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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 January 2009 (29.01.2009)

PCT

(10) International Publication Number
WO 2009/014612 A2

(51) International Patent Classification:
C07H 21/00 (2006.01) *C12P 19/30* (2006.01)

(21) International Application Number:
PCT/US2008/008613

(22) International Filing Date: 15 July 2008 (15.07.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/781,160 20 July 2007 (20.07.2007) US

(71) Applicants *(for all designated States except US)*: **VISIGEN BIOTECHNOLOGIES, INC.** [US/US]; 2575 West Bellfort, Suite 250, Houston, TX 77054 (US). **APPLERA CORPORATION** [US/US]; 850 Lincoln Dr., Foster City, CA 94404 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **WANG, Hongyi** [CN/US]; 5850 Gulfton, #1837, Houston, TX 77081 (US). **GAO, Xiaolian** [US/US]; 2212B Bellefontaine, Houston, TX 77030 (US). **YU, Peilin** [US/US]; 6245 Renwick #4210, Houston, TX 77081 (US). **REDDY, Mitsu S.** [IN/US]; 2203 Lake Wind Dr., Pearland, TX 77584 (US). **HARDIN, Susan H.** [US/US]; 4712 Nantucket Dr., College Station, TX 77845 (US). **LINCECUM, Tommie, Jr.** [US/US]; 6525 Rolla Street, Houston, Texas 77055 (US). **WILLIAMS, Amy** [US/US]; 5018 Stanhope Dr., Houston, TX 77084 (US). **DELUGE, Norha** [US/US]; 7447 Cambridge Street #123, Houston, TX 77054 (US). **BELOSLUDTSEV, Yuri** [US/US]; 47 Wood Scent Ct., The Woodlands, TX 77380 (US). **MENCHEN, Steven M.** [US/US]; 768 Vanda Way, Fremont, CA 94536 (US). **LAM, Joe Y. L.** [US/US]; 7867 Pineville Circle, Castro Valley, CA 94552 (US). **CHEN, Jer-Kang** [US/US]; 882 Ames Ave., Palo Alto, CA 94303 (US).

(74) Agent: **DOSHI, Nishita**; Invitrogen c/o Intellevate, P.O.Box 52050, Minneapolis, Minnesota 55402 (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, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

[Continued on next page]

(54) Title: MODIFIED NUCLEOTIDES, METHODS FOR MAKING AND USING SAME

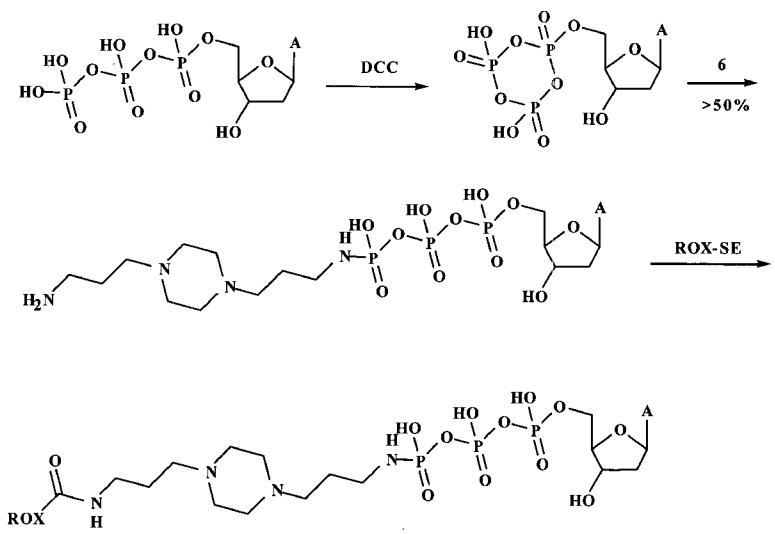


FIG. 10

(57) Abstract: Modified nucleotides are disclosed for use in single molecule sequencing, methods for making the modified nucleotides and method for using the modified nucleotides. Linkers for making the modified nucleotide are also disclosed.

WO 2009/014612 A2

GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

TITLE: **MODIFIED NUCLEOTIDES, METHODS FOR MAKING AND USING SAME**

INVENTOR: **Hongyi Wang, Xiaolian Gao, Peilin Yu, Mitsu Reddy, Susan H. Hardin, Tommie Lincecum, Jr., Amy Williams, Norha Deluge, Yuri Belosludtsev, Steven M. Menchen, Joe Y. L. Lam and Jer-Kang Chen**

ASSIGNEES: **VISIGEN BIOTECHNOLOGIES, INC.** and **APPLERA CORPORATION**

RELATED APPLICATIONS

[0001] This application claims priority to United States Patent Application No. 11/781,160 filed 07/20/2007 (20 July 2007).

BACKGROUND OF THE INVENTION

1. **Field of the Invention**

[0002] The present invention relates to modified nucleotides and methods for making and using same.

[0003] More particularly, the present invention relates to modified nucleotides including a natural or synthetic nucleotide having bonded to at least one site of the nucleotide a linker. The invention also relates to a modified nucleotide including a natural or synthetic nucleotide having bonded to at least one site a linker including at least one detectable group or moiety bonded to at one site of the linker. The invention also relates to method for making and using same.

2. **Description of the Related Art**

[0004] As single molecule sequencing advance ever closing to the ultimate goal of obtaining sequencing information from one or a large number of single molecule active sequencing sites in a field of view of a real time or near real time detection system, the need for modified nucleotide capable of detecting in such systems advance as well.

[0005] Although many modified nucleotides have been devised, there is a need in the art for modified nucleotides having a detectable group bonded thereto for use in such single molecule sequencing systems.

SUMMARY OF THE INVENTION

General Structures

[0006] The present invention provides modified nucleotides of the general formula (I):

$$\mathbf{DG - E' - G - E - Nu} \qquad \text{(I)}$$

where:

DG is a detectable group,

E and E' are the same and different group including a central main group element selected from the group consisting of boron (B), carbon (C), nitrogen (N), oxygen (O), silicon (Si), phosphorus (P),

sulfur (S), gallium (Ga) and germanium (Ge),

G is a linking group, and

Nu is a natural or synthetic nucleotide.

[0007] G can include a linear or branched alkenyl group or an alkenyl group including a central ring structure.

Structures with Ring Structure in the Core

[0008] The present invention provides modified nucleotides of the general formula (II):

$$\mathbf{DG-E'-R^2-A-R^1-E-Nu} \qquad \text{(II)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R^1 and R^2 are the same or different and are carbenyl groups,

A is a ring structure, and

Nu is a natural or synthetic nucleotide.

[0009] The present invention also provides modified nucleotides of the general formulas (III or IIIa) (γ-phosphate modified):

$$\mathbf{DG-E'-R^2-A-R^1-E-P(O_2)OP(O_2)OP(O_2)\text{-}Sugar\text{-}Base} \qquad \text{(III)}$$

$$\mathbf{DG-E'-R^2-A-R^1-E-P(O_2)OP(OZ^1)OP(OZ^2)\text{-}Sugar\text{-}Base} \qquad \text{(IIIa)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or

a carbon atom triply bonded to DG or R^2),

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group as described herein.

[00010] The present invention also provides modified nucleotides of the general formulas (IV or IVa) (β-phosphate modified):

$$\mathbf{DG\text{–}E'\text{–}R^2\text{–}A\text{–}R^1\text{–}E\text{–}P(O)(OP(O_2)OH)OP(O_2)\text{-}Sugar\text{-}Base} \quad \text{(IV)}$$

$$\mathbf{DG\text{–}E'\text{–}R^2\text{–}A\text{–}R^1\text{–}E\text{–}P(O)(OP(OZ^1)OH)OP(OZ^2)\text{-}Sugar\text{-}Base} \quad \text{(IVa)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group as described herein.

[0011] The present invention also provides modified nucleotides of the general formula (V) (α-phosphate modified):

$$\mathbf{DG\text{–}E'\text{–}R^2\text{–}A\text{–}R^1\text{–}E\text{–}P(O)(OP(O_2)OP(O_2)OH)\text{-}Sugar\text{-}Base} \quad \text{(V)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group

($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0012] The present invention also provides modified nucleotides of the general formula (VI) (sugar modified):

$$\mathbf{DG\text{–}E'\text{–}R^2\text{–}A\text{–}R^1\text{–}E\text{–}Sugar(P(O_2)OP(O_2)OP(O_2)OH)Base} \quad \text{(VI)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0013] The present invention also provides modified nucleotides of the general formula (VII) (base modified):

$$\mathbf{DG\text{–}E'\text{–}R^2\text{–}A\text{–}R^1\text{–}E\text{–}Base\text{-}Sugar\text{-}P(O_2)OP(O_2)OP(O_2)OH} \quad \text{(VII)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group

($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0014] In formulas (II-VII), the ring structure A can be saturated, unsaturated or aromatic or can include a mixture of saturated, unsaturated, or aromatic rings. Each ring in a ring structure include from 3 to about 12 main group elements. Of course, higher ordered rings are also included. Each carbyl group and each carbenyl group include from 1 to 40 carbon, where one or more of the carbon atoms can be replaced with a hetero atoms selected from the group consisting of B, C, Si, Ge, N, P, As, O, S, or Se and having sufficient hydrogen atoms to satisfy the valency of the group, where one or more hydrogen atoms can be replaced with F, Cl, Br, I, OR, SR, COR, COOR, $CONH_2$, CONHR, CONRR', or any other monovalent group inert or substantially inert under the substitution/displacement reaction conditions. It should be recognized that the linker group comprises $-\mathbf{R^2}-\mathbf{A}-\mathbf{R^1}-$ in the formulas (II-VII)

[0015] The present invention also provides a method for using the compounds of Formulas (II-VII) in single molecule sequencing including the step adding a compound of Formulas (II-VII) and detecting the detectable group before, during and/or after incorporation of one or a series of compounds of Formulas (II-VII).

[0016] The present invention also provides a method for using the compounds of Formulas (II-VII) in single molecule sequencing including the step adding a compound of Formulas (II-VII), where the detectable group is a fluorophore and detecting light from the fluorophore before, during and/or after incorporation of one or a series of compounds of Formulas (II-VII).

[0017] The present invention also provides a method for using the compounds of Formulas (II-VII) in single molecule sequencing including the step adding a compound of Formulas (II-VII), where the detectable group is an acceptor fluorophore and detecting light from the acceptor fluorophore after fluorescence resonance energy transfer from a donor fluorophore before, during and/or after incorporation of one or a series of compounds of Formulas (II-VII).

[0018] The Formulas (II-VII) can also includes other groups at different location of the nucleotide

including the phosphates, sugar and/or base. The additional groups are not intended to be detectable groups, but are groups designed to change the incorporation timing of the nucleotide modified with these additional groups. The additional groups can be atom replacements on the phosphates such as replacing an oxygen atom with a sulfur, nitrogen containing group, a carbon containing group, a boron containing group or any other group or atom that will change the incorporation timing of the nucleotide. In this way, sequencing can be performed with fewer distinct detectable groups, e.g., dATP and dTTP could be have the same detectable group, but modified with different additional groups so that one incorporates much faster than the other so that the detection signature of the incorporation will be distinguishable. These additional groups could also improve detectability of the detectable group by interacting with detectable group in a way that changes during the incorporation cycle - binding, incorporation, and pyrophosphate release.

Structures with Chains in the Core

[0019] The present invention provides modified nucleotides of the general formula (VIII):

$$\mathbf{DG-E'-R-E-Nu} \qquad \text{(VIII)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 2 and 10), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group, and

Nu is a natural or synthetic nucleotide.

[0020] The present invention also provides modified nucleotides of the general formulas (IX or IXa) (γ-phosphate):

$$\mathbf{DG-E'-R-E-P(O_2)OP(O_2)OP(O_2)\text{-}Sugar\text{-}Base} \qquad \text{(IX)}$$

$$\mathbf{DG-E'-R-E-P(O_2)OP(OZ^1)OP(OZ^2)\text{-}Sugar\text{-}Base} \qquad \text{(IXa)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an

oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group ($P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base, and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group as described herein.

[0021] The present invention also provides modified nucleotides of the general formulas (X or Xa)(β-phosphate):

$$\mathbf{DG\text{–}E'\text{–}R\text{–}E\text{–}P(O)(OP(O_2)OH)OP(O_2)\text{-}Sugar\text{-}Base} \quad (X)$$

$$\mathbf{DG\text{–}E'\text{–}R\text{–}E\text{–}P(O)(OP(OZ^1)OH)OP(OZ^2)\text{-}Sugar\text{-}Base} \quad (Xa)$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group ($P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base, and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group as described herein.

[0022] The present invention also provides modified nucleotides of the general formula (XI)(α-

phosphate):

$$\mathbf{DG\text{–}E'\text{–}R\text{–}E\text{–}P(O)(OP(O_2)OP(O_2)OH)\text{-}Sugar\text{-}Base} \quad \text{(XI)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), a phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0023] The present invention also provides modified nucleotides of the general formula (XII):

$$\mathbf{DG\text{–}E'\text{–}R\text{–}E\text{–}Sugar(P(O_2)OP(O_2)OP(O_2)OH)Base} \quad \text{(XII)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), a phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0024] The present invention also provides modified nucleotides of the general formula (XIII):

$$\mathbf{DG\text{–}E'\text{–}R\text{–}E\text{–}Base\text{-}Sugar\text{-}P(O_2)OP(O_2)OP(O_2)OH} \quad \text{(XIII)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0025] Each carbyl group and each carbenyl group include from 1 to 40 carbon, where one or more of the carbon atoms can be replaced with a hetero atoms selected from the group consisting of B, C, Si, Ge, N, P, As, O, S, or Se and having sufficient hydrogen atoms to satisfy the valency of the group, where one or more hydrogen atoms can be replaced with F, Cl, Br, I, OR, SR, COR, COOR, $CONH_2$, CONHR, CONRR', or any other monovalent group inert or substantially inert under the substitution/displacement reaction conditions. It should be recognized that the linker comprises – **R** – in formulas (VIII-XIII).

[0026] The present invention also provides a method for using the compounds of Formulas (VIII-XIII) in single molecule sequencing including the step adding a compound of Formulas (VIII-XIII) and detecting the detectable group before, during and/or after incorporation of one or a series of compounds of Formulas (VIII-XIII).

[0027] The present invention also provides a method for using the compounds of Formulas (VIII-XIII) in single molecule sequencing including the step adding a compound of Formulas (VIII-XIII), where the detectable group is a fluorophore and detecting light from the fluorophore before, during and/or after incorporation of one or a series of compounds of Formulas (VIII-XIII).

[0028] The present invention also provides a method for using the compounds of Formulas (VIII-XIII) in single molecule sequencing including the step adding a compound of Formulas (VIII-XIII), where the detectable group is an acceptor fluorophore and detecting light from the acceptor fluorophore after fluorescence resonance energy transfer from a donor fluorophore before, during and/or after incorporation of one or a series of compounds of Formulas (VIII-XIII).

