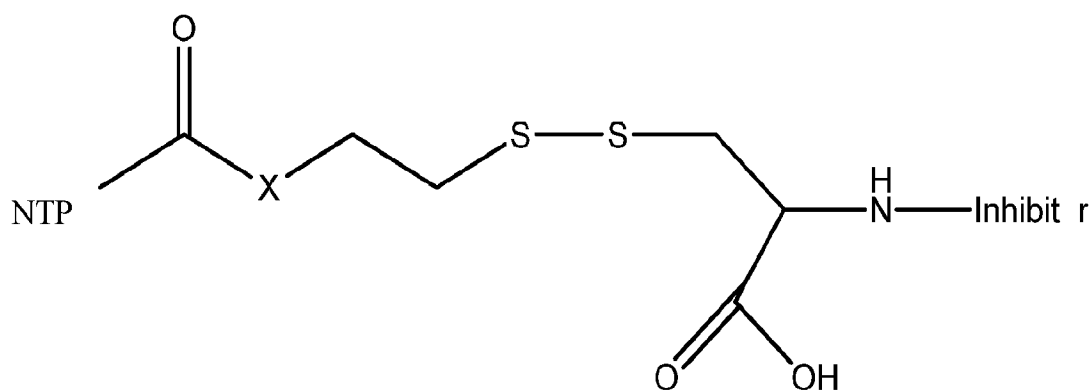
 $X = O, S, NH, CH_2$

FIG. 15

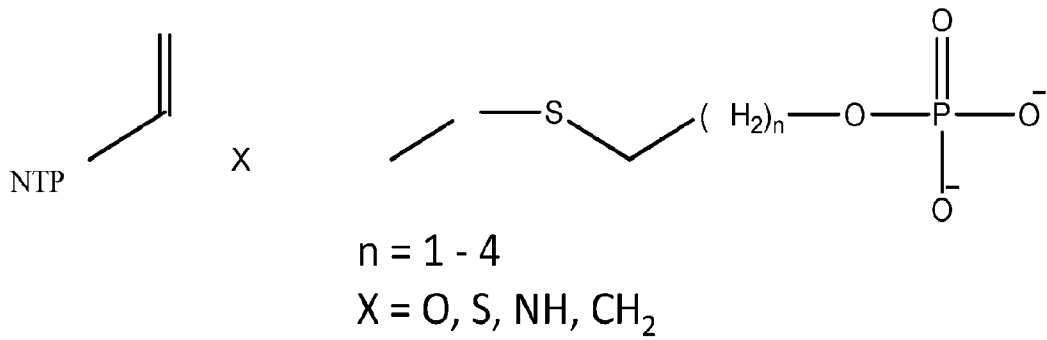


FIG. 16A

