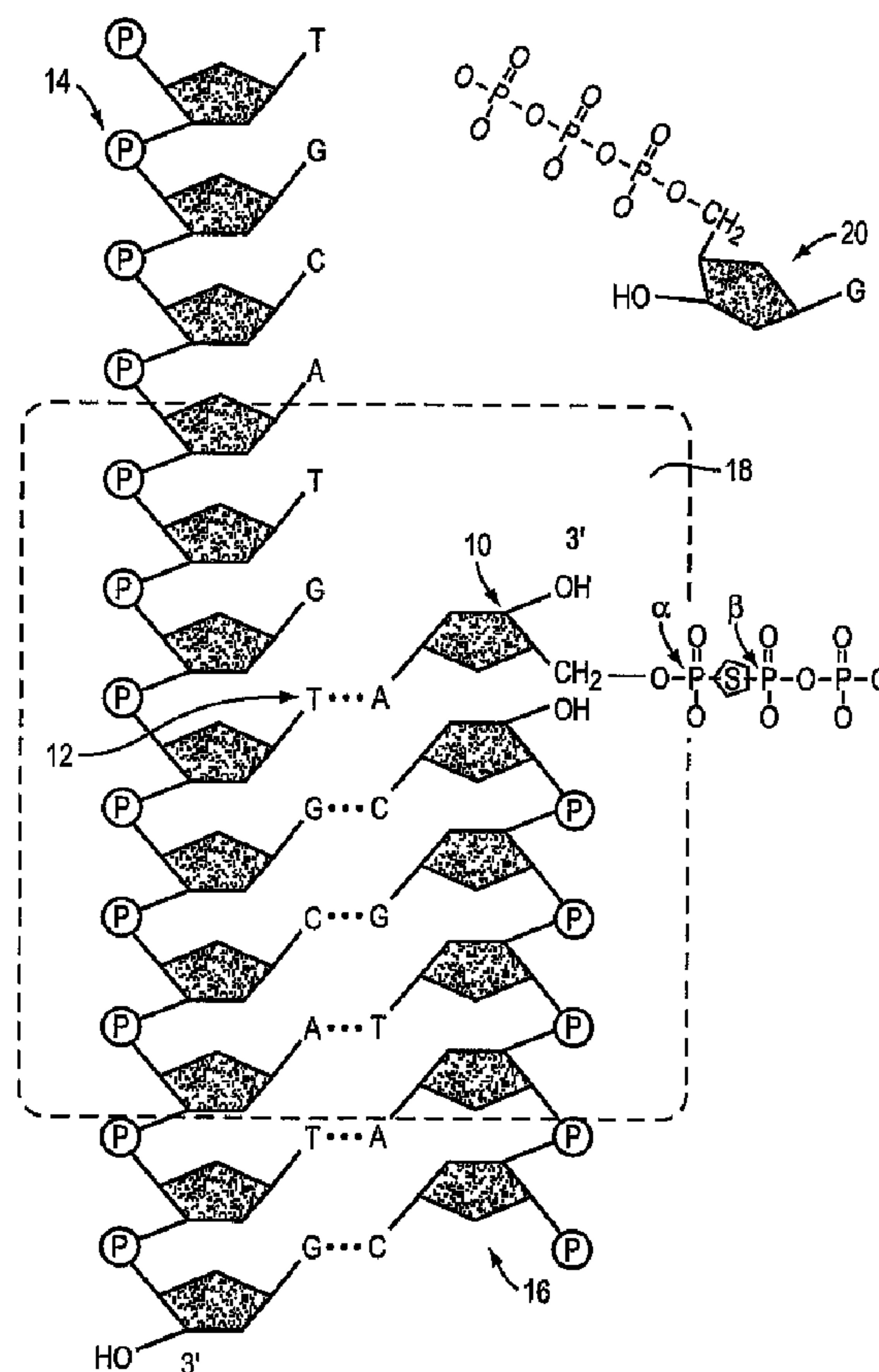




(86) Date de dépôt PCT/PCT Filing Date: 2006/01/30  
(87) Date publication PCT/PCT Publication Date: 2006/08/03  
(85) Entrée phase nationale/National Entry: 2007/08/24  
(86) N° demande PCT/PCT Application No.: US 2006/004638  
(87) N° publication PCT/PCT Publication No.: 2006/081583  
(30) Priorité/Priority: 2005/01/28 (US11/046,448)

(51) Cl.Int./Int.Cl. *C12Q 1/68* (2006.01)  
(71) Demandeur/Applicant:  
HELICOS BIOSCIENCES CORPORATION, US  
(72) Inventeur/Inventor:  
BUZBY, PHILIP RICHARD, US  
(74) Agent: ROBIC

(54) Titre : PROCÉDES ET COMPOSITIONS POUR AMÉLIORER LA FIDÉLITÉ D'UNE RÉACTION DE SYNTHÈSE D'ACIDE NUCLEIQUE  
(54) Title: METHODS AND COMPOSITIONS FOR IMPROVING FIDELITY IN A NUCLEIC ACID SYNTHESIS REACTION



(57) **Abrégé/Abstract:**

The invention provides methods and compositions for improving the fidelity of a sequencing-by-synthesis reaction by using a nucleotide derivative that forms a hydrogen bond with a complementary nucleotide on a template, but fails to form a



(57) **Abrégé(suite)/Abstract(continued):**

phosphodiester bond with the 3' hydroxyl group of a primer under conditions otherwise suitable for a polymerization reaction; thereby blocking incorporation of a mismatched nucleotide.