[0029] The Formulas (VIII-XIII) can also includes other groups at different location of the nucleotide including the phosphates, sugar and/or base. The additional groups are not intended to be detectable groups, but are groups designed to change the incorporation timing of the nucleotide modified with these additional groups. The additional groups can be atom replacements on the phosphates such as replacing an oxygen atom with a sulfur, nitrogen containing group, a carbon containing group, a boron containing group or any other group or atom that will change the incorporation timing of the nucleotide. In this way, sequencing can be performed with fewer distinct detectable groups, *e.g.*, dATP and dTTP could be have the same detectable group, but modified with different additional groups so that one incorporates much faster than the other so that the detection signature of the incorporation will be distinguishable. These additional groups could also improve detectability of the detectable group by interacting with detectable group in a way that changes during the incorporation cycle - binding, incorporation, and pyrophosphate release.

BRIEF DESCRIPTION OF THE DRAWINGS

[0030] The invention can be better understood with reference to the following detailed description together with the appended illustrative drawings in which like elements are numbered the same.

[0031] **Figure 1** depicts exemplary single ring linker structures of this invention.

[0032] **Figure 2** depicts other exemplary single ring linker structures of this invention.

[0033] **Figures 3&4** depicts exemplary binary ring linker structures of this invention.

[0034] **Figure 5** depicts exemplary trinary ring linker structures of this invention.

[0035] **Figure 6** depicts exemplary dyes for use in the modified nucleotide structures of this invention.

[0036] **Figure 7** depicts two synthetic schemes for preparing modified nucleotide triphosphates, where the modification is a linker terminating in a dye.

[0037] **Figure** depicts a synthetic scheme for preparing modified nucleotide triphosphates, where the modification is a linker terminating in a dye.

[0038] **Figure 9** depicts a synthetic scheme for preparing modified nucleotide triphosphates, where the modification is a linker terminating in a dye.

[0039] **Figure 10** depicts a synthetic scheme for preparing modified nucleotide triphosphates, where the modification is a linker terminating in a dye.

[0040] **Figure 11** depicts a synthetic scheme for preparing modified nucleotide triphosphates, where the modification is a linker terminating in a dye.

[0041] **Figure 12** depicts a synthetic scheme for preparing modified nucleotide triphosphates, where the modification is a linker terminating in a dye.

[0042] **Figure 13** depicts a synthetic scheme for preparing modified nucleotide triphosphates, where

the modification is a linker terminating in a dye.

DETAILED DESCRIPTION OF THE INVENTION

[0043] The inventors have found that modified nucleotide for use in sequencing experiments can be constructed from a linker group including a central group and terminal groups including a main group element . The central group can be a linear carbenyl group, a branched carbenyl group or a arenyl group. The hydroxy group is adapted to react with a nucleotide at a phosphate moiety, a sugar moiety and/or base moiety. The nucleotide can be naturally occurring or human created, where the human created nucleotide have altered incorporation rates and/or fidelities. The amino group is adapted to react with a detectable groups such as a fluorescent dye.

[0044] The present invention broadly relates to modified nucleotides of the general formula (I):

$$\mathbf{DG - E' - R^2 - G - R^1 - E - Nu} \quad \text{(I)}$$

where DG is a detectable group, E and E' are the same and different group including a central main group element, R^1 and R^2 are the same or different and are carbenyl groups, G is a central group, and Nu is a natural or synthetic nucleotide. The central group G can be a ring structure or an alkenyl group. If it is an alkenyl group, then R^2 - G - R^1 can be re-designated by the symbol R.

[0045] The present invention relates also broadly to modified nucleotide including a linker having a central ring structure, an amino terminated moiety and a hydroxy terminated moiety. The central ring structure can be a saturated ring structure, a partially unsaturated ring structure or an aromatic ring structure. The modified nucleotides including compounds of the general formula (II):

$$\mathbf{DG - E' - R^2 - A - R^1 - E - Nu} \quad \text{(II)}$$

where DG is a detectable group, E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), phosphito ($P(OR^3)O$), phosphate ($P(O_2)O$), polyphosphate $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), silyl ($Si(R^3)_2$), siloxyl ($Si(OR^3)_2$) carboxy group (C(O)O), keto (C(O)), amido group ($C(O)N(R^3)$), urea group ($N(R^3)C(O)N(R^3)$), carbonate (OC(O)O), urethane group ($OC(O)N(R^3)$), is a nitrogen atom, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2), R^1 and R^2 are the same or different and are is carbenyl groups, A is a ring structure, and Nu is a natural or synthetic nucleotide. The ring structure A is saturated, unsaturated or aromatic or can include a mixture of saturated, unsaturated, or aromatic rings. Each carbyl group and each carbenyl group include from about 1 to about 40 carbon atoms, where one or more of the carbon atoms is replaced with an hetero atom or an hetero atom containing group.

[0046] The present invention relates also broadly to modified nucleotides of the general formula (III):

$$\mathbf{DG - E' - R - E - Nu} \quad \text{(VIII)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group, and

Nu is a natural or synthetic nucleotide.

Methods for Preparing Modified Nucleotides

[0047] The present invention also relates to methods for preparing modified nucleotides, especially gamma (γ) phosphate modified nucleotides. One such method includes the step of reacting a nucleotide triphosphate with a diamine in N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide (EDC) or pre-cyclizing a nucleotide triphosphate in N,N-dicyclohexylcarbodiimide (DCC) and then reacted with a diamine. Both routes produce a diamine functionalized gamma (γ) phosphate modified nucleotide terminating in a free amino group in yields greater than about 50%. The free amino group can then be treated with an acid, an anhydride or an acid chloride to produce an amide functionalized gamma (γ) phosphate modified nucleotide, where the amido group (which can bear a fluorophore) is separated from the gamma (γ) phosphate by a linker or linking group – the portion of the diamine excluding the two terminal amino groups – H_2N - L - NH_2, where L is the linking group which can be a R^2 - G - R^1 motif, R^2 - A - R^1 motif or an R motif as shown in Formulas (I), (II), and (VIII) above. These two methods are shown in pictorially in **Figure 7**.

[0048] The second method set forth above can also be used to prepare modified gamma (γ) phosphate nucleotide triphosphates, where the linker molecule is of the general motif **E' – R^2 – A – R^1 – E** as set forth in Formula (II) and shown pictorially in **Figure 10**.

[0049] Another such method includes the step of reacting a linker molecule including an N-protected, terminal amino group and a hydroxy terminal group (Protector - HN - L - OH) with phosphate to produce a linker molecule bearing a terminal phosphate group (Protector - HN - L - $OP(O)OH_2$). Next, the terminal phosphate linker molecule is activated with carbonyldiimidazole to produce an imidazole activated terminal phosphate linker molecule. The imidazole activated terminal phosphate

linker molecule is reacted with a nucleotide diphosphate to produce a protected-amino terminated, functionalized gamma (γ) phosphate modified nucleotide triphosphate. The protected-amino terminated, functionalized gamma (γ) phosphate modified nucleotide triphosphate is then deprotected and the free amine is then treated with an acid, an anhydride or an acid chloride to produce an amide functionalized gamma (γ) phosphate modified nucleotide, where the amido group (which can bear a fluorophore) is separated from the gamma (γ) phosphate by a linker or linking group – the portion of the diamine excluding the two terminal amino groups – H_2N - L - OH, where L is the linking group which can be a R^2 - G - R^1 motif, R^2 - A - R^1 motif or an R motif as shown in Formulas (I), (II), and (VIII) above. This method is shown in pictorially in **Figure 8**.

[0050] The above multi-step reaction can also be used to produce functionalized nucleotide polyphosphates. This method is shown in pictorially in **Figure 9**, which evidences a general synthesis for functionalized nucleotide tetraphosphates.

[0051] An alternate multi-step reaction similar to the multi-step reaction above can also be used to produce functionalized nucleotide polyphosphates. The alternate reaction starts with an amino and phosphate terminated linker molecule, where the amino group is then protected before reacting the phosphate linker with carbonyldiimidazole. This method is shown in pictorially in **Figure 11**, which evidences a general synthesis for functionalized nucleotide tetraphosphates.

[0052] Another such method includes the step of reducing an amine terminated alkylated benzoic acid to produce an amine terminated alkylated, a hydroxy terminated alkylate benzene linker molecule. The linker is then amine protected and the hydroxy group is sulfonated. The sulfonated, protected linker molecule is then reacted with phosphate to form a phosphate, protected linker molecule. The phosphate, protected linker molecule is then activated with imidazole and reacted with a nucleotide diphosphate to form a gamma phosphate functionalized nucleotide triphosphate. However, deprotecting of the amino group resulted in very poor yields. Thus, this method is of little utility in forming gamma phosphate functionalized nucleotide triphosphate. However, an alternate reaction scheme did result in a general synthetic scheme to prepare gamma phosphate functionalized nucleotide triphosphates. The alternate synthesis includes reducing an amine terminated alkylated benzoic acid to produce an amine terminated alkylated, a hydroxy terminated alkylate benzene linker molecule. The linker is then amine protected with TFA protecting group. The TFA protected linker molecule is then reacted with a cyclized nucleotide triphosphate to produce a TFA protected gamma phosphate functionalized nucleotide triphosphate. Deprotecting and dye treatment produces dye gamma phosphate functionalized nucleotide triphosphates.

[0053] For additional information on DNA sequencing, data acquisition and analysis, monomers, monomers synthesis, or other features of system that are amenable to detection using the apparatuses

and methods of this invention, the reader is referred to United States Patent, Published Patent Application and Pending Patent Application Nos. 09/901,782; 10/007621; 11/007794; 11/671,956; 11/694605; 2006-0078937; 6,982,146; 7,169,560; 7,220,549, 20070070349; 20070031875; 20070012113; 20060286566; 20060252077; 20060147942; 200601336144; 20060024711; 20060024678; 20060012793; 20060012784; 20050100932; incorporated herein by reference.

Suitable Reagents

[0054] Suitable detectable agents include, without limitation, any group that is detectable by a known or yet to be invented analytical technique. Exemplary examples include, without limitation, fluorophores or chromophorers, group including one or a plurality of nmr active atoms (^{2}H, ^{11}B, ^{13}C, ^{15}N, ^{17}O, ^{19}F, ^{27}Al, ^{29}Si, ^{31}P, nmr active transition metals, nmr active actinide metals, nmr active lanthanide metals), IR active groups, nearIR active groups, Raman active groups, UV active groups, X-ray active groups, light emitting quantum dots, light emitting nano-structures, or other structures or groups capable of direct detection or that can be rendered detectable or mixtures or combinations thereof.

[0055] Suitable atomic tag for use in this invention include, without limitation, any atomic element amenable to attachment to a specific site in a polymerizing agent or dNTP, especially Europium shift agents, nmr active atoms or the like.

[0056] Suitable atomic tag for use in this invention include, without limitation, any atomic element amenable to attachment to a specific site in a polymerizing agent or dNTP, especially fluorescent dyes such as d-Rhodamine acceptor dyes including dichloro[R110], dichloro[R6G], dichloro[TAMRA], dichloro[ROX] or the like, fluorescein donor dye including fluorescein, 6-FAM, or the like;Acridine including Acridine orange, Acridine yellow, Proflavin, or the like; Aromatic Hydrocarbon including 2-Methylbenzoxazole, Ethyl p-dimethylaminobenzoate, Phenol, benzene, toluene, or the like; Arylmethine Dyes including Auramine O, Crystal violet, Crystal violet, Malachite Green or the like; Coumarin dyes including 7-Methoxycoumarin-4-acetic acid, Coumarin 1, Coumarin 30, Coumarin 314, Coumarin 343, Coumarin 6 or the like; Cyanine Dye including 1,1'-diethyl-2,2'-cyanine iodide, Cryptocyanine, Indocarbocyanine (C3) dye, Indodicarbocyanine (C5) dye, Indotricarbocyanine (C7) dye, Oxacarbocyanine (C3) dye, Oxadicarbocyanine (C5) dye, Oxatricarbocyanine (C7) dye, Pinacyanol iodide, Stains all, Thiacarbocyanine (C3) dye, Thiacarbocyanine (C3) dye, Thiadicarbocyanine (C5) dye, Thiatricarbocyanine (C7) dye, or the like; Dipyrrin dyes including N,N'-Difluoroboryl-1,9-dimethyl-5-(4-iodophenyl)-dipyrrin, N,N'-Difluoroboryl-1,9-dimethyl-5-[(4-(2-trimethylsilylethynyl), N,N'-Difluoroboryl-1,9-dimethyl-5-phenydipyrrin, or the like; Merocyanines including 4-(dicyanomethylene)-2-methyl-6-(p-dimethylaminostyryl)-4H-pyran (DCM), 4-(dicyanomethylene)-2-methyl-6-(p-dimethylaminostyryl)-

4H-pyran (DCM), 4-Dimethylamino-4'-nitrostilbene, Merocyanine 540, or the like; Miscellaneous Dye including 4',6-Diamidino-2-phenylindole (DAPI), 4',6-Diamidino-2-phenylindole (DAPI), 7-Benzylamino-4-nitrobenz-2-oxa-1,3-diazole, Dansyl glycine, Dansyl glycine, Hoechst 33258, Hoechst 33258, Lucifer yellow CH, Piroxicam, Quinine sulfate, Quinine sulfate, Squarylium dye III, or the like; Oligophenylenes including 2,5-Diphenyloxazole (PPO), Biphenyl, POPOP, p-Quaterphenyl, p-Terphenyl, or the like; Oxazines including Cresyl violet perchlorate, Nile Blue, Nile Red, Nile blue, Oxazine 1, Oxazine 170, or the like; Polycyclic Aromatic Hydrocarbons including 9,10-Bis(phenylethynyl)anthracene, 9,10-Diphenylanthracene, Anthracene, Naphthalene, Perylene, Pyrene, or the like; polyene/polyynes including 1,2-diphenylacetylene, 1,4-diphenylbutadiene, 1,4-diphenylbutadiyne, 1,6-Diphenylhexatriene, Beta-carotene, Stilbene, or the like; Redox-active Chromophores including Anthraquinone, Azobenzene, Benzoquinone, Ferrocene, Riboflavin, Tris(2,2'-bipyridyl)ruthenium(II), Tetrapyrrole, Bilirubin, Chlorophyll a, Chlorophyll b, Diprotonated-tetraphenylporphyrin, Hematin, Magnesium octaethylporphyrin, Magnesium octaethylporphyrin (MgOEP), Magnesium phthalocyanine (MgPc), Magnesium phthalocyanine (MgPc), Magnesium tetramesitylporphyrin (MgTMP), Magnesium tetraphenylporphyrin (MgTPP), Octaethylporphyrin, Phthalocyanine (Pc), Porphin, Tetra-t-butylazaporphine, Tetra-t-butylnaphthalocyanine, Tetrakis(2,6-dichlorophenyl)porphyrin, Tetrakis(o-aminophenyl)porphyrin, Tetramesitylporphyrin (TMP), Tetraphenylporphyrin (TPP), Vitamin B12, Zinc octaethylporphyrin (ZnOEP), Zinc phthalocyanine (ZnPc), Zinc tetramesitylporphyrin (ZnTMP), Zinc tetramesitylporphyrin radical cation, Zinc tetraphenylporphyrin (ZnTPP), or the like; Cy3, Cy3B, Cy5, Cy5.5, Atto590, Atto610, Atto611, Atto611x, Atto620, Atto655, Alexa488, Alexa546, Alexa594, Alexa610, Alexa610x, Alexa633, Alexa647, Alexa660, Alexa680, Alexa700, Bodipy630, DY610, DY615, DY630, DY632, DY634, DY647, DY680, DyLight647, HiLyte647, HiLyte680, LightCycler (LC) 640, Oyster650, ROX, TMR, TMR5, TMR6; Xanthenes including Eosin Y, Fluorescein, Fluorescein, Rhodamine 123, Rhodamine 6G, Rhodamine B, Rose bengal, Sulforhodamine 101, or the like; or mixtures or combination thereof or synthetic derivatives thereof or FRET fluorophore-quencher pairs including DLO-FB1 (5'-FAM/3'-BHQ-1) DLO-TEB1 (5'-TET/3'-BHQ-1), DLO-JB1 (5'-JOE/3'-BHQ-1), DLO-HB1 (5'-HEX/3'-BHQ-1), DLO-C3B2 (5'-Cy3/3'-BHQ-2), DLO-TAB2 (5'-TAMRA/3'-BHQ-2), DLO-RB2 (5'-ROX/3'-BHQ-2), DLO-C5B3 (5'-Cy5/3'-BHQ-3), DLO-C55B3 (5'-Cy5.5/3'-BHQ-3), MBO-FB1 (5'-FAM/3'-BHQ-1), MBO-TEB1 (5'-TET/3'-BHQ-1), MBO-JB1 (5'-JOE/3'-BHQ-1), MBO-HB1 (5'-HEX/3'-BHQ-1), MBO-C3B2 (5'-Cy3/3'-BHQ-2), MBO-TAB2 (5'-TAMRA/3'-BHQ-2), MBO-RB2 (5'-ROX/3'-BHQ-2); MBO-C5B3 (5'-Cy5/3'-BHQ-3), MBO-C55B3 (5'-Cy5.5/3'-BHQ-3) or similar FRET pairs available from Biosearch Technologies, Inc. of Novato, CA, fluorescent quantum dots (stable long lived fluorescent