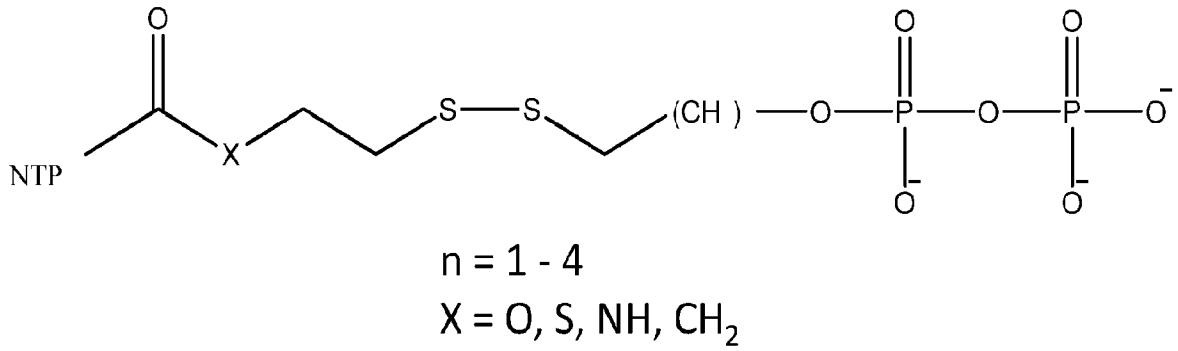


FIG. 16B

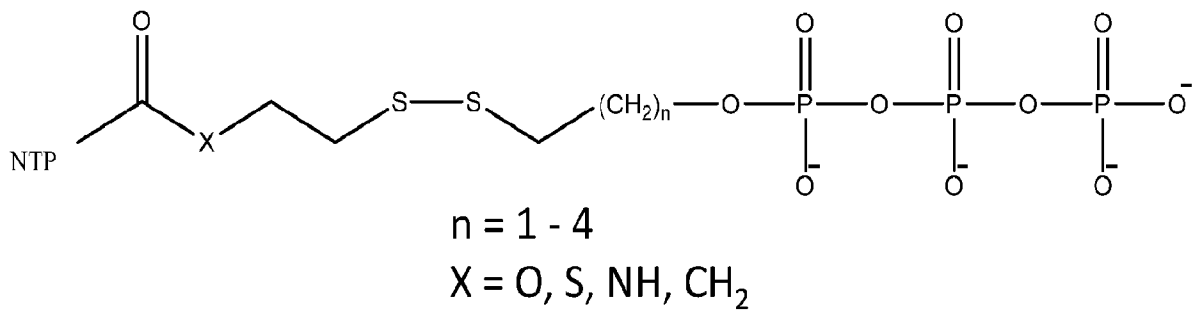


FIG. 16C