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

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 August 2006 (03.08.2006)

PCT

(10) International Publication Number  
**WO 2006/081583 A3**

(51) International Patent Classification:  
*C12Q 1/68* (2006.01)

(21) International Application Number:  
PCT/US2006/004638

(22) International Filing Date: 30 January 2006 (30.01.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
11/046,448 28 January 2005 (28.01.2005) US

(71) Applicant (for all designated States except US): **HELI-COS BIOSCIENCES CORPORATION** [US/US]; ONE KENDALL SQUARE, Building 200, Cambridge, Massachusetts 02139 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **BUZBY, Philip, Richard** [US/US]; 31 Harlan Drive, Brockton, Massachusetts 02301 (US).

(74) Agents: **GREENHALGH, Duncan, A. et al.**; Goodwin Procter LLP, Exchange Place, Boston, MA 02109 (US).

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

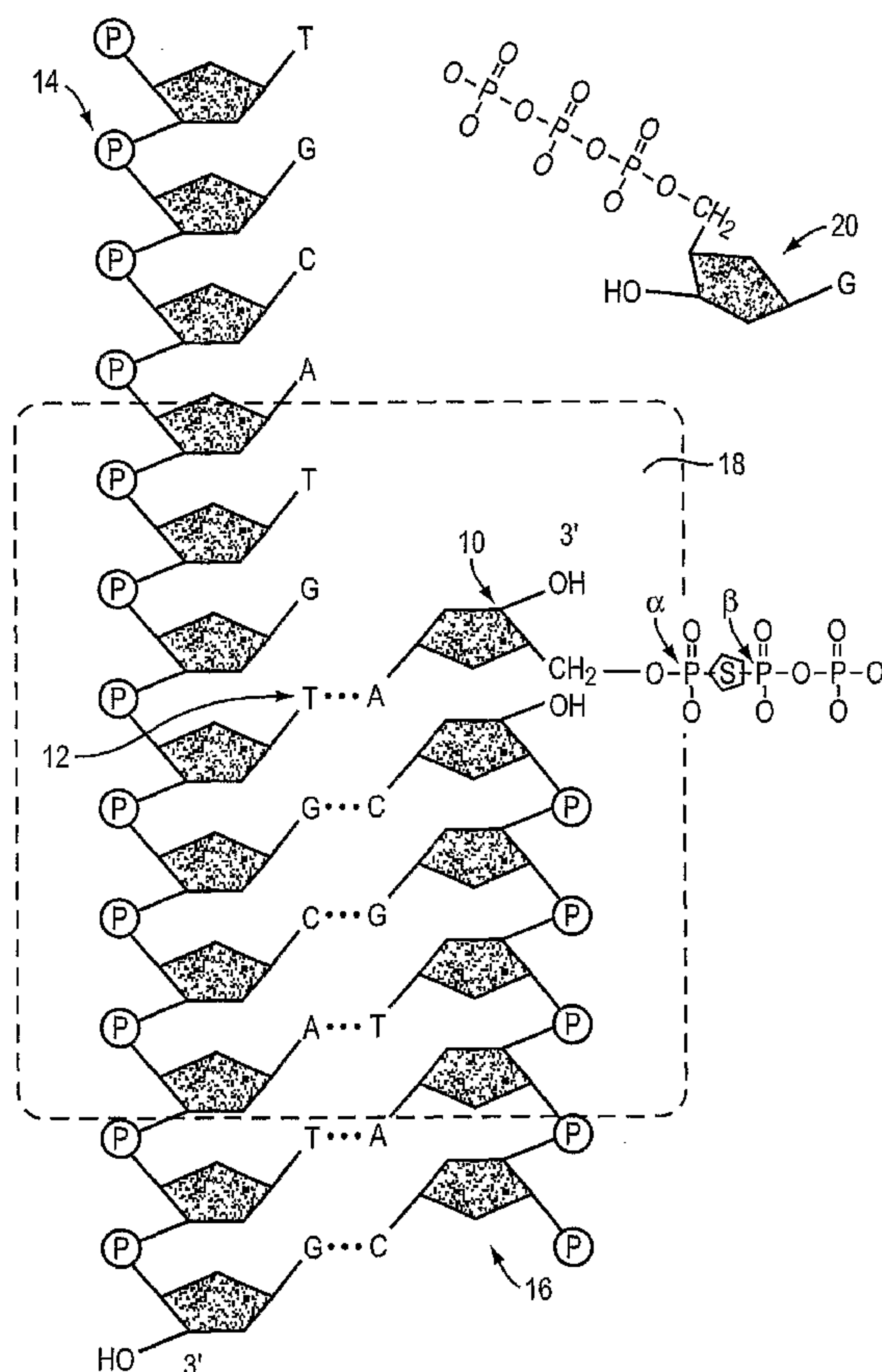
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, 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:

— with international search report

[Continued on next page]

(54) Title: METHODS AND COMPOSITIONS FOR IMPROVING FIDELITY IN A NUCLEIC ACID SYNTHESIS REACTION



(57) Abstract: The invention provides methods and compositions for improving the fidelity of a sequencing-by-synthesis reaction by using a nucleotide derivative that forms a hydrogen bond with a complementary nucleotide on a template, but fails to form a phosphodiester bond with the 3' hydroxyl group of a primer under conditions otherwise suitable for a polymerization reaction; thereby blocking incorporation of a mismatched nucleotide.

**WO 2006/081583 A3**



**(88) Date of publication of the international search report:**

19 October 2006

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



**METHODS AND COMPOSITIONS FOR IMPROVING FIDELITY  
IN A NUCLEIC ACID SYNTHESIS REACTION**

5

**Technical Field of the Invention**

[0001] The invention generally relates to methods for improving fidelity in a nucleic acid synthesis reaction, and more particularly to methods for sequencing a nucleic acid using a nucleotide derivative that decreases the rate of misincorporation of nucleotides in a nucleic acid synthesis reaction.

10

**Background of the Invention**

[0002] In vitro nucleic acid synthesis is a foundation of many fundamental research and diagnostic tools, such as nucleic acid amplification and sequencing. In a template-dependent nucleic acid synthesis reaction, the sequential addition of nucleotides is catalyzed by a nucleic acid polymerase. Depending on the template and the nature of the reaction, the nucleic acid polymerase may be a DNA polymerase, an RNA polymerase, or a reverse transcriptase. DNA synthesis requires a 3' hydroxyl on the primer for addition at the  $\alpha$ -phosphate of a free deoxynucleotide triphosphate (dNTP). A phosphodiester bond between the nucleotide on the 3' terminus of the primer and the free nucleotide is formed and diphosphate is released. This reaction is shown schematically in Figure 1.

15

20

[0003] The fidelity of template-dependent nucleic acid synthesis depends in part on the ability of the polymerase to discriminate between complementary and non-complementary nucleotides. Normally, the conformation of the polymerase enzyme favors incorporation of the complementary nucleotide. However, there is still an identifiable rate of misincorporation that depends upon factors such as local sequence and the base to be incorporated.

25

[0004] Synthetic or modified nucleotides and analogs, such as labeled nucleotides, tend to be incorporated into a primer less efficiently than naturally-occurring dNTPs. The reduced efficiency with which the unconventional nucleotides are incorporated by the polymerase can adversely affect the performance of sequencing techniques that depend upon faithful incorporation of such unconventional nucleotides.

30

[0005] Single molecule sequencing techniques allow the evaluation of individual nucleic acid molecules in order to identify changes and/or differences affecting genomic function. In single molecule techniques, individual, optically-resolvable nucleic acid fragments are attached

to a solid support, and sequencing is conducted on the individual strands. Sequencing events are detected and correlated to the individual strands. See Braslavsky et al., Proc. Natl. Acad. Sci., 100: 3960-64 (2003), incorporated by reference herein. Because single molecule techniques do not rely on ensemble averaging as do bulk techniques, errors due to misincorporation can have a significant deleterious effect on the sequencing results. The incorporation of a nucleotide that is incorrectly paired, under standard Watson and Crick base-pairing, with a corresponding template nucleotide during primer extension may result in sequencing errors. The presence of misincorporated nucleotides may also result in prematurely terminated strand synthesis, reducing the number of template strands for future rounds of synthesis, and thus reducing the efficiency of sequencing.

[0006] There is, therefore, a need in the art for improved methods for reducing the frequency of misincorporation and improving fidelity in nucleic acid synthesis reactions, especially in single molecule sequencing.