donors), tags with nmr active groups, Raman active tags, tags with spectral features that can be easily identified such as IR, far IR, near IR, visible UV, far UV or the like. It should be recognized that any molecule, nano-structure, or other chemical structure that is capable of chemical modification and includes a detectable property capable of being detected by a detection system. Such detectable structure can include one presently known and structures that are being currently designed and those that will be prepared in the future.

[0057] Referring now to **Figure 1**, a set of exemplary single ring linkers are shown, where E^R is a main element containing group such as CH, SiH, N, P, or the like. The ring structure can also be saturated or unsaturated, but not aromatic in which case E^R is a main element containing group such as CH, SiH, N, P, O, S, or the like. R is a carbyl group and n is an integer having a value between 1 and the maximum number of R groups that the ring structure can accommodate and still be a compound known or capable of synthesis by known synthetic methods.

[0058] Referring now to **Figure 2**, a set of exemplary single ring linkers are shown, where E^{R1}, E^{R2} and E^{R3} are the same or different main element containing groups such as CH, SiH, N, P, or the like. The ring structure can also be saturated or unsaturated, but not aromatic in which case E^{R1}, E^{R2} and E^{R3} are the same or different main element containing groups such as CH, SiH, N, P, O, S, or the like. R is a carbyl group and n is an integer having a value between 1 and the maximum number of R groups that the ring structure can accommodate and still be a compound known or capable of synthesis by known synthetic methods.

[0059] Referring now to **Figures 3&4**, a set of exemplary binary ring linkers are shown, where E^{R1} and E^{R2} are the same or different main element containing groups such as CH, SiH, N, P, or the like. The ring structure can also be saturated or unsaturated, but not aromatic in which case $E^{R1}$and E^{R2} are the same or different main element containing groups such as CH, SiH, N, P, O, S, or the like. R is a carbyl group and n is an integer having a value between 1 and the maximum number of R groups that the ring structure can accommodate and still be a compound known or capable of synthesis by known synthetic methods. It should be recognized that the second ring can be any other sized ring besides a six membered ring.

[0060] Referring now to **Figure 5**, a set of exemplary trinary ring linkers are shown, where E^R is a main element containing group such as CH, SiH, N, P, or the like. The ring structure can also be saturated or unsaturated, but not aromatic in which case E^R is a main element containing group such as CH, SiH, N, P, O, S, or the like. R is a carbyl group and n is an integer having a value between 1 and the maximum number of R groups that the ring structure can accommodate and still be a compound known or capable of synthesis by known synthetic methods. It should be recognized that the second ring can be any other sized ring besides a six membered ring.

EXPERIMENTS OF THE INVENTION

Example 1

[0061] This example illustrates the preparation of dATP bonded to 1,4-Bis-(3-aminopropyl)piperazine to form dATP-BAPP.

[0062] 20 μmol of the sodium salt of dATP was treated with Dowex resin/TEAB, lyophilized and dried under vacuum. Dry DCC (75μmol) was added to a DMF (200 μL) solution of the above nucleotide and the resulting mixture was stirred under argon for 2 hrs. Pyridine (17 μL) was added and the resulting mixture was slowly evaporated. To the pellet was added a solution of 1,4-Bis-(3-aminopropyl)piperazine (150 μmol) in DMF (200μL) and the solution was stirred for 12 hrs. The mixture was then quenched with water and centrifuged to remove the solid. The clear solution was subject to HPLC (SAX, TEAB) purification. The product was collected and lyophilized. The pellet was dissolved in HEPES buffer (10mM, pH 8.5). Yield 4.2 μmol, 21%. Note: The yield is not accurate because aliquots from the DCC-reacted mixture were transferred into several reactions including the top one. Normally, the yield of this synthesis is much higher (>50%).

Example 2

[0063] This example illustrates the preparation of the compound of Example 1 bonded to ROX to form dATP-BAPP-ROX.

[0064] The above nucleotide dATP-11 (0.5 μmol) was reacted with ROX-SE (2μmol) overnight in the following mixture: DMF (20μL) + $NaHCO_3$ (1 M, pH 9). The product was purified on a Sephadex G25 column and then on HPLC (C18, TEAA/MeOH). Yield 0.41 μmol, 82%.

Example 3

[0065] This example illustrates an enzymatic tests on dATP-BAPP-ROX.

[0066] This nucleotide was tested upon calf intestinal alkaline phosphatese (CIAP) and phosphodiesterase 1 (PDE1) and the result was analyzed on PEI cellulose thin-layer chromatography. It was inert to CIAP and readily hydrolyzed by PDE1.

Fluorophores

[0067] Referring now to **Figure 6**, fluorophores used in the synthesis of fluorophore modified dNTPs are shown.

Pictorial Examples of dNTP Modification Schemes

[0068] Referring now to **Figure 7-11**, a number of synthetic schemes for prepare modified dNTPs are shown.

I. General Synthetic Scheme for Nitrogen Terminated Linkers

[0069] This general scheme involves coupling an dNTP to a nitrogen-terminated linker in the

presence of N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide (EDC) in two steps.

Step 1 - dNTP-1(TEA^+)

[0070] Nucleotide dNTP Na_2 (12.7μmol) is reacted with linker 1 (110μmol) in the presence of EDC (110μmol) at rt for 3hr and pH is maintained at ~5.7 over the time. The product is purified on HPLC (C18) with TEAA/MeOH or on HPLC (SAX) with TEAB. The product after lyophilization is dissolved in HEPES buffer (10mM, pH 8.5). Yield varies from 35% to 55%.

Step 1 - dNTP-1-dye (TEA^+)

[0071] The intermediate dNTP-1 (1μmol) in $NaHCO_3$ buffer (1M, pH9, 50μL) and dye-NHS (2.5μmol) in DMF (100μL) are mixed and reacted overnight. After a Sephadex G-25 column purification, the product-containing sample is purified on HPLC (C18) with TEAA/MeOH or on HPLC (SAX) with TEAB. The product after lyophilization is dissolved in HEPES buffer (10mM, pH8.5). Yield varies from 15% to 70%. Enzymatic assay and MS are performed when necessary.

II. Alternate General Synthetic Scheme for Nitrogen-Terminated Linkers

[0072] This general scheme involves activation of dNTP by DCC and then coupling the intermediate to a linker.

Step 1 - dNTP(TEA^+)

[0073] Nucleotide dNTP·Na_2 (57 μmol) is passed through a TEAB-equilibrated Dowex resin (H^+) column. The sample is lyophilized.

Step 2 - dNTP-2 (TEA^+)

[0074] Nucleotide dNTP·TEA (20 μmol) is coevaporated with TEA and methanol 3 times before dried under vacuum overnight. DCC (75 μmol) is dried under vacuum 2 hrs. Linker compound (200μmol) is coevaporated with TEA and methanol and dried under vacuum overnight.

[0075] DCC is transferred to dNTP·TEA in DMF/MeOH (200μL/20μL) and the mixture is stirred at r.t. for 3-4hrs before coevaporated with pyridine (17μL). Linker compound in DMF (200-300μL) is then added to the pellet and the resulting solution is stirred at r.t. overnight. The product is purified on HPLC (SAX column, TEAB). After lyophilization the product is dissolved in HEPES buffer (10mM, pH 8.5). Yield varies from 30% to 50%.

Step 3 - dNTP-2-dye (TEA^+)

[0076] This step is similar to Step 2 of the first general scheme.

III. General Synthetic Scheme for Linkers Terminated by Nitrogen at One End and Oxygen at the Other End

[0077] This scheme involves activating a monophosphate with CDI and coupling the intermediate to a dNDP.

Step 1 - dNDP (TBA+)

[0078] Nucleotide dNDP sodium salt (43 umol) is passed through Dowex resin (H+) into cooled TBA. It is coevaporated with DMF 3 times and dried under vacuum overnight.

Step 2 - Pi-5-Cbz (TBA+)

[0079] Alcohol 5-Cbz is phosphorylated with $POCl_3/P(OMe)_3$ system and purified on Sephadex G25 DEAE anion exchanger with a gradient of AB buffer. After lyophilization it was transformed into TBA+ salt as described in Step 1. Yield varies from 50% to 70%.

Step 3 - dNTP-5-Cbz (TEA+)

[0080] All reagents are dried under vacuum. Monophosphate Pi-5-Cbz (TBA+, 33μmol)) is treated with CDI (165umol) in DMF (250μL) for 6 hrs before MeOH (264umol) is added to quench the excess of CDI. Nucleotide dNDP (TBA+, 43umol) is added in DMF (400μL) and the reaction is allowed overnight. Purification was achieved on HPLC (SAX, TEAB) followed by lyophilization. Yields vary from 20% to 50%.

Step 4 - dNTP-5 (TEA+)

[0081] Nucleotide dNTP-5-Cbz (TEA+) is treated with ammonium formate and Pd/C for 10-20minutes. The product is purified on HPLC (SAX, TEAB) followed by lyophilization. Yields are above 90%.

Step 5 - dNTP-5-dye (TEA+)

[0082] Same procedure as performed for the second step of the first general procedure.

Overview

[0083] 1. dNDP (1) is converted to tetrabutylammonium salt (2) by cation exchange. This converts the diphosphate into a reagent that is very soluble in dry organic solvents

[0084] 2. Potential hydroxyl terminated linker (3) (containing a nitrogen protecting group, as example here, trifluoroactetate (TFA)) is phosphorylated with phosphorous oxytrichloride to yield activated chlorophosphate (4). Excess $POCl_3$ is removed by evaporation yielding the dichloride ester

[0085] 3. (2) and (4) are reacted in an anhydrous solvent (for example dry DMF), followed by hydrolysis of the resulting (monochloro)triphosphate ester to the triphosphate (5)

Procedure for Preparing the dNTP with Linker Attached to γ-Phosphate Through P-O Linkage

Step 1 - See **Figure 13.**

[0086] This step illustrates the conversion of a dNDP-sodium salt (1) to a dNDP-tetrabutylammonium salt (2).

[0087] An aqueous solution (2 mL) of the commercially available dNDP-sodium salt (100 to 150 mg)

was loaded onto a strong cation exchange (-SO_3H) packed column. The column was eluted with gravity. Fractions were collected and checked by spotting on TLC and visualized by UV lamp (dNDP will have show blue spot under UV). The desired fractions were pooled together and quenched with tetrabutylammonium hydroxide (1.01 eq in ~10 mL H_2O) immediately at 0 °C. The solution was evaporated to dryness. The residue was re-dissolved in DMF and dried down. When this material was dried down 3X with DMF, the dNDP-Tetrabutylammonium salt was ready for the coupling reaction.

Step 2

[0088] This step illustrates the coupling of the linker (3) to dNDP (2) using dCDP and neutral EO linker as an example.

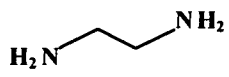
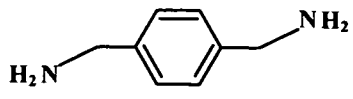
Material

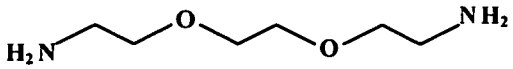
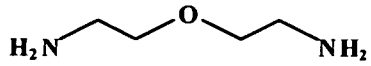
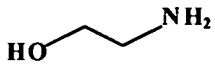
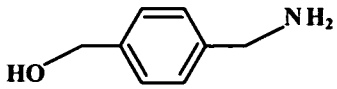
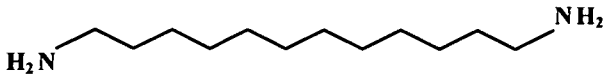
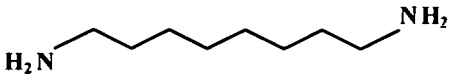
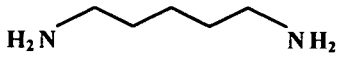
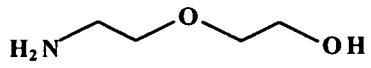
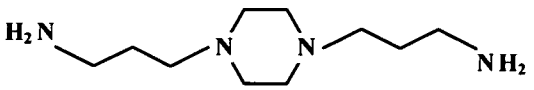
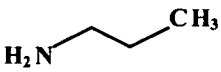
[0089] TFA protected linker (3): 10 mg (0.0497 mmole; dried by co-evaporating with DMF three times before use); $POCl_3$: 9.2 mL (0.0994 mmole, 2x); dCDP-tetrabutylammonium salt (2) (24.85 μmole, quantity determined by UV absorbance at 260 nm with ε~9,300); dry methylene dichloride (DCM); dry dimethylformamide (DMF); 1.5 M triethylammonium bicarbonate (TEAB) buffer (pH ~7.5 to 8).