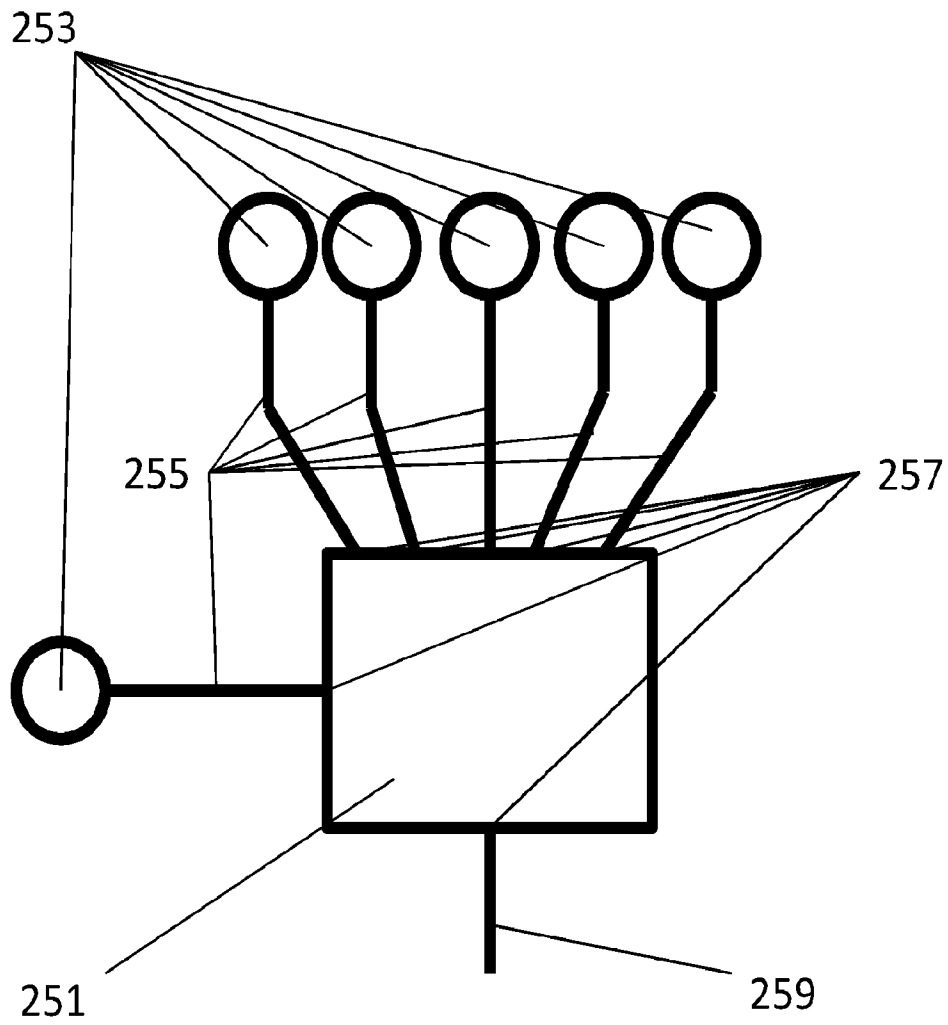


FIG. 17

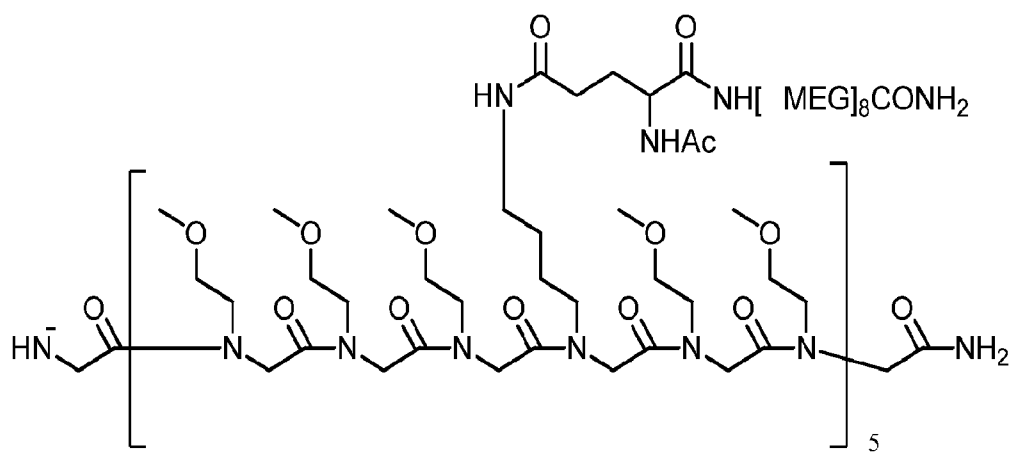


FIG. 18

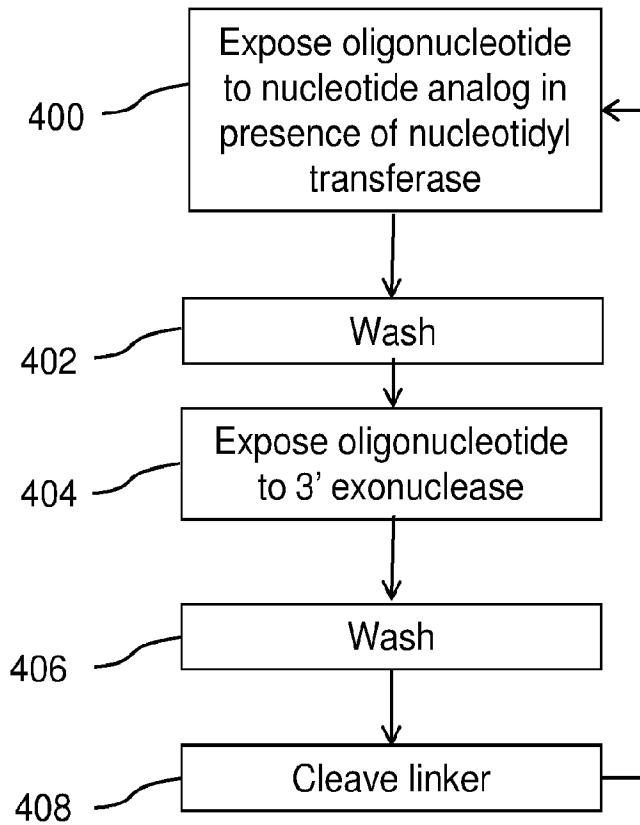