### **Summary of the Invention**

[0007] The invention addresses the problem of misincorporations in nucleic acid synthesis reactions. The invention improves the fidelity of nucleic acid synthesis reactions by selectively blocking incorporation of non-complementary nucleotides in a template-dependent sequencing-by-synthesis reaction. In order to block misincorporation, nucleotide derivatives that hydrogen bond to their complement on the template, but do not engage in phosphodiester bonding with adjacent primer nucleotides, are added to the reaction along with labeled standard nucleotides for incorporation into the primer. A standard (labeled) nucleotide typically will out-compete derivatives for complementary binding. In cases in which a derivative attaches at an incorporation site intended for the complementary standard nucleotide, the derivative will be washed out and the position will be occupied by a standard nucleotide in a subsequent round. Because a complementary derivative will typically out-compete a mismatched standard nucleotide for incorporation, the derivatives end up blocking misincorporation. Derivatives, which are only hydrogen-bonded to their complement, are labile in comparison to standard nucleotides, which engage in phosphodiester bonding to adjacent primer nucleotides (in addition to complement hydrogen bonding). Thus derivatives can be washed out, making their complement available for binding in subsequent nucleotide addition cycle.

[0008] Methods of the invention prevent some or all of the misincorporation errors that result in erroneous base calling and/or chain termination in sequencing reactions. According to



the invention, a polymerization reaction is conducted on a nucleic acid duplex that comprises a primer hybridized to a template nucleic acid. The reaction is conducted in the presence of a polymerase, at least one free dNTP corresponding to a first nucleotide species, and at least one nucleotide derivative corresponding to a nucleotide species different from the dNTP. The nucleotide derivative comprises a modification that inhibits the formation of a phosphodiester bond between the derivative and the free 3' hydroxyl group of the terminal nucleotide on the primer. As such, the nucleotide derivative cannot incorporate into the primer. However, if the template comprises a nucleotide at a position immediately adjacent to the 3' terminus of the primer that is complementary to the nucleotide derivative, the nucleotide derivative forms hydrogen bonds with its complementary nucleotide and prevents the incorporation of the free nucleotide into the primer. According to the invention, derivatives that are complementary to an available template nucleotide out-compete misincorporating dNTPs for template binding.

[0009] In a single molecule sequencing-by-synthesis reaction, primer/template duplexes are bound to a solid support such that each duplex is individually optically detectable.

According to the invention, a primer/template duplex is exposed to a polymerase, a labeled nucleotide corresponding to at least a first nucleotide species, and at least one nucleotide derivative corresponding to a nucleotide species different than the labeled species. The duplex may be simultaneously exposed to the polymerase, labeled nucleotide and nucleotide derivative; or it may be first exposed to the nucleotide derivative and then to the polymerase and labeled nucleotide. In a preferred aspect, the duplex is simultaneously exposed to the polymerase, the free nucleotide and three nucleotide derivatives, one corresponding to each of the three remaining nucleotide species. Unincorporated labeled nucleotides are washed away, and because hydrogen bonds are relatively weak, the nucleotide derivatives are removed along with the unincorporated labeled nucleotides. The incorporation of the labeled nucleotide is determined, as well the identity of the nucleotide that is complementary to a nucleotide on the template at a position that is opposite the incorporated nucleotide. The polymerization reaction is serially repeated in the presence of labeled nucleotide that corresponds to each of the other nucleotide species, and appropriate nucleotide derivatives, until a sequence of incorporated nucleotides is compiled from which the sequence of the template nucleic acid can be determined.

[0010] Single molecule sequencing methods of the invention preferably comprise template/primer duplex attached to a surface. Individual nucleotides added to the surface comprise a detectable label – preferably a fluorescent label. Each nucleotide species can comprise a different label, or they can comprise the same label. In a preferred embodiment, each



duplex is individually optically resolvable in order to facilitate single molecule sequence discrimination. The choice of a surface for attachment of duplex depends upon the detection method employed. Preferred surfaces for methods of the invention include epoxide surfaces and polyelectrolyte multilayer surfaces, such as those described in Braslavsky, et al., supra. Surfaces  
5 preferably are deposited on a substrate that is amenable to optical detection of the surface chemistry, such as glass or silica. The precise surface and substrate used in methods of the invention is immaterial to the functioning of the invention.

[0011] The invention provides also compositions for use in a nucleic acid synthesis reaction. In one embodiment, the compositions comprise labeled nucleotide corresponding to a  
10 first nucleotide species and at least one nucleotide derivative corresponding to a different nucleotide sequence. Methods and compositions of the invention may take into account that different nucleotides incorporate in an extending primer at different rates and that misincorporation rates similarly vary with the type of nucleotide being incorporated. Accordingly, the identity of the one or more nucleotide derivatives used in methods and  
15 compositions of the invention can be varied with the identity of the correct nucleotide sought to be incorporated. For example, if guanine triphosphate frequently misincorporates at an incorporation site in which adenosine triphosphate should be incorporated, but never misincorporates at an incorporation site in which thymidine triphosphate or cytidine triphosphate should be incorporated, there is no need to add nucleotide derivatives corresponding to  
20 thymidine triphosphate or cytidine triphosphate in a synthesis reaction in which guanine triphosphate is used as the free nucleotide.