[0090] A solution of the linker (3) in dry DCM (0.5 mL) was added into the solution of $POCl_3$ in DCM (1 mL). at 0 °C. The reaction was then stirred at 0 °C for three hours. The solution was then evaporated to dryness under reduced pressure and was further dried down with high vacuum for another 10 minutes to remove the residual $POCl_3$. The residue (4) was then re-dissolved in dry DMF (1 mL). To this solution, dCDP (2) (in dry 0.5 mL of dry DMF) was added in at 0 °C. The reaction was then stirred at 0 °C initially and then the temperature was gradually raised to ambient. The reaction was then stirred at room temperature overnight, followed by quenching with the addition of TEAB buffer (5mL) at 0 °C. The mixture was then stirred at 0 °C for three more hours. The product (5) was evaporated to dryness, re-dissolved in water, material and purified by reverse phase HPLC (C-18 column).

[0091] The linkers tabulated in Table 1 were used in the above articulated preparatory methods.

TABLE 1
Linker Used in the Various Preparatory Methods

Linker	Structure
1	
2	

Linker	Structure
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

[0092] The modified nucleotides tabulated in Table 2 were prepared using the above articulated preparatory methods. In Table 2, the numbers represent the linkers tabulated above. The structures with Cbz are protected linker structures where the terminal amino group has been reacted with benzoic acid.

TABLE 2
Modified Nucleotides Prepare Using the Various Preparatory Methods

Modified Nucleotide	Modified Nucleotide	Modified Nucleotide	Modified Nucleotide
ATP-1	ATP-6	dATP-6-ROX-2	dGTP-2-Cy5
ATP-1-ATTO620	ATP-6-Cbz	dATP-10	dGTP-2-Oyster650
ATP-1-Biotin	ATP-7	dATP-10-Cbz	dGTP-2-ROX
ATP-1-Cy3	ATP-8	dATP-10-ROX	dGTP-2-ROX
ATP-1-Cy5	ATP-9	dATP-11	dGTP-3
ATP-1-Fluorescein	ATP-P5-Cbz	dATP-11-ROX	dGTP-3-Cy5
ATP-1-Fluoresecein1	ATPP-5	dATP-12	dGTP-3-ROX
ATP-1-Fluoresecein2	ATPP-5-biotin	dATP-12-Cbz	dGTP-4

Modified Nucleotide	Modified Nucleotide	Modified Nucleotide	Modified Nucleotide
ATP-1-ROX	ATPP-5Cbz	dATP-12-ROX	dGTP-4-Cy5
ATP-1-ROX (Na+)	ATPP-5ROX	dCTP-1	dGTP-4-ROX
ATP-1-ROX(5)	dATP-1	dCTP-1-Cy5	dTTP-1
ATP-1-ROX(6)	dATP-1-BODIPY	dCTP-1-LC	dTTP-1-Cy5
ATP-1-TMR1	dATP-1-Cy5	dCTP-1-ROX	dTTP-1-LC
ATP-1-TMR2	dATP-1-LC	dCTP-1-TMR	dTTP-1-ROX
ATP-2	dATP-1-ROX	dCTP-1-TMR1	dTTP-1-TMR
ATP-2-biotin	dATP-2	dCTP-1-TMR2	dTTP-1-TMR1
ATP-2-Cy3	dATP-2-Alx594	dCTP-2	dTTP-1-TMR2
ATP-2-Cy5	dATP-2-Alx610	dCTP-2	dTTP-2
ATP-2-Flu	dATP-2-Alx633	dCTP-2-Alx610	dTTP-2-Alx610
ATP-2-ROX	dATP-2-BODIPY	dCTP-2-Alx633	dTTP-2-Alx633
ATP-2-ROX(5)	dATP-2-Cy5	dCTP-2-Cy5	dTTP-2-Cy5
ATP-2-ROX(6)	dATP-2-Oys650	dCTP-2-ROX	dTTP-2-ROX
ATP-2-TMR-1	dATP-2-ROX	dCTP-3	dTTP-3
ATP-2-TMR-2	dATP-3	dCTP-3-Cy5	dTTP-3-Cy5
ATP-3	dATP-3-Bodipy	dCTP-3-ROX	dTTP-3-ROX
ATP-3-biotin-1	dATP-3-Cy5	dCTP-4	dTTP-4
ATP-3-Cy3	dATP-3-ROX	dCTP-4-Cy5	dTTP-4-Cy5
ATP-3-Cy5	dATP-4	dCTP-4-ROX	dTTP-4-ROX
ATP-3-Flu	dATP-4-Cy5	dCTP-6	dTTP-5
ATP-3-ROX	dATP-4-ROX	dCTP-6-Alx610	dTTP-5-Cbz
ATP-3-TMR	dATP-5	dGTP-1	dTTP-5-Cbz
ATP-4	dATP-5-Cbz	dGTP-1-ATTO620	dTTP-5-ROX
ATP-4-Biotin	dATPP-5-Cbz	dGTP-1-Cy5	dUTP-1
ATP-4-Cy3	dATPP-5	dGTP-1-LC	dUTP-1-Cy5
ATP-4-Cy5	dATPP-5-Cbz	dGTP-1-LC	dUTP-1-LC
ATP-4-Flu	dATPP-5-ROX	dGTP-1-ROX	dUTP-1-LC1
ATP-4-ROX	dATPP-5-ROX	dGTP-2	dUTP-1-LC2
ATP-4-TMR	dATP-6	dGTP-2-Alx594	dUTP-1-ROX
ATP-5	dATP-6-Cbz	dGTP-2-Alx610	dUTP-1-TMR1
ATP-5-Cbz	dATP-6-ROX-1	dGTP-2-Alx633	dUTP-1-TMR2
ATP-5-ROX			

Linker 10 - 1,1'-Carbonyldiimidazole (CDI) chemistry

[0093] Pi-10-Cbz (Bu_3NH^+) was vigorously dried down to get the weight for quantification. Its coupling with dADP (20μmol scale) was tried again with new reagents and still ended with a low yield (TLC & HPLC). The product, however, was purified on HPLC (SAX, TEAB) to give 0.53 μmol of dATP-10-Cbz. After hydrogenolysis deprotection (0.4umol scale), an aliquot of dATP-10 (44nmol) was directly labeled with ROX-SE. After Sephadex G-25 column and C18 HPLC purification, dATP-10-ROX (3.8nmol) was submitted to enzymology team to evaluate its polymerase incorporation.

Linker 10 – Trifluoroacetic anhydride / methylimidazole chemistry

[0094] The method reacts monophosphate with trifluoroacetic anhydride and the resulting mixed

anhydride is reacted with methylimidazole followed by dADP quenching. It was tested at 20μmol scale for the preparation of dATP-10-Cbz and gave a low yield of 0.5μmol. Although not widely used this chemistry takes advantage of the volatility of $(CF_3CO)_2O$ and CF_3COOH and represents a fast coupling method (< 3 hrs) compared to other known methods.

[0095] All references cited herein are incorporated by reference. Although the invention has been disclosed with reference to its embodiments, from reading this description those of skill in the art may appreciate changes and modification that may be made which do not depart from the scope and spirit of the invention as described above and claimed hereafter.

CLAIMS

We claim:

[0096] 1. A modified nucleotide of the general formula (I):

$$\mathbf{DG-E'-G-E-Nu} \quad (I)$$

where:

DG is a detectable group,

E and E' are the same and different group,

G is a linking group, and

Nu is a natural or synthetic nucleotide,

where G comprises a linear or branched alkenyl group or an alkenyl group including a central ring structure.

[0097] 2. The nucleotide of claim 1, wherein the E and E' group include a central main group element.

[0098] 3. The nucleotide of claim 1, wherein the central main group elements are selected from the group consisting of boron (B), carbon (C), nitrogen (N), oxygen (O), silicon (Si), phosphorus (P), sulfur (S), gallium (Ga) and germanium (Ge).

[0099] 4. The nucleotide of claim 1, wherein G includes a central ring structure.

[0100] 5. The nucleotide of claim 1, wherein G includes a linear alkenyl group.

[0101] 6. The nucleotide of claim 1, wherein the nucleotide has the general formula (II):

$$\mathbf{DG-E'-R^2-A-R^1-E-Nu} \quad (II)$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure,

R^1 and R^2 are the same or different and are carbenyl groups,

A is a ring structure, and

Nu is a natural or synthetic nucleotide.

[0102] 7. The nucleotide of claim 1, wherein the nucleotide has the general formulas (III or IIIa):

$$\mathbf{DG\text{-}E'\text{-}R^2\text{-}A\text{-}R^1\text{-}E\text{-}P(O_2)OP(O_2)OP(O_2)\text{-}Sugar\text{-}Base} \quad \text{(III)}$$

$$\mathbf{DG\text{-}E'\text{-}R^2\text{-}A\text{-}R^1\text{-}E\text{-}P(O_2)OP(OZ^1)OP(OZ^2)\text{-}Sugar\text{-}Base} \quad \text{(IIIa)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure,

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group.

[0103] 8. The nucleotide of claim 1, wherein the nucleotide has the general formulas (IV or IVa):

$$\mathbf{DG\text{-}E'\text{-}R^2\text{-}A\text{-}R^1\text{-}E\text{-}P(O)(OP(O_2)OH)OP(O_2)\text{-}Sugar\text{-}Base} \quad \text{(IV)}$$

$$\mathbf{DG\text{-}E'\text{-}R^2\text{-}A\text{-}R^1\text{-}E\text{-}P(O)(OP(OZ^1)OH)OP(OZ^2)\text{-}Sugar\text{-}Base} \quad \text{(IVa)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O),

a keto group (C(O)), an amido group (C(O)N(R^3)), an urea group (N(R^3)C(O)N(R^3)), a carbonate group (OC(O)O), or an urethane group (OC(O)N(R^3), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure,

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group.

[0104] 9. The nucleotide of claim 1, wherein the nucleotide has the general formula (V):

DG –E' – R^2 – A – R^1 – E – P(O)(OP(O_2)OP(O_2)OH)-Sugar-Base (V)

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group (N(R^3)), an phosphano group (P(R^3)), a phosphito group (P(OR^3)O), a phosphate group (P(O_2)O), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group (C(O)N(R^3)), an urea group (N(R^3)C(O)N(R^3)), a carbonate group (OC(O)O), or an urethane group (OC(O)N(R^3), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure,

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0105] 10. The nucleotide of claim 1, wherein the nucleotide has the general formula (VI):

DG –E' – R^2 – A – R^1 – E – Sugar(P(O_2)OP(O_2)OP(O_2)OH)Base (VI)

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group (N(R^3)), an phosphano group (P(R^3)), a phosphito group (P(OR^3)O), a phosphate group (P(O_2)O), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O),

a keto group (C(O)), an amido group (C(O)N(R^3)), an urea group (N(R^3)C(O)N(R^3)), a carbonate group (OC(O)O), or an urethane group (OC(O)N(R^3), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure,

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0106] 11. The nucleotide of claim 1, wherein the nucleotide has the general formula (VII):

DG –E' – R^2 – A – R^1 – E – Base -Sugar-$P(O_2)OP(O_2)OP(O_2)OH$ (VII)

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group (N(R^3)), an phosphano group (P(R^3)), a phosphito group (P(OR^3)O), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group (C(O)N(R^3)), an urea group (N(R^3)C(O)N(R^3)), a carbonate group (OC(O)O), or an urethane group (OC(O)N(R^3), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure,

R^1 and R^2 are the same or different and are is carbenyl groups,

A is a ring structure,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0107] 12. The nucleotide of claim 1, wherein the ring structure A is saturated, unsaturated or aromatic or can include a mixture of saturated, unsaturated, or aromatic rings.

[0108] 13. The nucleotide of claim 1, wherein the each ring in the ring structure includes from 3 to about 12 main group elements.

[0109] 14. The nucleotide of claim 1, wherein each carbyl group and each carbenyl group include from 1 to 40 carbon, where one or more of the carbon atoms can be replaced with a hetero atoms selected from the group consisting of B, C, Si, Ge, N, P, As, O, S, or Se and having sufficient hydrogen atoms to satisfy the valency of the group, where one or more hydrogen atoms can be

replaced with F, Cl, Br, I, OR, SR, COR, COOR, $CONH_2$, CONHR, CONRR', or any other monovalent group inert or substantially inert under the substitution/displacement reaction conditions.

[0110] 15. The nucleotide of claim 1, wherein the nucleotide has the general formula (VIII):

$$\mathbf{DG-E'-R-E-Nu} \qquad \text{(VIII)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), a phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group, and

Nu is a natural or synthetic nucleotide.

[0111] 16. The nucleotide of claim 1, wherein the nucleotide has the general formulas (IX or IXa):

$$\mathbf{DG-E'-R-E-P(O_2)OP(O_2)OP(O_2)\text{-}Sugar\text{-}Base} \qquad \text{(IX)}$$

$$\mathbf{DG-E'-R-E-P(O_2)OP(OZ^1)OP(OZ^2)\text{-}Sugar\text{-}Base} \qquad \text{(IXa)}$$

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), a phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base, and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group as described herein.

[0112] 17. The nucleotide of claim 1, wherein the nucleotide has the general formulas (X or Xa):

DG –E' – R – E – P(O)(OP(O_2)OH)OP(O_2)-Sugar-Base (X)

DG –E' – R – E – P(O)(OP(OZ^1)OH)OP(OZ^2)-Sugar-Base (Xa)

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$, R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety,

Base is a natural or synthetic nucleotide base, and

Z^1 or Z^2 are the same or different and are groups that either modify incorporation timing or enhancing detection of the detectable group as described herein.

[0113] 18. The nucleotide of claim 1, wherein the nucleotide has the general formula (XI):

DG –E' – R – E – P(O)(OP(O_2)OP(O_2)OH)-Sugar-Base (XI)

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group

($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0114] 19. The nucleotide of claim 1, wherein the nucleotide has the general formula (XII):

DG –E' – R – E – Sugar($P(O_2)OP(O_2)OP(O_2)OH$)Base (XII)

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group ($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0115] 20. The nucleotide of claim 1, wherein the nucleotide has the general formula (XIII):

DG –E' – R – E – Base -Sugar-$P(O_2)OP(O_2)OP(O_2)OH$ (XIII)

where:

DG is a detectable group,

E and E' are the same and different and are a carbon group ($C(H)_2$, $C(HR^3)$ or $C(R^3)_2$), an oxygen atom (O), a sulfur atom (S), an amino group ($N(R^3)$), an phosphano group ($P(R^3)$), a phosphito group ($P(OR^3)O$), a phosphate group ($P(O_2)O$), a polyphosphate group $(P(O_2)O)_n$ (n is an integer having a value between 3 and 12), a silyl group ($Si(R^3)_2$), a siloxyl group

($Si(OR^3)_2$), a carboxy group (C(O)O), a keto group (C(O)), an amido group ($C(O)N(R^3)$), an urea group ($N(R^3)C(O)N(R^3)$), a carbonate group (OC(O)O), or an urethane group ($OC(O)N(R^3)$), R^3 is a hydrogen atom, a carbyl group or is absent depending on the structure (*e.g.*, E' or E is a nitrogen atom doubly bonded to DG or to R^2 or a carbon atom triply bonded to DG or R^2),

R is a carbenyl group,

Sugar is a sugar moiety, and

Base is a natural or synthetic nucleotide base.