FIG. 19

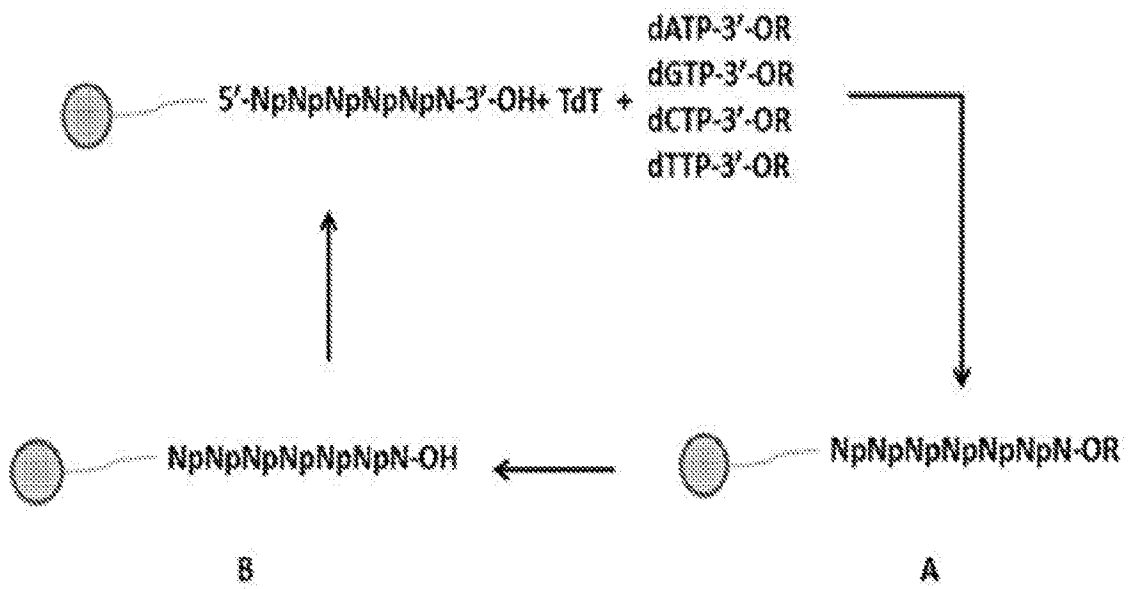


FIG. 20

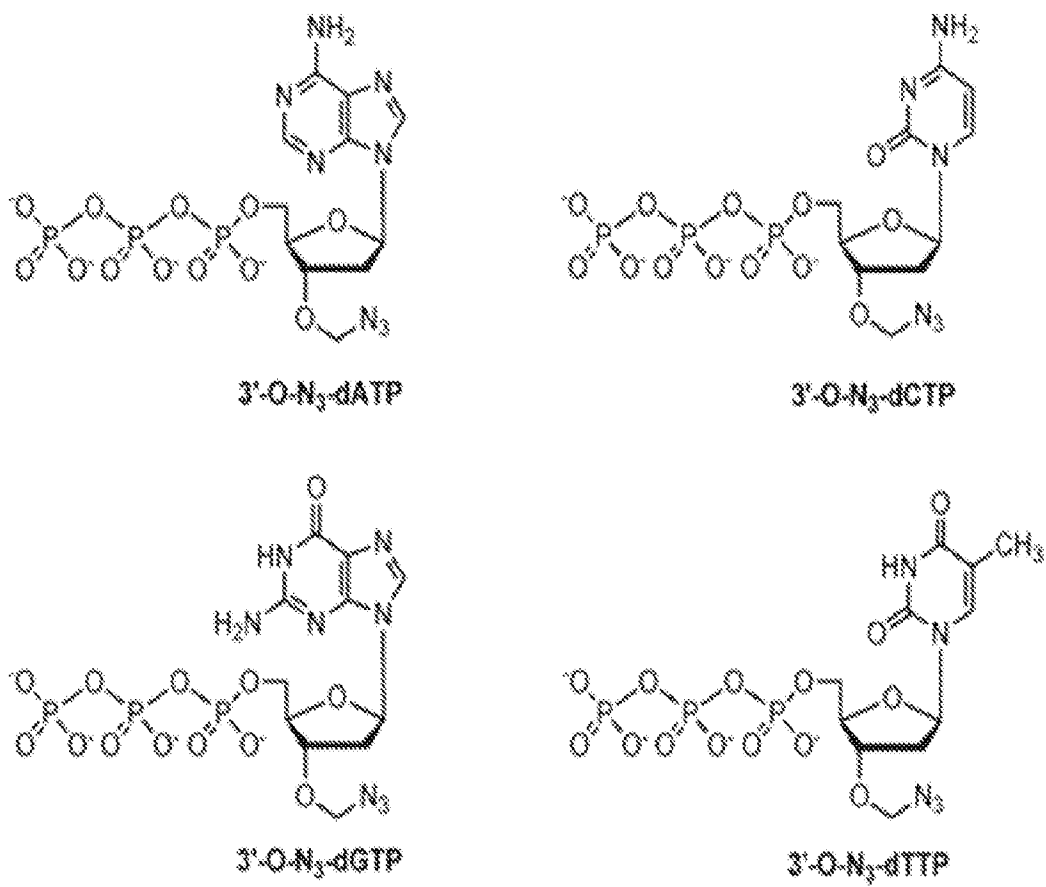


FIG. 21