[0012] According to the invention, a nucleotide derivative forms a hydrogen bond with a complementary nucleotide on a template nucleic acid, but comprises a modification that prohibits the polymerase from incorporating the derivative into the primer because it prevents the  
25 formation of a phosphodiester bond with the 3' hydroxyl group of the primer. Preferred nucleotide derivatives comprise a modification to the normal nucleotide triphosphate group to prevent formation of a phosphodiester bond. Preferably, the oxygen bridging the  $\alpha$ - and  $\beta$ -phosphates is replaced with another molecule that prevents formation of the phosphodiester bond. Particularly preferred derivatives comprise a substitution of the oxygen molecule bridging  
30 the  $\alpha$ - and  $\beta$ -phosphate with a carbon, nitrogen or sulfur molecule or a methylene group. Other nucleotide derivatives useful in the invention comprise a modification of the  $\alpha$ -,  $\beta$ - or  $\gamma$ -phosphate group. The modification may comprise, for example, the substitution of a bridging or



non-bridging oxygen molecule with a thiol, alkyl, carbonyl, amine, alcohol, aryl or an amino acid group.

[0013] Nucleotides useful in the invention include any nucleotide or nucleotide analog, whether naturally-occurring or synthetic. For example, preferred nucleotides include phosphate  
5 esters of deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine, adenosine, cytidine, guanosine, and uridine.

[0014] Polymerases useful in the invention include any nucleic acid polymerase capable of catalyzing a template-dependent addition of a nucleotide or nucleotide analog to a primer. Depending on the characteristics of the target nucleic acid, a DNA polymerase, an RNA  
10 polymerase, a reverse transcriptase, or a mutant or altered form of any of the foregoing can be used. According to one aspect of the invention, a thermophilic polymerase is used, such as ThermoSequenase<sup>®</sup>, 9<sup>°</sup>N<sup>™</sup>, Terminator<sup>™</sup>, Taq, Tne, Tma, Pfu, Tfl, Tth, Tli, Stoffel fragment, Vent<sup>™</sup> and Deep Vent<sup>™</sup> DNA polymerase.

### **Brief Description of the Drawing**

15 [0015] Figure 1 is a schematic representation of a polymerization reaction that results in the addition of a nucleotide to a polynucleotide.

[0016] Figure 2 is a schematic representation of a DNA synthesis reaction performed in the presence of a nucleotide derivative.

### **Detailed Description**

20 [0017] The invention provides methods and compositions for improving the fidelity of a nucleic acid sequencing-by-synthesis reaction by reducing misincorporation of nucleotides. While applicable to bulk sequencing methods, the invention is particularly useful in connection with single molecule sequencing methods. Methods of the invention improve the fidelity of DNA synthesis by blocking misincorporation of a nucleotide triphosphate in template/primer  
25 duplexes. According to the invention, when DNA synthesis reaction is performed in the presence of a nucleotide derivative that is complementary to the template nucleotide, the nucleotide derivative forms hydrogen bonds with the template nucleotide, but does not form a phosphodiester bond with the 3' hydroxyl group of the primer. Without incorporating into the primer, the nucleotide derivative is transiently held in place by the hydrogen bonds and blocks  
30 the misincorporation of the nucleotide triphosphate.

[0018] Figure 2 is a schematic representation of DNA synthesis reaction in which a nucleotide derivative (as shown  $\alpha,\beta$ -S-2'-deoxyadenosine 5'-triphosphate) (10) has formed hydrogen bonds with its complementary base (12) located on the template (14) opposite the incorporation site immediately adjacent the 3' terminus of the primer (16). Because the  
5 nucleotide derivative lacks an oxygen molecule bridging the  $\alpha$  and  $\beta$  phosphate groups, the polymerase (18) is not able to catalyze the formation of a phosphodiester bond between the nucleotide derivative (10) and the 3' hydroxyl group of the primer (16), but because the derivative (10) is blocking the incorporation site at the 3' terminus of the primer (16), the polymerase (18) can not misincorporate an incorrect (i.e., non-complementary) nucleotide(20).

10 [0019] According to the invention, a polymerization reaction is conducted in the presence of a polymerase, at least one labeled dNTP corresponding to a first nucleotide species, and at least one nucleotide derivative corresponding to a different nucleotide species. The nucleotide derivative comprises a modification that inhibits the formation of a phosphodiester bond between the derivative and the 3' hydroxyl group of a primer on a primer/template duplex.  
15 If the template comprises a nucleotide opposite the incorporation site that is not complementary to the dNTP, but is complementary to the nucleotide derivative, the polymerase recognizes the nucleotide derivative as complementary to the template nucleotide and brings the derivative into close proximity with the template nucleotide, which facilitates the formation of hydrogen bonds between the derivative and its complementary nucleotide. With the nucleotide derivative  
20 blocking the incorporation site, misincorporation of the dNTP is prevented.