[0116] 21. The nucleotide of claim 1, wherein the each carbyl group and each carbenyl group include from 1 to 40 carbon, where one or more of the carbon atoms can be replaced with a hetero atoms selected from the group consisting of B, C, Si, Ge, N, P, As, O, S, or Se and having sufficient hydrogen atoms to satisfy the valency of the group, where one or more hydrogen atoms can be replaced with F, Cl, Br, I, OR, SR, COR, COOR, $CONH_2$, CONHR, CONRR', or any other monovalent group inert or substantially inert under the substitution/displacement reaction conditions.

[0117] 22. A method for preparing gamma phosphate modified nucleotide triphosphates comprising the steps of:

cyclizing a nucleotide triphosphate in N,N-dicyclohexylcarbodiimide (DCC) to form a cyclized nucleotide triphosphate,

contacting the cyclized nucleotide triphosphate with an α,ω-diamino linker, where the linker includes a linking group comprising a linear or branched carbenyl group or a carbenyl group including a central ring structure, to form a linker gamma phosphate modified nucleotide triphosphate, and

contacting the linker gamma phosphate modified nucleotide triphosphate with a carboxyic acid, a carboxylic acid chloride or a carboxylic acid anhydride including a detectable group having a detectable property to form a detectable group, linker gamma phosphate modified nucleotide triphosphate.

[0118] 23. A method for preparing gamma phosphate modified nucleotide triphosphates comprising the steps of:

contacting a nucleotide triphosphate with an α,ω-diamino linker, where the linker includes a linking group comprising a linear or branched carbenyl group or a carbenyl group including a central ring structure, in N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide (EDC) to form a linker gamma

phosphate modified nucleotide triphosphate, and

contacting the linker gamma phosphate modified nucleotide triphosphate with a carboxyic acid, a carboxylic acid chloride or a carboxylic acid anhydride including a detectable group having a detectable property to form a detectable group, linker gamma phosphate modified nucleotide triphosphate.

[0119] 24. A method for preparing gamma phosphate modified nucleotide triphosphates comprising the steps of:

contacting a protected linker of the general formula Q - NH - G - OH with a phosphate donor to form a phosphate terminated linker of the general formula Q - NH - G - $OP(O)(OH)_2$, where Q is a protecting group and G a linker comprising a linear or branched carbenyl group or a carbenyl group including a central ring structure,

contacting the phosphate terminated linker, Q - NH - G - $OP(O)(OH)_2$, with imidazole to form an imidazole activated phosphate terminated linker, Q - NH - G - OP(O)(OH) - imidazole,

contacting the imidazole activated phosphate terminated linker, Q - NH - G - OP(O)(OH) - imidazole with a nucleotide polyphosphate to form a protected, linker terminal phosphate modified nucleotide polyphosphate, Q - NH - G - O - $[P(O)(OH)]_n$ - O - Nuc, where n is an integer having a value between 3 and 12,

deprotecting the protected gamma phosphate functionalized nucleotide polyphosphate, Q - NH - G - O - $[P(O)(OH)]_n$ - O - Nuc, to form an unprotected linker, gamma phosphate modified nucleotide polyphosphate, H_2N - G - O - $[P(O)(OH)]_n$ - O - Nuc, and

contacting the unprotected linker, gamma phosphate modified nucleotide polyphosphate, H_2N - G - O - $[P(O)(OH)]_n$ - O - Nuc, with a carboxyic acid, a carboxylic acid chloride or a carboxylic acid anhydride including a detectable group (DG) having a detectable property to form a detectable group gamma phosphate modified nucleotide triphosphate, DG - HN - G - O - $[P(O)(OH)]_n$ - O - Nuc.

[0120] 25. A method for preparing gamma phosphate modified nucleotide triphosphates comprising the steps of:

contacting a linker of the general formula Q - NH - G - OH with a sulfonate donor to form a sulfonate terminated linker, Q - NH - G - $OS(O)_2CH_3$,

contacting the sulfonate terminated linker, Q - NH - G- $OS(O)_2CH_3$, with phosphate donor to form a phosphate terminated linker, Q - NH - G - $OP(O)(OH)_2$,

contacting the phosphate terminated linker, Q - NH - G - $OP(O)(OH)_2$, with imidazole to form

an activated phosphate terminated linker, Q - NH - G - OP(O)(OH) - Imidazole,

contacting the imidazole activated phosphate terminated linker, Q - NH - G - OP(O)(OH) - Imidazole, with a nucleotide polyphosphate to form a protected linker, terminal phosphate modified nucleotide polyphosphate, Q - NH - G - O - $[P(O)(OH)]_n$ - O - Nuc, where n is an integer having a value between 3 and 12,

deprotecting the protected gamma phosphate functionalized nucleotide polyphosphate, Q - NH - G - O - $[P(O)(OH)]_n$ - O -Nuc, to form an unprotected linker, terminal phosphate modified nucleotide polyphosphate, H_2N - G - O - $[P(O)(OH)]_n$ - O - Nuc, and

contacting the unprotected linker, terminal phosphate modified nucleotide polyphosphate, H_2N - G - O - $[P(O)(OH)]_n$ - O -Nuc, with a carboxyic acid, a carboxylic acid chloride or a carboxylic acid anhydride including a detectable group having a detectable property to form a detectable group, linker, terminal phosphate modified nucleotide triphosphate, DG - HN - G - O - $[P(O)(OH)]_n$ - O - Nuc.

[0121] 26. A method for preparing gamma phosphate modified nucleotide triphosphates comprising the steps of:

contacting a linker of the general formula H_2N - G - OH with trifluoro acetic acid (TFA) to form a TFA terminated linker, TFA - NH - G - OH,

contacting the TFA terminated linker, TFA - NH - G - OH, with a cyclized nucleotide triphosphate in the presence of a base to form a TFA terminated linker, gamma phosphate modified nucleotide triphosphate, TFA - NH - G - $[P(O)(OH)]_3$ - O -Nuc,

deprotecting the TFA terminated linker, gamma phosphate modified nucleotide triphosphate, TFA - NH - G - $[P(O)(OH)]_3$ - O -Nuc, to form a linker, gamma phosphate modified nucleotide triphosphate, H_2N - G - $[P(O)(OH)]_3$ - O -Nuc, and

contacting the a linker, gamma phosphate modified nucleotide triphosphate, H_2N - G - $[P(O)(OH)]_3$ - O -Nuc, with a carboxyic acid, a carboxylic acid chloride or a carboxylic acid anhydride including a detectable group having a detectable property to form a detectable group, linker, terminal phosphate modified nucleotide triphosphate, DG - HN - G - O - $[P(O)(OH)]_3$ - O - Nuc.

[0122] 27. A method for preparing gamma phosphate modified nucleotide triphosphates comprising the steps of:

contacting a nucleotide diphosphate salt, Nuc - O - $P(O)(O^-)$ - O - $P(O)(O^-)_2M_3$, with a tetracarbyl ammonium salt, $R_4N^+X^-$, to form a nucleotide diphosphate tetracarbyl ammonium salt,

Nuc - O - P(O)(O^-) - O - P(O)(O^-)$_2$(R_4N^+)$_3$,

contacting an N-TFA-protected, α-amino, ω-hydroxy linker, TFA-N(H)-G-OH, with sufficient $POCl_3$ to form an N-TFA-protected, α-amino, ω-dichlorophosphite linker, TFA-N(H)-G-OP(O)Cl_2,

contacting the nucleotide diphosphate tetracarbyl ammonium salt, Nuc - O - P(O)(O^-) - O - P(O)(O^-)$_2$(R_4N^+)$_3$, with the N-TFA-protected, α-amino, ω-dichlorophosphite linker, TFA-N(H)-G-OP(O)Cl_2, to form a TFA-protected, α-amino, ω-gamma phosphate modified nucleotide triphosphate tetracarbyl ammonium salt, TFA - NH - G -O - [P(O)(O^-)]$_3$ - O -Nuc((R_4N^+)$_3$,

deprotecting the TFA-protected, α-amino, ω-gamma phosphate modified nucleotide triphosphate tetracarbyl ammonium salt, TFA - NH - G -O - [P(O)(O^-)]$_3$ - O -Nuc((R_4N^+)$_3$, to form a linker, gamma phosphate modified nucleotide triphosphate, α-amino, ω-gamma phosphate modified nucleotide triphosphate tetracarbyl ammonium salt, H_2N - G -O - [P(O)(O^-)]$_3$ - O -Nuc((R_4N^+)$_3$.

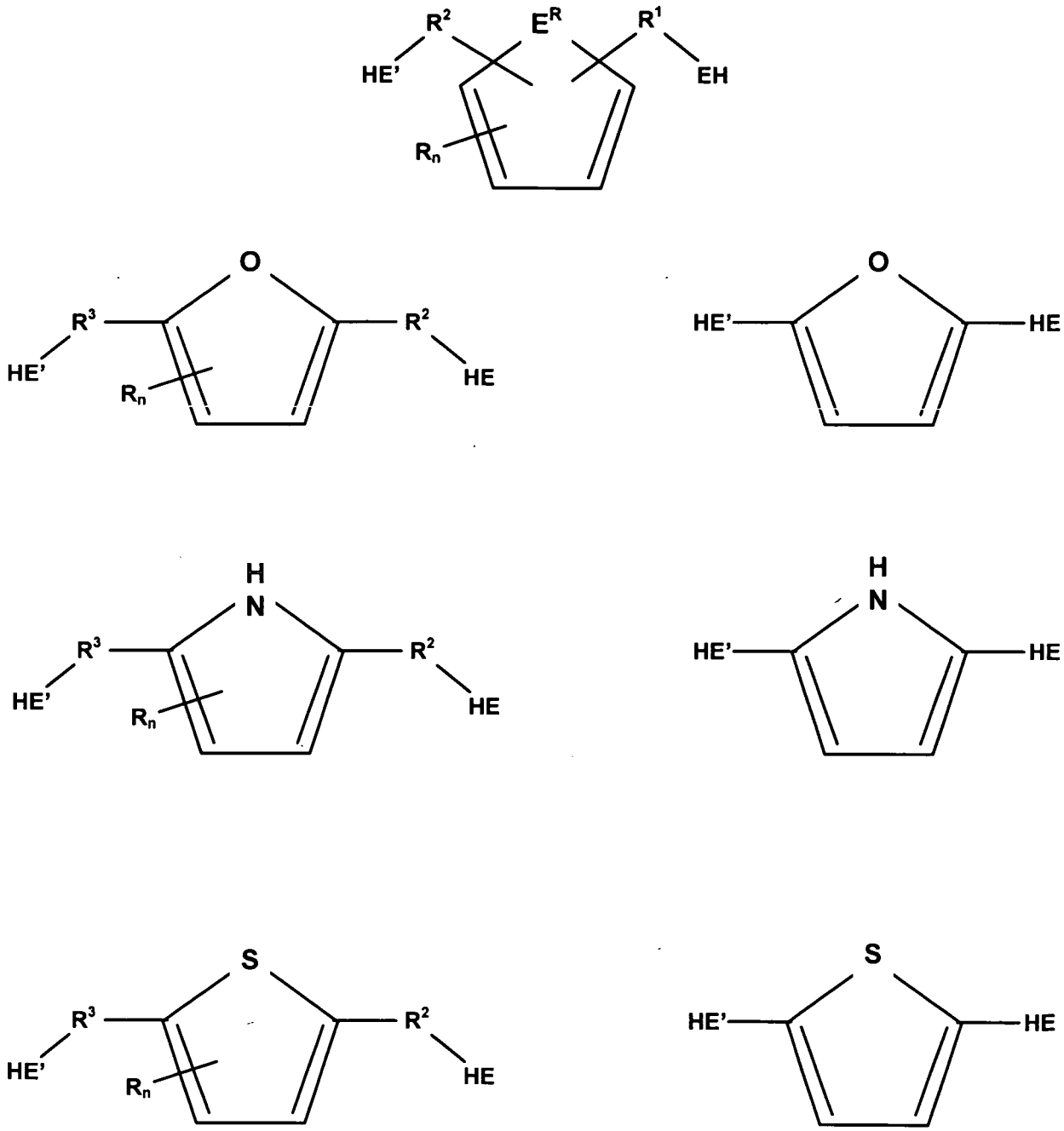

R2
ER
R1
HE'
EH
Rn
O
R3
R2
HE'
HE
Rn
O
HE'
HE
H
N
R3
R2
HE'
HE
Rn
H
N
HE'
HE
S
R3
R2
HE'
HE
Rn
S
HE'
HE

FIG. 1

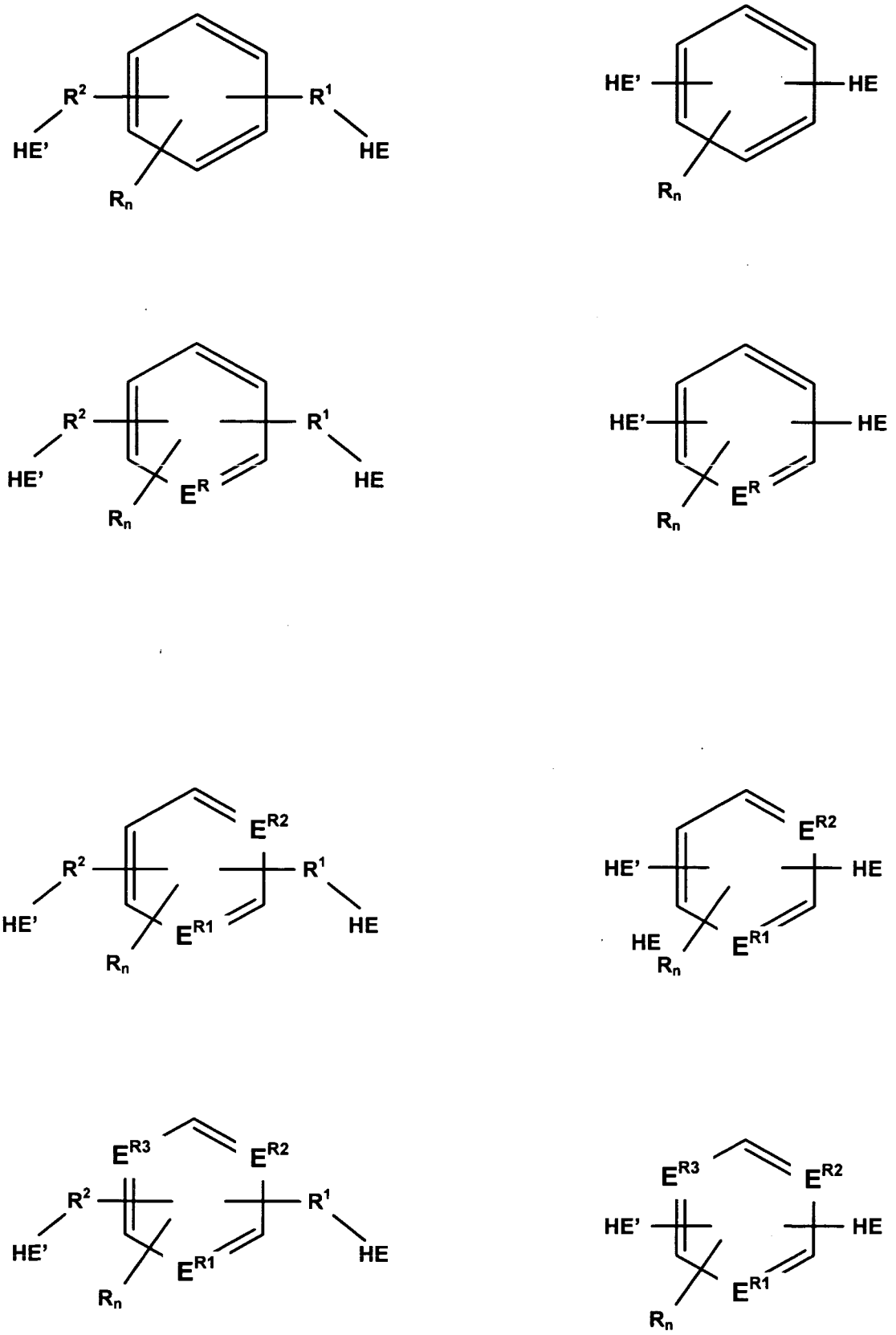

R2
R1
HE'
HE
Rn
HE'
HE
Rn
R2
R1
HE'
HE
ER
Rn
HE'
HE
ER
Rn
ER2
R2
R1
HE'
HE
ER1
Rn
ER2
HE'
HE
HE
ER1
Rn
ER3
ER2
R2
R1
HE'
HE
ER1
Rn
ER3
ER2
HE'
HE
ER1
Rn