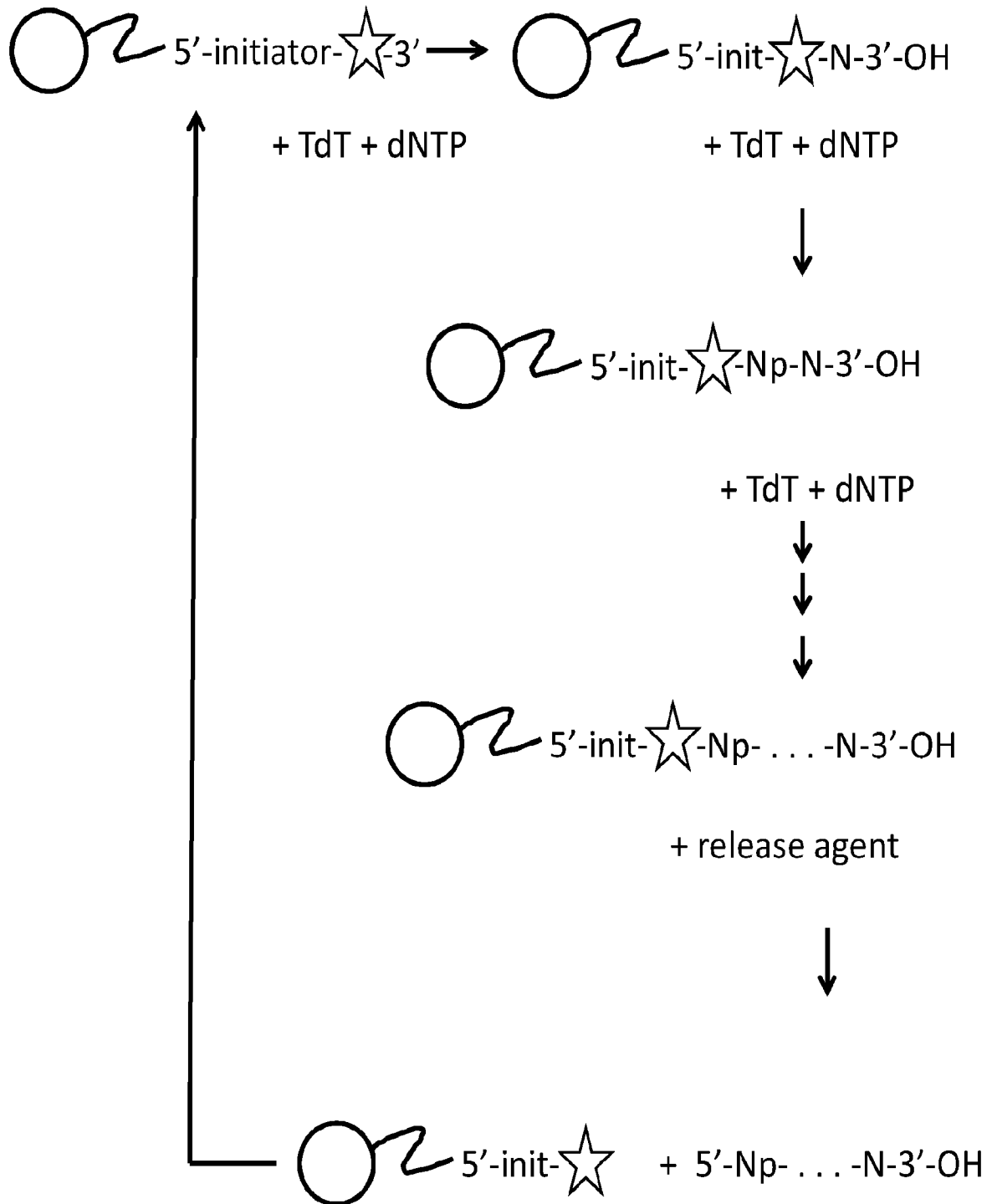


FIG. 22

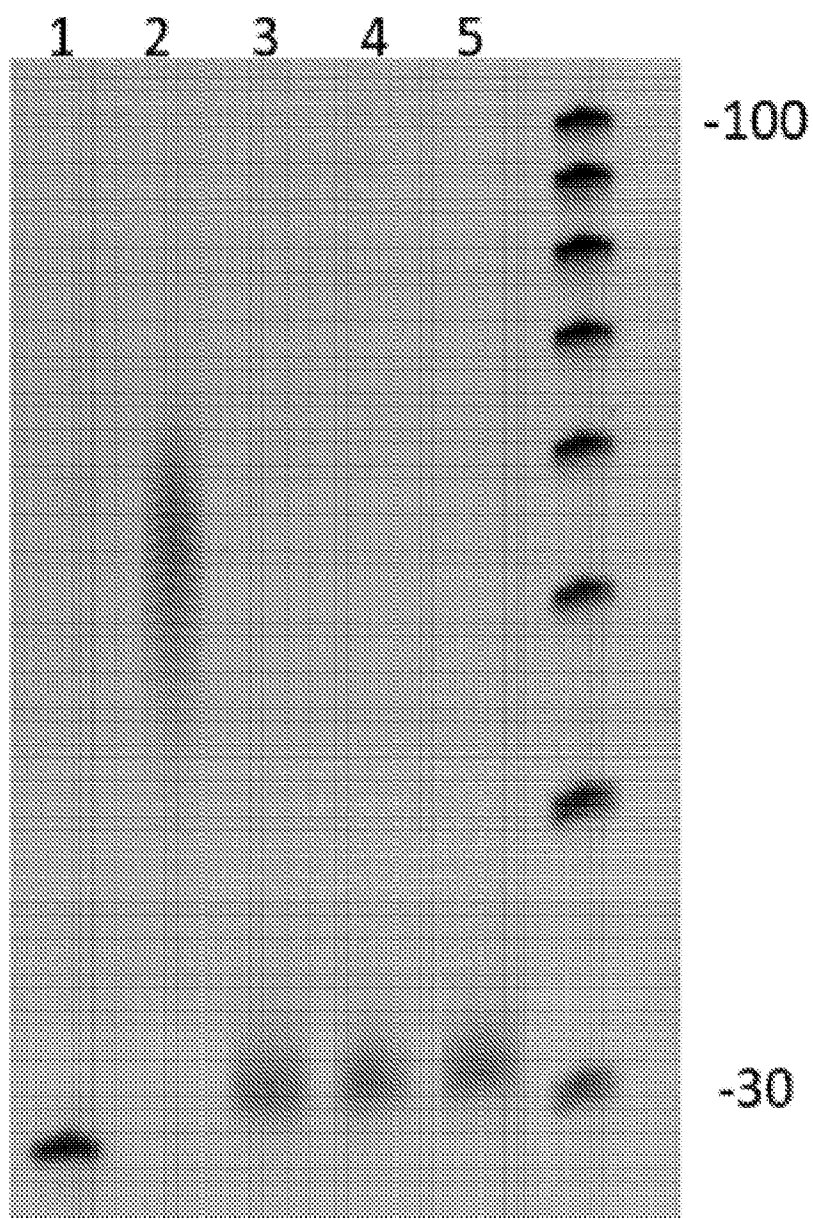


FIG. 23