[0020] Methods and compositions of the invention are particularly well-suited for use in single molecule sequencing techniques. Substrate-bound primer/template duplexes are exposed to a polymerase, a labeled nucleotide corresponding to a first nucleotide species, and at least one nucleotide derivative corresponding to a different nucleotide species. The duplex may be  
25 simultaneously exposed to the polymerase, the labeled nucleotide, and the nucleotide derivative; or it may be first exposed to the nucleotide derivative and then to the polymerase and labeled nucleotide. Typically, however, the duplex is simultaneously exposed to the polymerase, the labeled nucleotide, and three nucleotide derivatives, one corresponding to each of the three remaining nucleotide species. The duplexes are washed of unincorporated labeled nucleotides  
30 and nucleotide derivatives, and the incorporation of labeled nucleotide is determined. The identity of the nucleotide positioned on the template opposite the incorporate nucleotide is likewise determined. The polymerization reaction is serially repeated in the presence of a



labeled nucleotide that corresponds to each of the other nucleotide species in order to compile a sequence of incorporated nucleotides that is representative of the complement to the template nucleic acid.

[0021] Certain non-limiting aspects of the invention are further described below in terms  
5 of general considerations and examples.

## I. General Considerations

### A. Nucleotides

[0022] Nucleotides useful in the invention as free, unbound nucleotides for incorporation into a primer/template duplex include any nucleotide or nucleotide analog, whether naturally-occurring or synthetic. For example, preferred nucleotides include phosphate esters of deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine, adenosine, cytidine, guanosine, and uridine. Other nucleotides useful in the invention comprise an adenine, cytosine, guanine, thymine base, an xanthine or hypoxanthine; 5-bromouracil, 2-aminopurine, deoxyinosine, or methylated cytosine, such as 5-methylcytosine, and N4-methoxydeoxycytosine. Also included are bases of polynucleotide mimetics, such as methylated nucleic acids, *e.g.*, 2'-O-methRNA, peptide nucleic acids, modified peptide nucleic acids, locked nucleic acids and any other structural moiety that can act substantially like a nucleotide or base, for example, by exhibiting base-complementarity with one or more bases that occur in DNA or RNA and/or being capable of base-complementary incorporation, and includes chain-terminating analogs. A nucleotide corresponds to a specific nucleotide species if they share base-complementarity with respect to at least one base.

[0023] Nucleotides for nucleic acid sequencing according to the invention preferably comprise a detectable label. Labeled nucleotides include any nucleotide that has been modified to include a label that is directly or indirectly detectable. Preferred labels include optically-detectable labels, including fluorescent labels or fluorophores, such as fluorescein, rhodamine, derivatized rhodamine dyes, such as TAMRA, phosphor, polymethadine dye, fluorescent phosphoramidite, Texas Red, green fluorescent protein, acridine, cyanine, cyanine 5 dye, cyanine 3 dye, 5-(2'-aminoethyl)-aminonaphthalene-1-sulfonic acid (EDANS), BODIPY, 120 ALEXA or a derivative or modification of any of the foregoing, and also include such labeling systems as hapten labeling. Accordingly, methods of the invention further provide for exposing

the primer/target nucleic acid duplex to a digoxigenin, a fluorescein, an alkaline phosphatase or a peroxidase.

### ***B. Nucleotide Derivatives***

[0024] Nucleotide derivatives useful in the invention include any nucleotide, including those described above in Section A, that forms hydrogen bonds with a complementary nucleotide on a template nucleic acid, but comprises a modification that prevents the formation of a phosphodiester bond with the 3' hydroxyl group of the primer. Preferred nucleotide derivatives will be recognized by the catalytic domain of the polymerase and brought in close proximity with its complementary nucleotide where hydrogen bonding can occur. Accordingly, particularly useful nucleotide derivatives most closely resemble naturally-occurring substrates for polymerases in both chemical formula and structure. Thus, particularly useful nucleotide derivatives comprise phosphate esters of deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine, adenosine, cytidine, guanosine, and uridine. In a preferred embodiment, nucleotide derivatives comprise a triphosphate ester, such as dATP, dCTP, dGTP, dTTP or dUTP. A nucleotide derivative corresponds to a specific nucleotide species if they exhibit the same base-complementarity with respect to at least one base.

[0025] According to one feature, a nucleotide derivative comprises a modification of the oxygen bridging the  $\alpha$ -phosphate and the  $\beta$ -phosphate in a standard nucleotide triphosphate. For example, preferred modifications include the substitution of the oxygen molecule bridging the  $\alpha$ - and  $\beta$ -phosphate groups with a carbon, nitrogen or sulfur molecule or a methylene group. Other nucleotide derivatives useful in the invention comprise a modification of the  $\alpha$ -,  $\beta$ - or  $\gamma$ -phosphate group, such as, for example, the substitution of a bridging or non-bridging oxygen molecule with a thiol, alkyl, carbonyl, amine, alcohol, aryl or an amino acid group; or a bulky group that physically interferes with polymerase function. Custom modified nucleotides are commercially available from, for example, TriLink BioTechnologies, Inc., San Diego, California, Alexis Biochemicals, Inc., Carlsbad, California and BIOLOG Life Science Institute, Germany.