FIG. 2

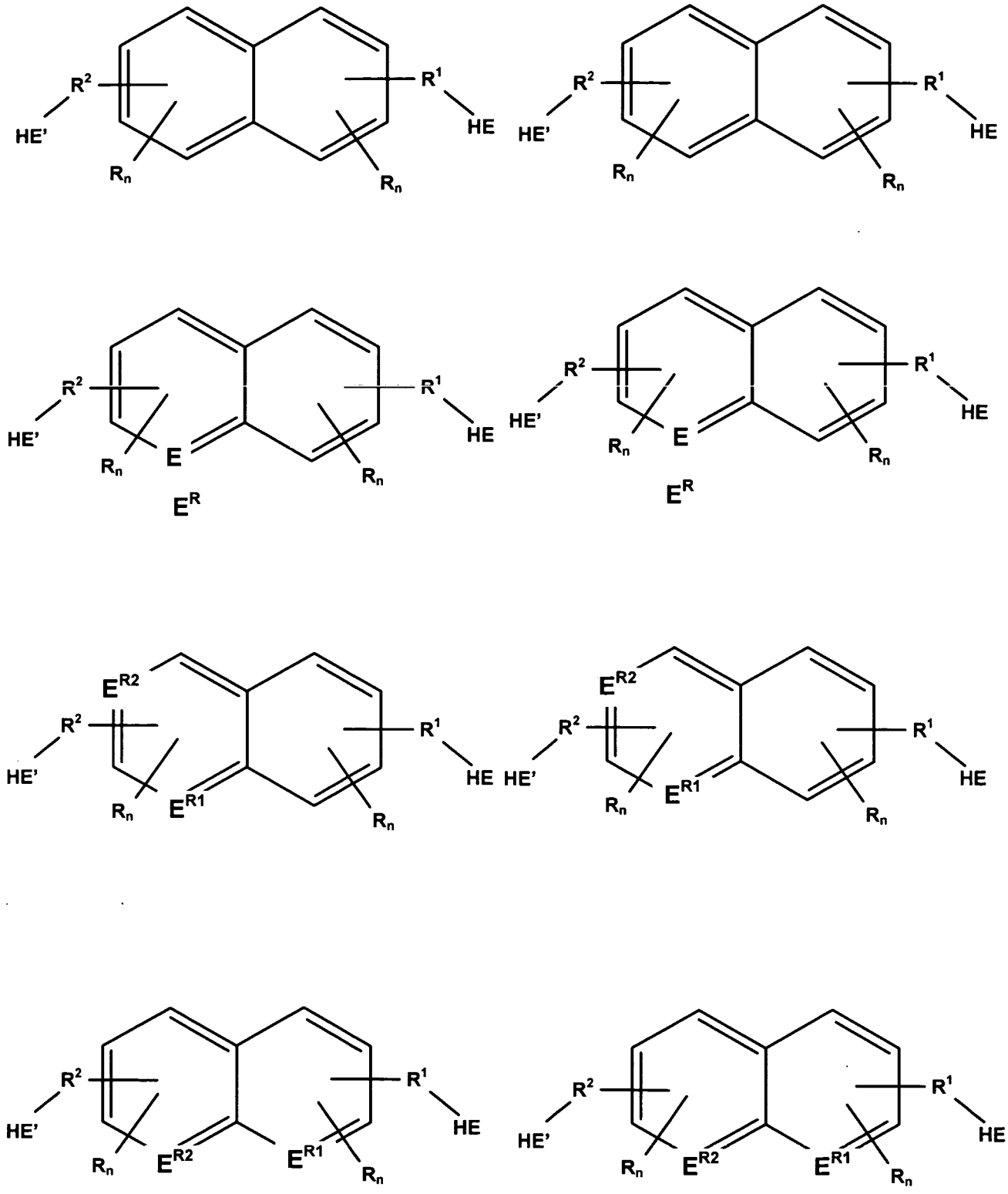


FIG. 3

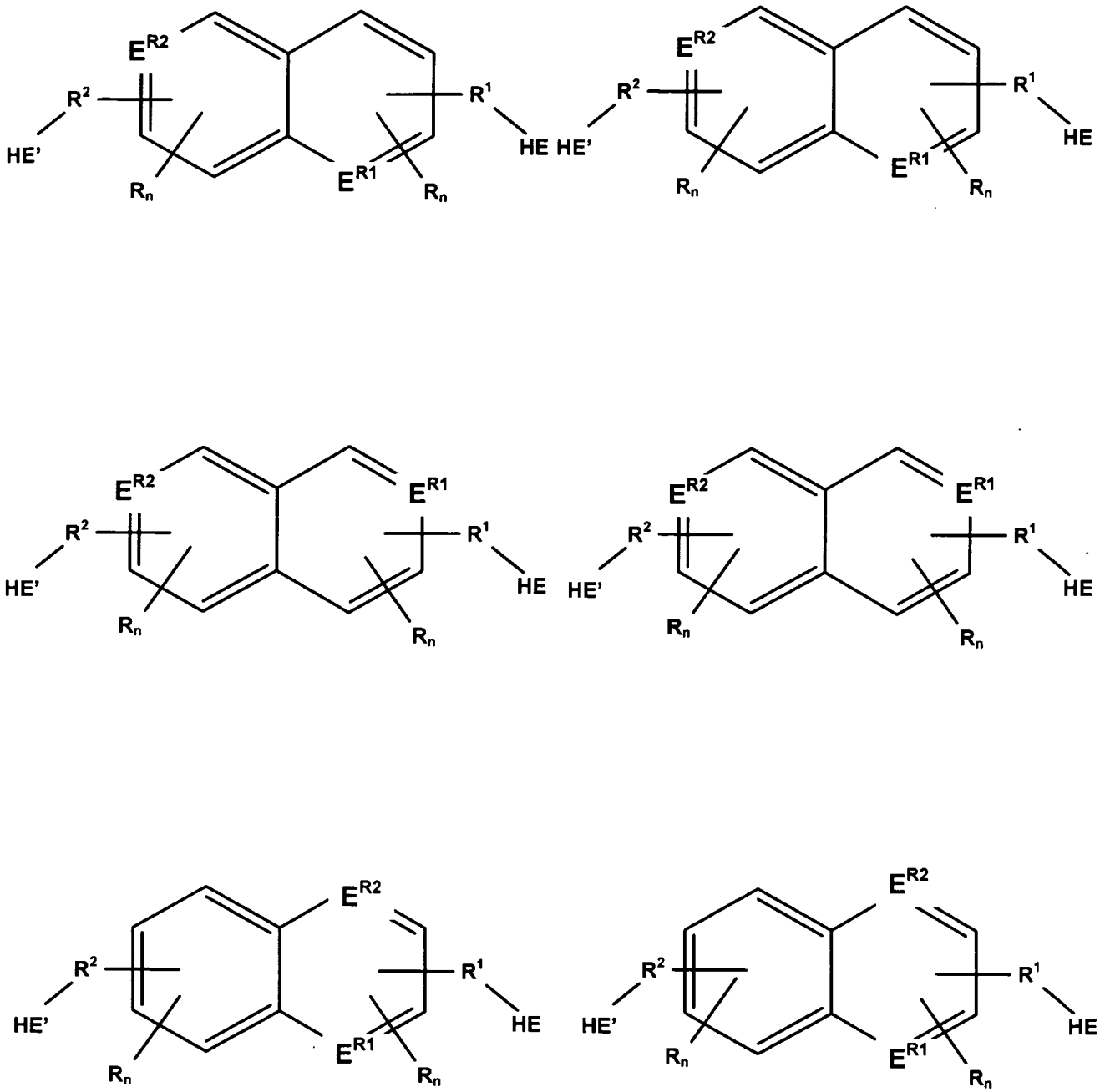


FIG. 4

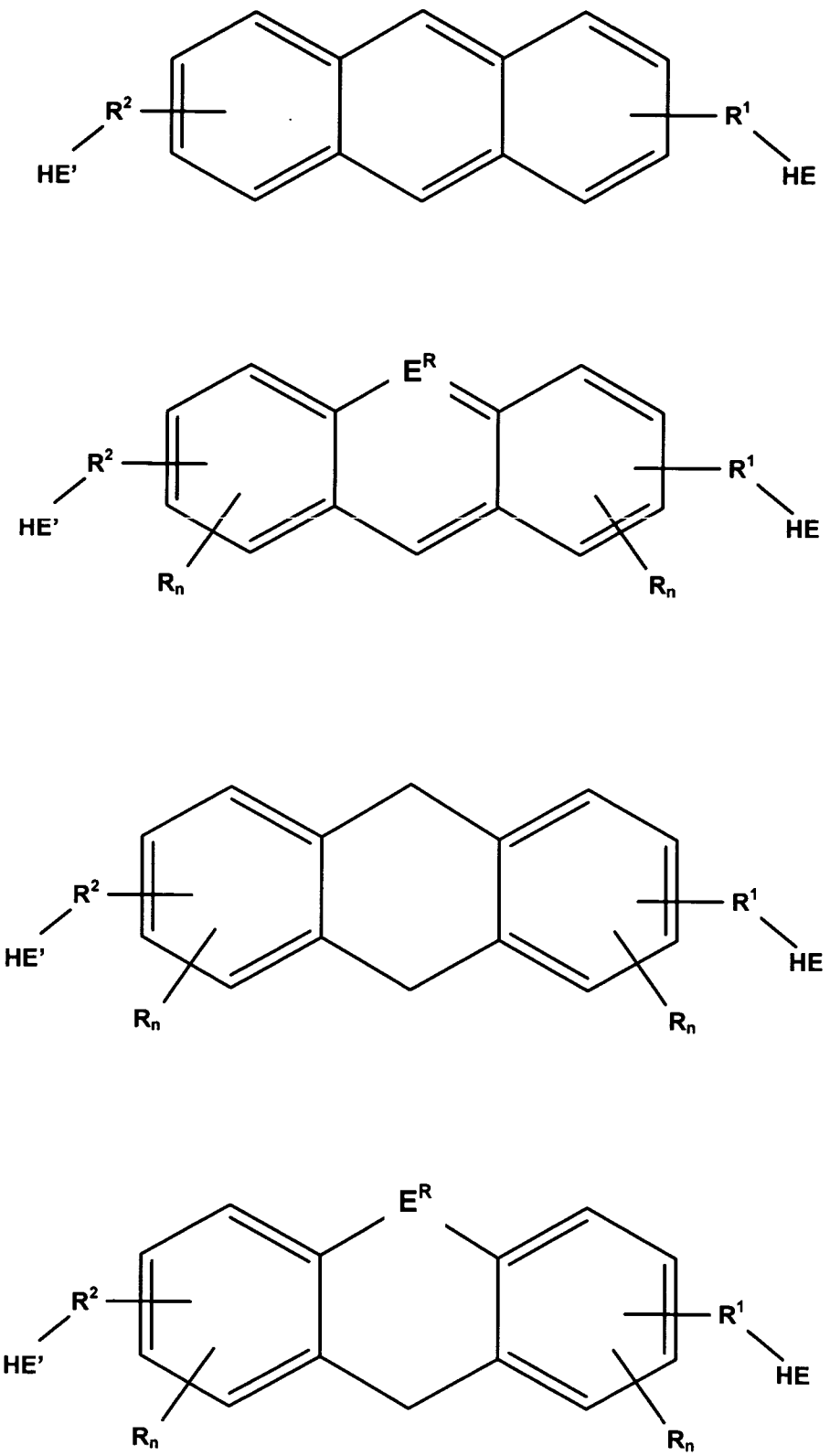


FIG. 5

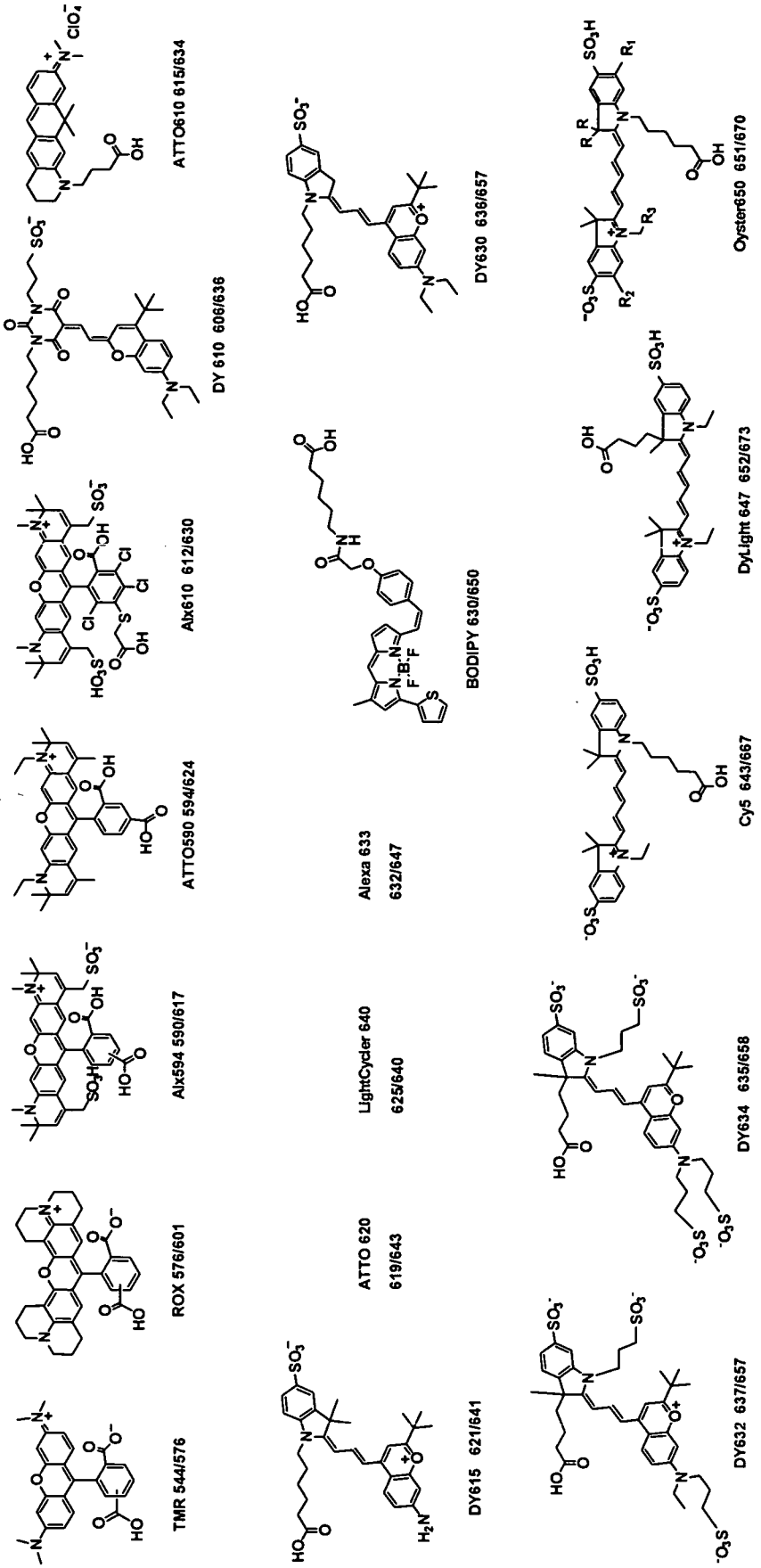

TMR 544/576
ROX 576/601
Alx594 590/617
ATTO590 594/624
Abx610 612/630
DY 610 606/636
ATTO610 615/634
DY615 621/641
ATTO 620
619/643
LightCycler 640
625/640
Alexa 633
632/647
BODIPY 630/650
DY630 636/657
DY632 637/657
DY634 635/658
Alexa 647
650/665
HiLyte 647
650/675
DY 647
653/672
ATTO 655
663/684
Alexa 660
663/690
HiLyte 680
678/699
Cy5 643/667
DyLight 647 652/673
Oyster650 651/670
DY 680 690/709