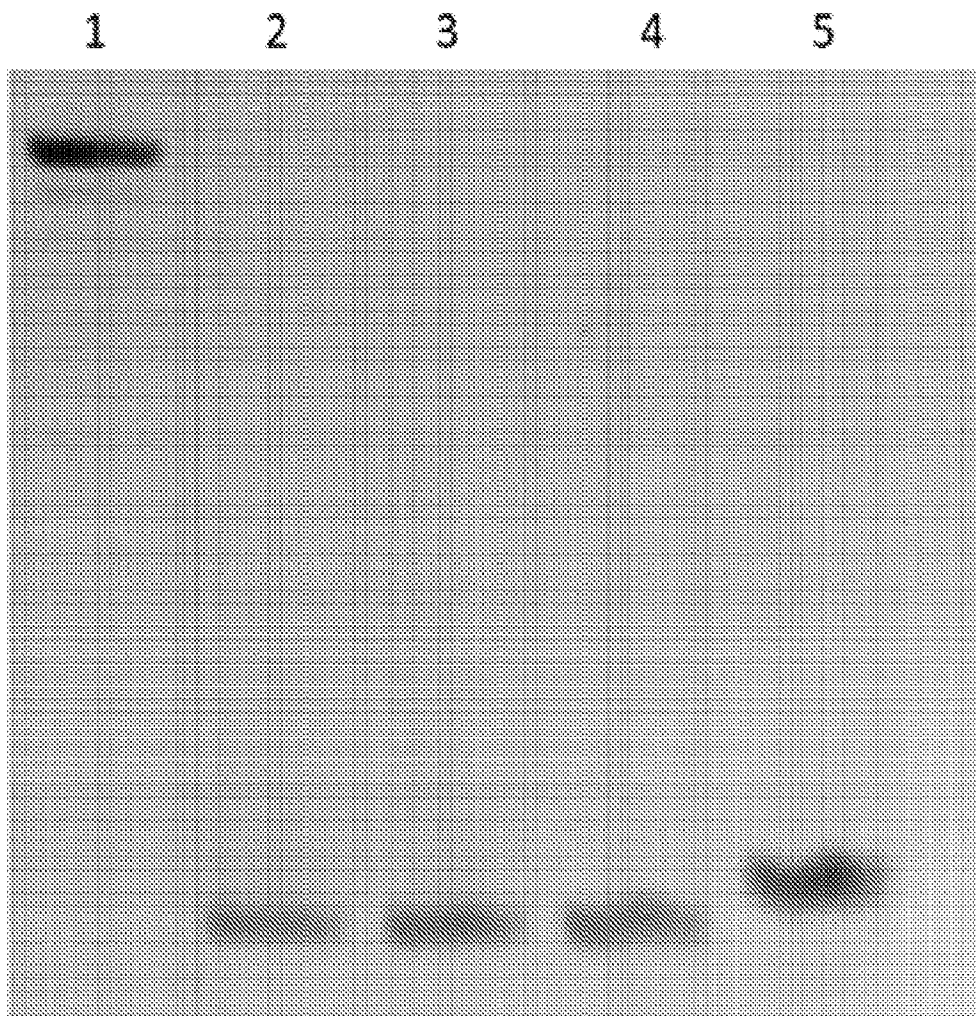


FIG. 24

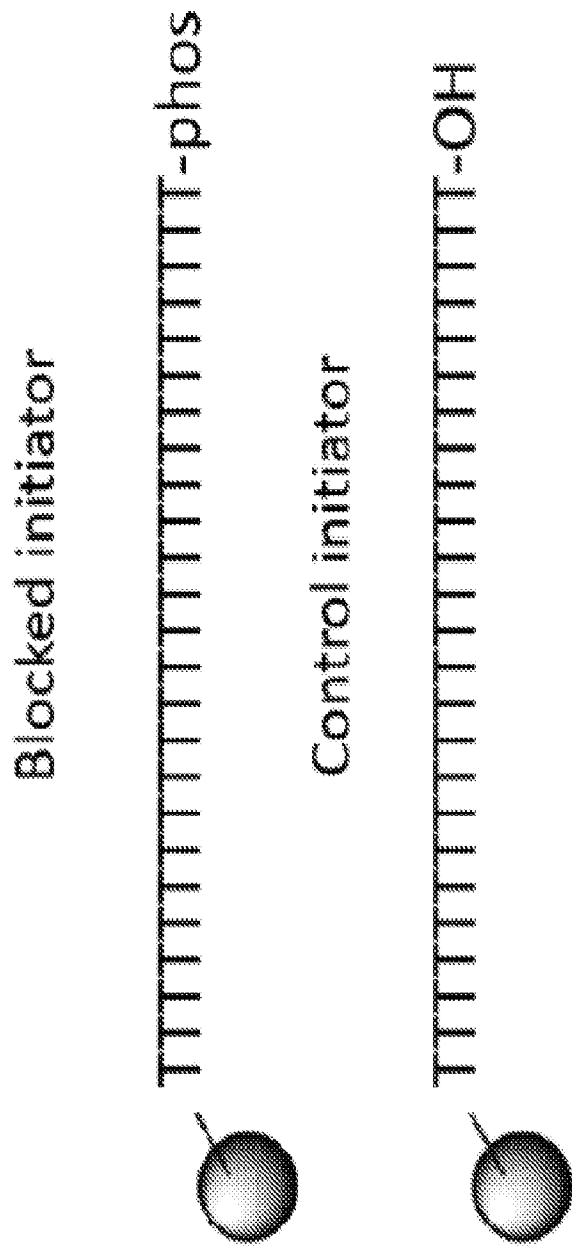
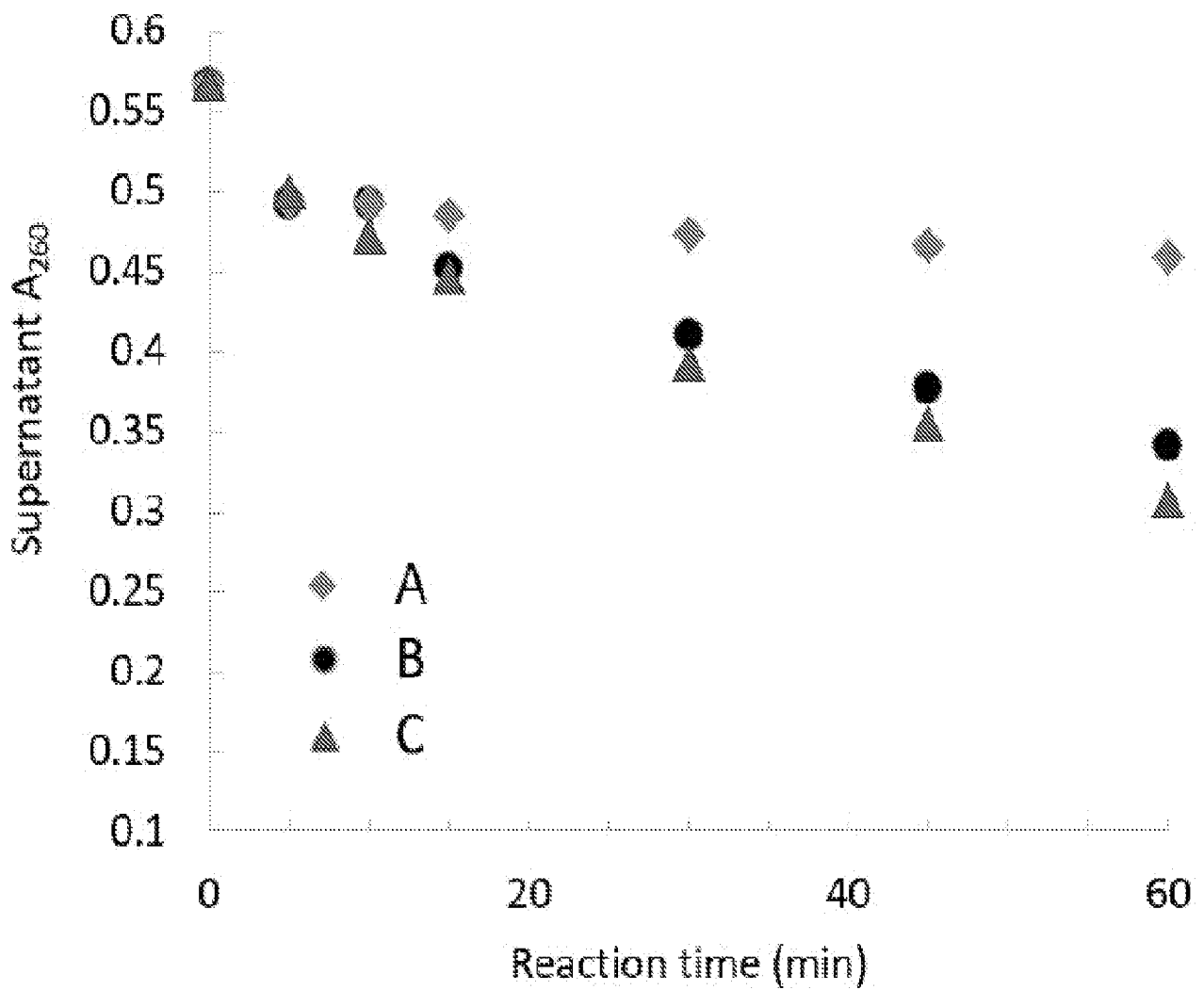