### ***C. Nucleic Acid Polymerases***

[0026] Nucleic acid polymerases generally useful in the invention include DNA polymerases, RNA polymerases, reverse transcriptases, and mutant or altered forms of any of the foregoing. DNA polymerases and their properties are described in detail in, among other places,



DNA Replication 2nd edition, Komberg and Baker, W. H. Freeman, New York, N.Y. (1991). Known conventional DNA polymerases useful in the invention include, but are not limited to, *Pyrococcus furiosus* (Pfu) DNA polymerase (Lundberg et al., 1991, Gene, 108: 1, Stratagene), *Pyrococcus woesei* (Pwo) DNA polymerase (Hinnisdaels et al., 1996, Biotechniques, 20:186-8, Boehringer Mannheim), *Thermus thermophilus* (Tth) DNA polymerase (Myers and Gelfand 1991, Biochemistry 30:7661), *Bacillus stearothermophilus* DNA polymerase (Stenesh and McGowan, 1977, Biochim Biophys Acta 475:32), *Thermococcus litoralis* (Tli) DNA polymerase (also referred to as Vent<sup>TM</sup> DNA polymerase, Cariello et al., 1991, Polynucleotides Res, 19: 4193, New England Biolabs), 9<sup>o</sup>Nm<sup>TM</sup> DNA polymerase (New England Biolabs), Stoffel 10 fragment, ThermoSequenase<sup>®</sup> (Amersham Pharmacia Biotech UK), Terminator<sup>TM</sup> (New England Biolabs), *Thermotoga maritima* (Tma) DNA polymerase (Diaz and Sabino, 1998 Braz J Med. Res, 31:1239), *Thermus aquaticus* (Taq) DNA polymerase (Chien et al., 1976, J. Bacteriol, 127: 1550), DNA polymerase, *Pyrococcus kodakaraensis* KOD DNA polymerase (Takagi et al., 1997, Appl. Environ. Microbiol. 63:4504), JDF-3 DNA polymerase (from 15 *thermococcus* sp. JDF-3, Patent application WO 0132887), *Pyrococcus* GB-D (PGB-D) DNA polymerase (also referred as Deep Vent<sup>TM</sup> DNA polymerase, Juncosa-Ginesta et al., 1994, Biotechniques, 16:820, New England Biolabs), UITma DNA polymerase (from thermophile *Thermotoga maritima*; Diaz and Sabino, 1998 Braz J. Med. Res, 31:1239; PE Applied Biosystems), Tgo DNA polymerase (from *thermococcus gorgonarius*, Roche Molecular 20 Biochemicals), *E. coli* DNA polymerase I (Lecomte and Doubleday, 1983, Polynucleotides Res. 11:7505), T7 DNA polymerase (Nordstrom et al., 1981, J Biol. Chem. 256:3112), and archaeal DP1I/DP2 DNA polymerase II (Cann et al., 1998, Proc Natl Acad. Sci. USA 95:14250-->5).

[0027] While mesophilic polymerases are contemplated by the invention, preferred polymerases are thermophilic. Thermophilic DNA polymerases include, but are not limited to, 25 ThermoSequenase<sup>®</sup>, 9<sup>o</sup>Nm<sup>TM</sup>, Terminator<sup>TM</sup>, Taq, Tne, Tma, Pfu, Tfl, Tth, Tli, Stoffel fragment, Vent<sup>TM</sup> and Deep Vent<sup>TM</sup> DNA polymerase, KOD DNA polymerase, Tgo, JDF-3, and mutants, variants and derivatives thereof.

[0028] Reverse transcriptases useful in the invention include, but are not limited to, reverse transcriptases from HIV, HTLV-1, HTLV-II, FeLV, FIV, SIV, AMV, MMTV, 30 MoMuLV and other retroviruses (see Levin, Cell 88:5-8 (1997); Verma, Biochim Biophys Acta. 473:1-38 (1977); Wu et al., CRC Crit Rev Biochem. 3:289-347(1975)).



[0029] Certain embodiments of the invention are described in the following examples, which are not meant to be limiting.