FIG. 6

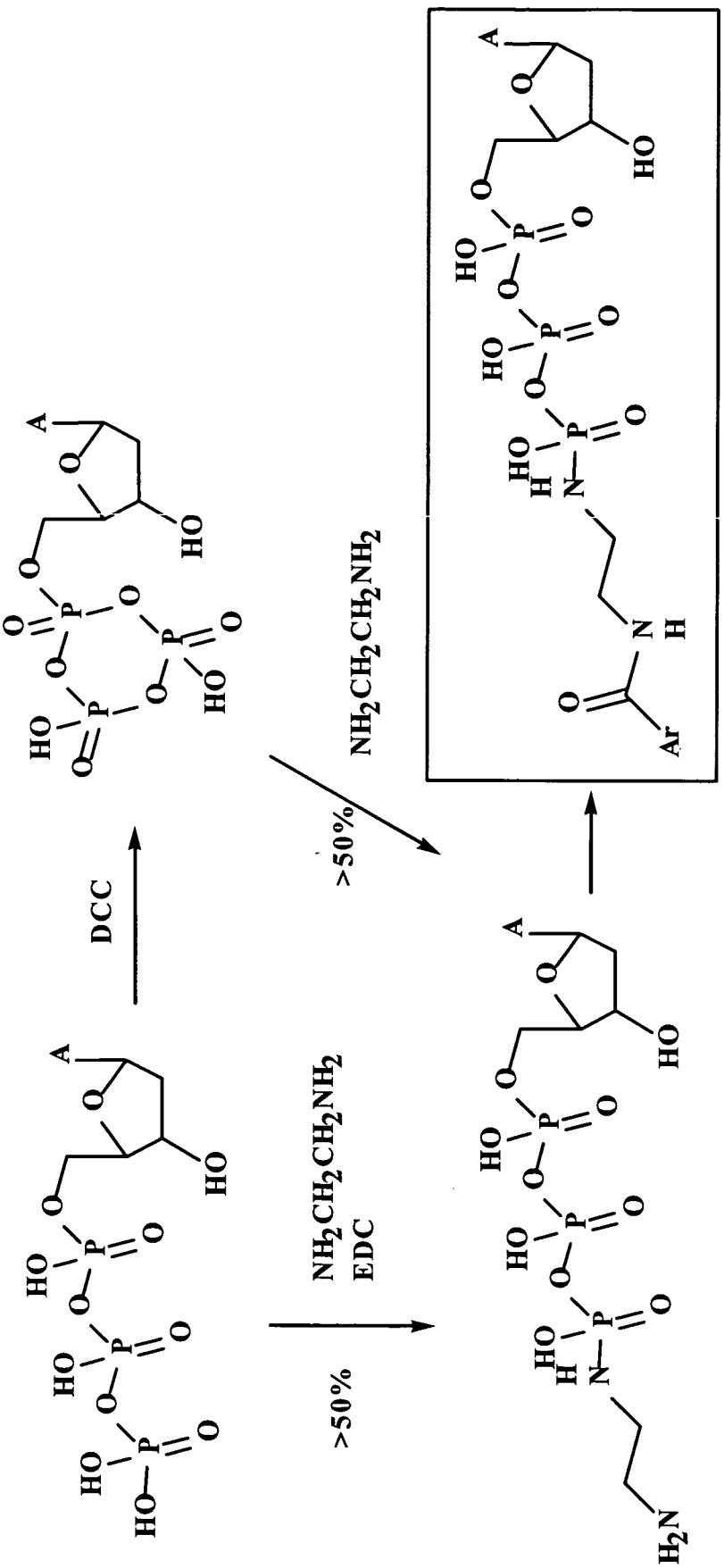


FIG. 7

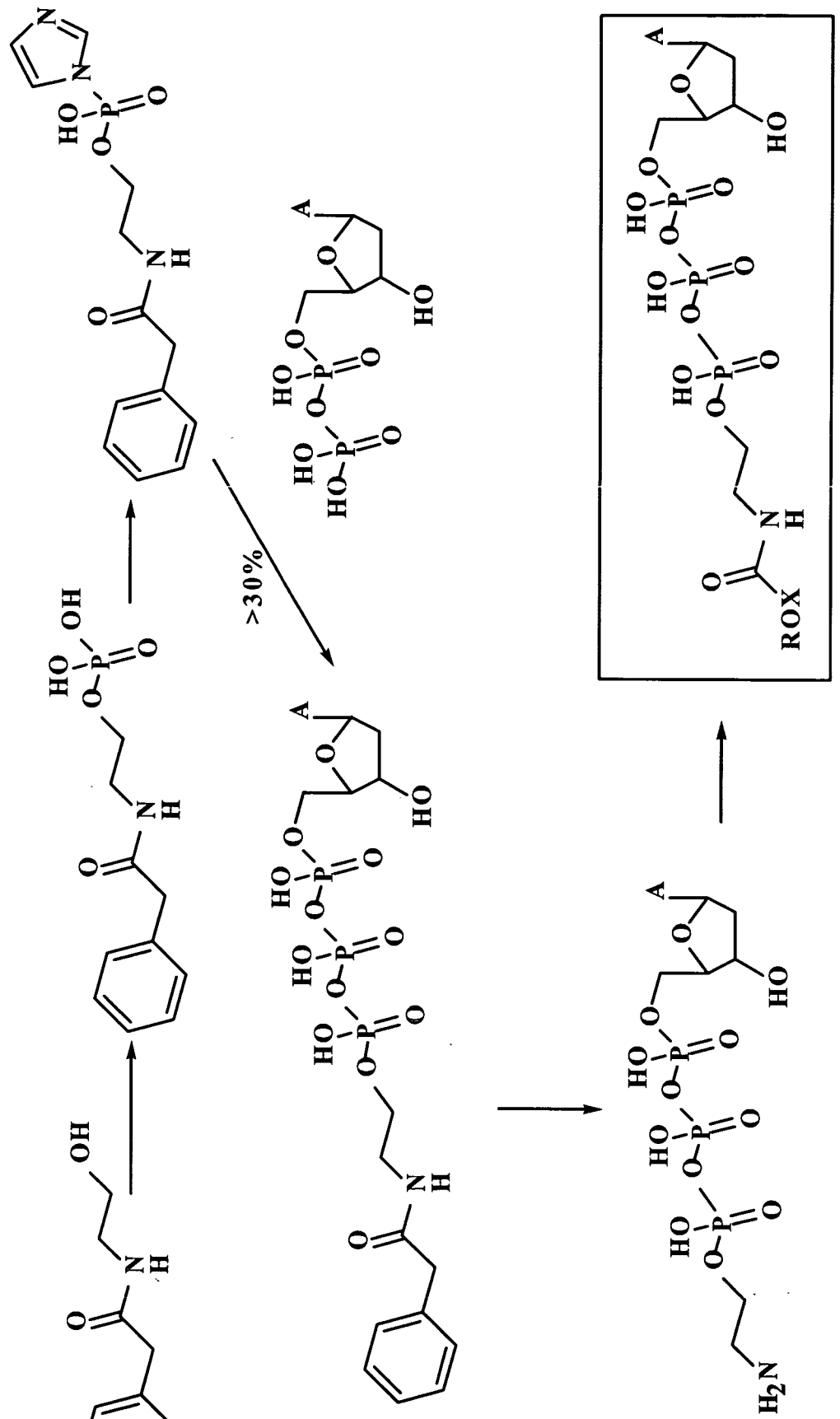


FIG. 8

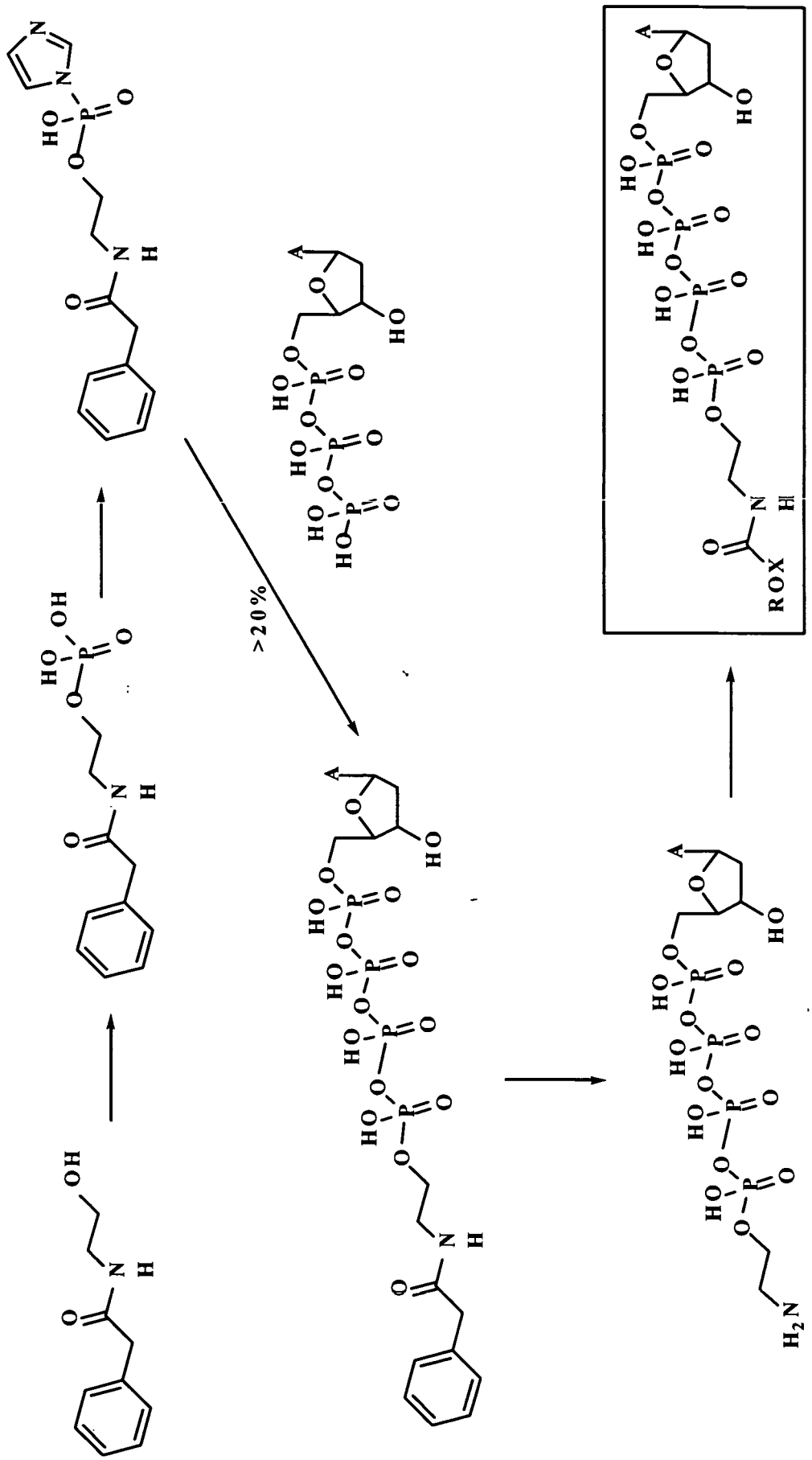


FIG. 9

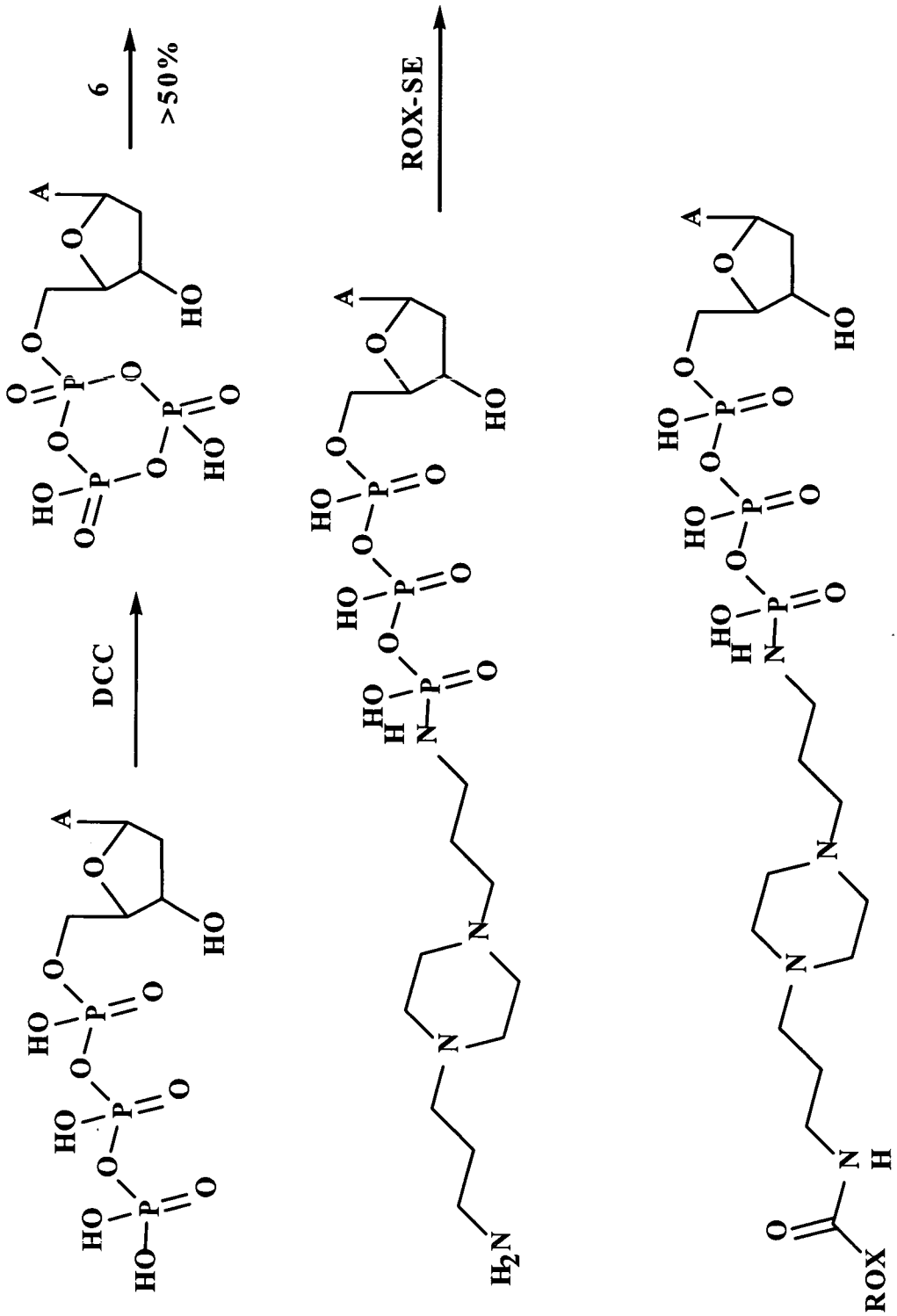


FIG. 10

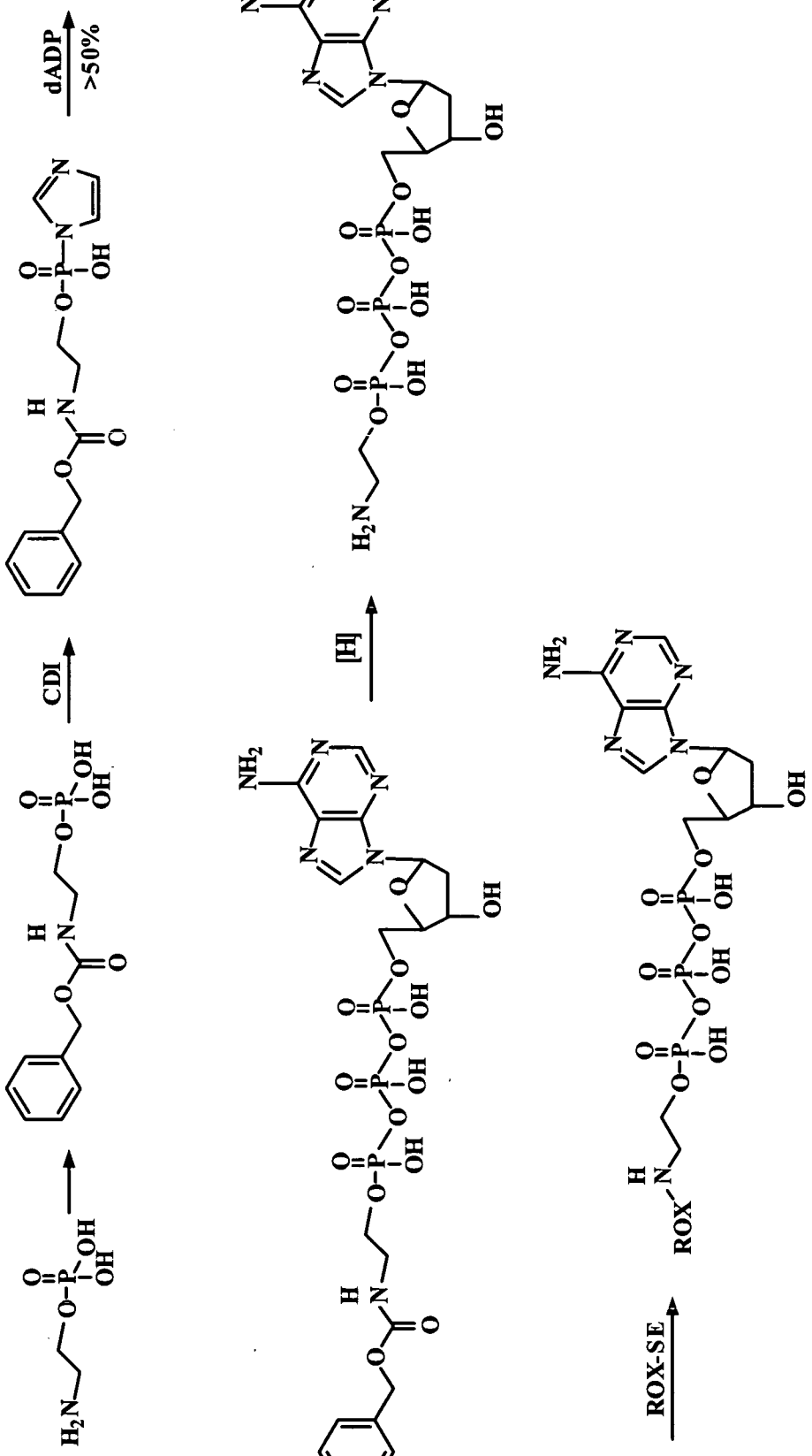


FIG. 11

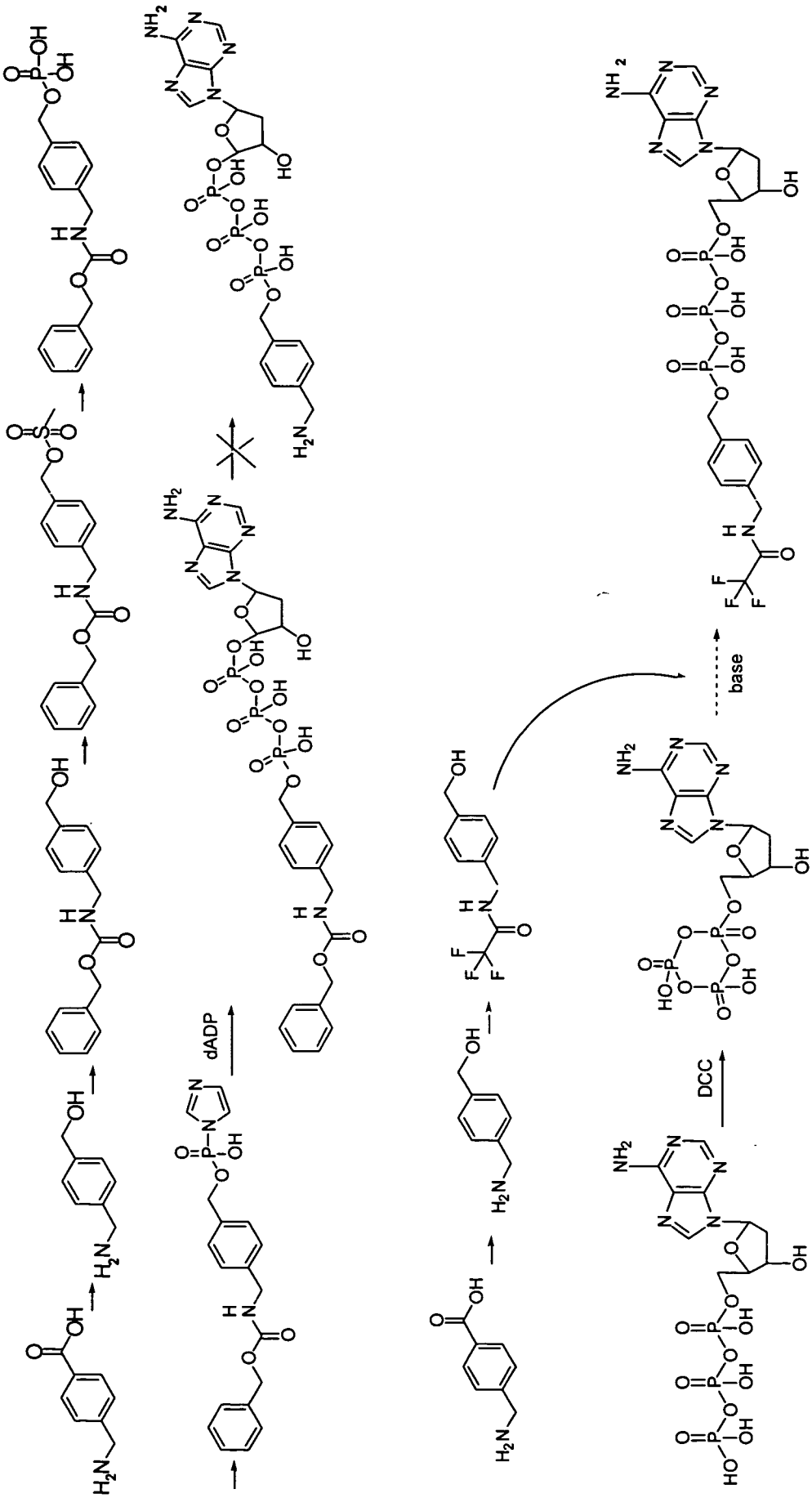


FIG. 12

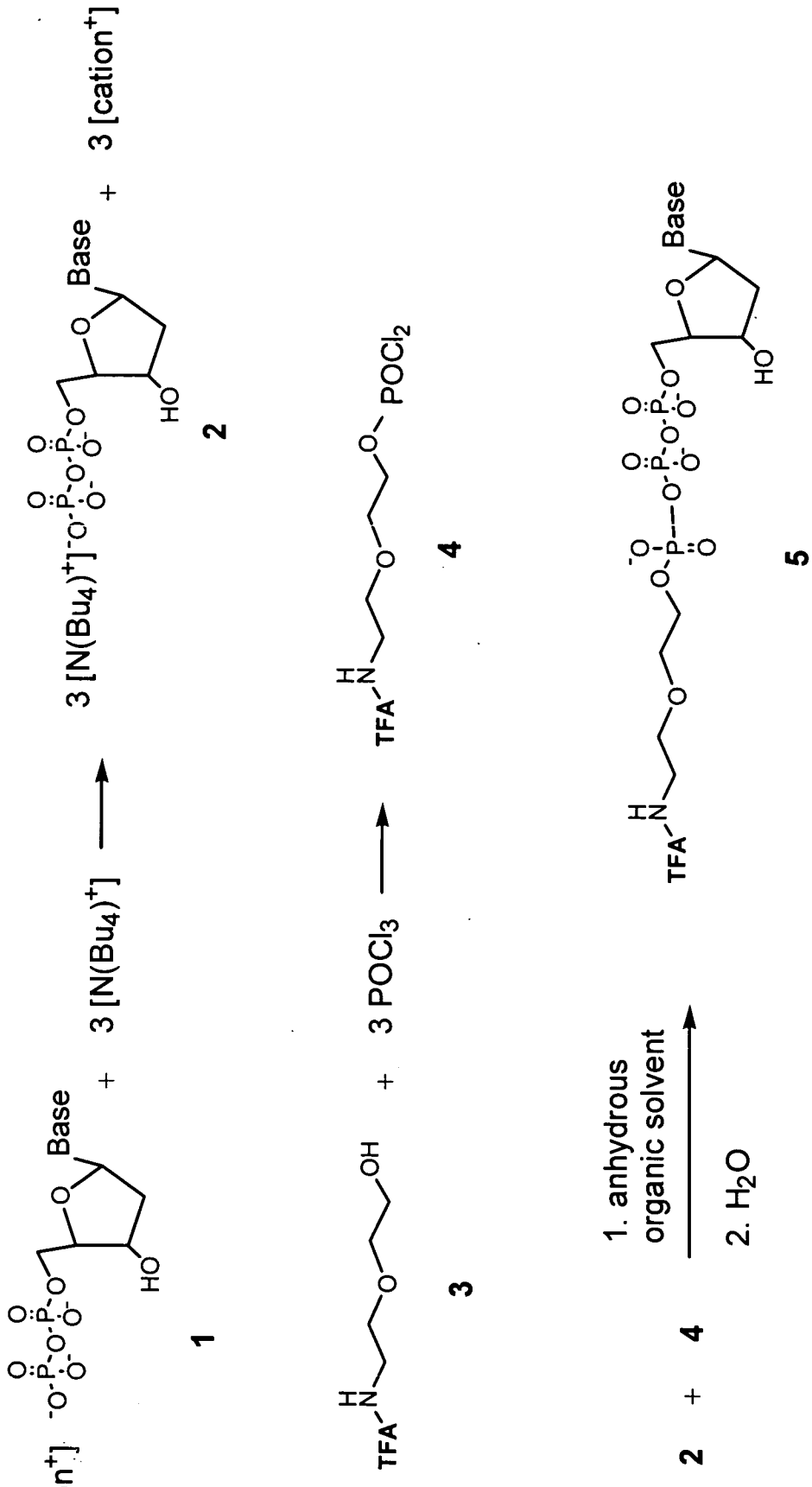

3 [cation+] -O-P(=O)(-O-)-O-P(=O)(-O-)-O-P(=O)(-O-)-O-Base
1
+ 3 [N(Bu4)+]
3 [N(Bu4)+] -O-P(=O)(-O-)-O-P(=O)(-O-)-O-P(=O)(-O-)-O-Base
2
+ 3 [cation+]
TFA-H2N-CH2-CH2-O-CH2-CH2-OH
3
+ 3 POCl3
TFA-H2N-CH2-CH2-O-CH2-CH2-O-POCl2
4
1. anhydrous organic solvent
2. H2O
2 + 4
TFA-H2N-CH2-CH2-O-CH2-CH2-O-P(=O)(-O-)-O-P(=O)(-O-)-O-P(=O)(-O-)-O-Base
5

FIG. 13