FIG. 25



- A) Blocked initiator after heat killing only
- B) Blocked initiator after alkaline phosphatase and heat killing
- C) Control initiator

FIG. 25 (Continued)

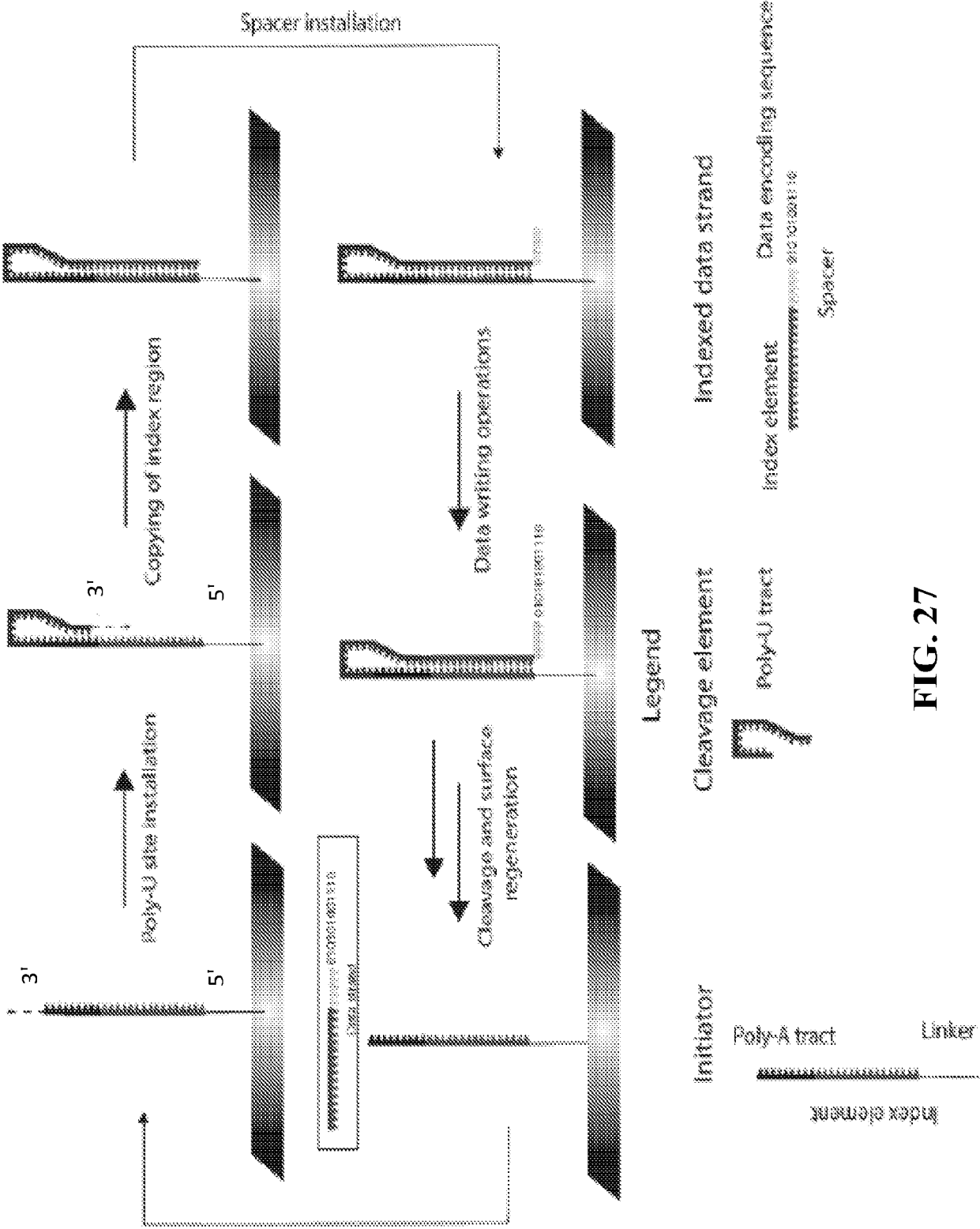


FIG. 27

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2020/015640**A. CLASSIFICATION OF SUBJECT MATTER****B01J 19/00(2006.01)i, C12P 19/34(2006.01)i, C12Q 1/6806(2018.01)i, C12Q 1/6811(2018.01)i, C12Q 1/6837(2018.01)i, C12N 9/12(2006.01)i**

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)

B01J 19/00; C07H 021/04; C07H 21/02; C07H 21/04; C12P 019/34; C12P 19/34; C12Q 001/68; C12Q 1/68; C12Q 1/6806; C12Q 1/6811; C12Q 1/6837; C12N 9/12

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

Korean utility models and applications for utility models
Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & Keywords: solid support, oligonucleotide, nucleic acid, nucleotide, polymerase

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004-0161749 A1 (HALL, J. G. et al.) 19 August 2004 paragraphs [0007], [0014], [0065], [0089]; claims 1, 7, 14	1-5,7
Y		6,8-20
Y	US 2010-0304368 A1 (CHERKASOV, D. et al.) 02 December 2010 paragraphs [0074], [0133]-[0138], [0141], [0184], [0192], [0378], [0610]; figures 15-17	6,8-20
Y	US 2004-0005614 A1 (KURN, N. et al.) 08 January 2004 paragraph [0008]	8-20
A	US 2016-0046974 A1 (MOLECULAR ASSEMBLIES, INC.) 18 February 2016 claims 1-48	1-20
A	US 2005-0260609 A1 (LAPIDUS, S. N.) 24 November 2005 claims 1-23	1-20
PX	US 2019-0275492 A1 (MOLECULAR ASSEMBLIES, INC.) 12 September 2019 claims 1-20 * This is an open-laid gazette of priority application of this PCT application.	1-20

 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

"D" document cited by the applicant in the international application

"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

25 May 2020 (25.05.2020)

Date of mailing of the international search report

25 May 2020 (25.05.2020)

Name and mailing address of the ISA/KR

International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

MIN, In Gyou

Telephone No. +82-42-481-3326



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2020/015640

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004-0161749 A1	19/08/2004	US 2008-0015349 A1 US 2008-0293930 A1 US 7582436 B2	17/01/2008 27/11/2008 01/09/2009
US 2010-0304368 A1	02/12/2010	DE 102007045173 A1 EP 2074133 A2 WO 2008-043426 A2 WO 2008-043426 A3	15/05/2008 01/07/2009 17/04/2008 28/05/2009
US 2004-0005614 A1	08/01/2004	AU 2003-279697 A1 CA 2486283 A1 EP 1573056 A2 EP 1573056 A4 JP 2005-534304 A JP 2010-011872 A JP 4551216 B2 WO 2004-011665 A2 WO 2004-011665 A3 ZA 200409152 B	16/02/2004 05/02/2004 14/09/2005 28/11/2007 17/11/2005 21/01/2010 22/09/2010 05/02/2004 21/07/2005 28/11/2007
US 2016-0046974 A1	18/02/2016	AU 2014-247928 A1 AU 2014-247928 B2 BR 112015025264 A2 CA 2908690 A1 CA 2958581 A1 CN 105264085 A CN 107109452 A DK 2796552 T3 EP 2796552 A2 EP 2796552 A3 EP 2796552 B1 EP 3115462 A1 EP 3115462 B1 EP 3183356 A2 EP 3183365 A1 EP 3425053 A1 JP 2016-515391 A JP 2017-525391 A JP 2019-054810 A JP 6448097 B2 US 10041110 B2 US 2014-0363851 A1 US 2014-0363852 A1 US 2016-0046973 A1 US 2016-0168611 A1 US 2018-0305746 A1 US 8808989 B1 US 9279149 B2 US 9695470 B2	29/10/2015 24/08/2017 18/07/2017 09/10/2014 25/02/2016 20/01/2016 29/08/2017 05/09/2016 29/10/2014 04/03/2015 13/07/2016 11/01/2017 25/07/2018 28/06/2017 28/06/2017 09/01/2019 30/05/2016 07/09/2017 11/04/2019 09/01/2019 07/08/2018 11/12/2014 11/12/2014 18/02/2016 16/06/2016 25/10/2018 19/08/2014 08/03/2016 04/07/2017

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2020/015640

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		US 9771613 B2	26/09/2017
		WO 2014-162307 A2	09/10/2014
		WO 2014-162307 A3	23/04/2015
		WO 2014-165864 A2	09/10/2014
		WO 2014-165864 A3	20/11/2014
		WO 2016-028802 A1	25/02/2016
		WO 2016-028803 A2	25/02/2016
		WO 2016-028803 A3	14/04/2016
US 2005-0260609 A1	24/11/2005	CA 2567822 A1	08/12/2005
		EP 1766064 A1	28/03/2007
		JP 2008-500047 A	10/01/2008
		US 2008-0287306 A1	20/11/2008
		WO 2005-116262 A1	08/12/2005
US 2019-0275492 A1	12/09/2019	None	