***Example 1: Single Molecule Sequencing Using Nucleotide Derivatives***

[0030] The following nucleotide derivatives, each comprising a substitution of the oxygen molecule bridging the  $\alpha$  and  $\beta$  phosphate groups with a sulfur molecule, are prepared:  $\alpha,\beta$ -S-2'-deoxyadenosine 5'-triphosphate;  $\alpha,\beta$ -S-2'-deoxycytidine 5'-triphosphate;  $\alpha,\beta$ -S-2'-deoxyguanosine 5'-triphosphate; and  $\alpha,\beta$ -S-2'-deoxythymidine 5'-triphosphate. Primer/template duplexes are bound to a solid support in a concentration that results in individually optically resolvable duplexes. The bound duplexes are subjected to serial sequencing-by-synthesis cycles in which a polymerase, labeled standard nucleotide, and nucleotide derivatives corresponding to each of the three non-standard nucleotide species are combined. The incorporation of a labeled nucleotide is determined, recorded and the reaction serially repeated with labeled nucleotide corresponding to each of the different nucleotide species and the appropriate nucleotide derivatives in order to compile a sequence that is representative of the complement of the target nucleic acid.

***Example 2: Assay for the Inhibition of Polymerase Activity***

[0031] A simple assay may be performed to assess the ability of a usefulness of nucleotide derivative in methods and compositions of the invention.

[0032] Primed DNA template is diluted into an appropriate volume of 20 mM Tris-HCl, pH 7.5 and the enzyme is diluted into an appropriate volume of 20 mM Tris-HCl, containing 2 mM  $\beta$ -mercaptoethanol, and 100 mM KCl. 0.05 mg/mL primed DNA template and 0.01 U/ $\mu$ L DNA polymerase are pipetted into microcentrifuge tubes or a 96-well plate in a 50  $\mu$ L reaction volume. 1.6  $\mu$ M nucleotide derivative is added to the test reactions. Blank reactions excluding enzyme and control reactions excluding test compound are also prepared using enzyme dilution buffer and test compound solvent, respectively. The reaction is initiated with reaction buffer with the following components 20 mM Tris-HCl, pH 7.5; 200  $\mu$ g/mL bovine serum albumin; 100 mM KCl; 2 mM  $\beta$ -mercaptoethanol; 10 mM  $MgCl_2$ , 1.6  $\mu$ M dATP, dGTP, dCTP, and dTTP; and  $\alpha$ - $^{33}P$ -dAT. The reaction is incubated for sixty minutes at 37° C and then quenched by the addition of 20  $\mu$ L of 0.5M EDTA. 50  $\mu$ L of the quenched reaction is spotted onto Whatman DE81 filter disks and air dried. The filter disks are repeatedly washed with 150 mL 0.3M



ammonium formate, pH 8 until 1 mL of wash is <100 cpm. The disks are washed twice with 150 mL absolute ethanol and once with 150 mL anhydrous ether, dried and counted in 5 mL scintillation fluid.

[0033] The percentage of inhibition of the test is calculated according to the following  
5 equation:  $\% \text{ inhibition} = \{1 - (\text{cpm in test reaction} - \text{cpm in blank}) / (\text{cpm in control reaction} - \text{cpm in blank})\} \times 100$ .

[0034] The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative rather than limiting on the invention described herein.  
10 Scope of the invention is thus indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

**Claims**

We claim:

1. A method for inhibiting nucleotide misincorporation in a nucleic acid synthesis reaction, the method comprising  
  
conducting a polymerization reaction on a nucleic acid duplex comprising a template and a primer in the presence of a nucleic acid polymerase, a first nucleotide corresponding to a first nucleotide species, and at least one nucleotide derivative corresponding to a different nucleotide species,  
  
wherein said nucleotide derivative comprises a modification that inhibits formation of a phosphodiester bond between the nucleotide derivative and a free 3' hydroxyl on a primer nucleotide.
2. The method of claim 1, further comprising removing nucleotide derivatives and nucleotides that are not incorporated into the primer.
3. The method of claim 2, further comprising the step of repeating said conducting and removing steps using a second nucleotide species and a nucleotide derivative corresponding to a nucleotide species different from said second species.
4. The method of claim 1, wherein said nucleotide derivative comprises a triphosphate group lacking an oxygen between an  $\alpha$ -phosphate and a  $\beta$ -phosphate such that said nucleotide derivative is prevented from forming a phosphodiester bond with said free 3' hydroxyl group on the primer.
5. The method of claim 4 wherein said derivative comprises substitution of an oxygen molecule with a carbon, nitrogen or sulfur molecule.
6. The method of claim 4 wherein said derivative comprises substitution of an oxygen molecule with a methylene group.
7. The method of claim 4 wherein said derivative comprises substitution of an oxygen molecule with a thiol, alkyl, carbonyl, amine, alcohol, aryl or an amino acid group.
8. The method of claim 1 wherein, if said template comprises a nucleotide opposite a position immediately adjacent to a 3' terminus of the primer that is complementary to said



nucleotide derivative but not complementary to said first nucleotide, said nucleotide derivative forms a hydrogen bond with its complementary nucleotide so as to prevent said first nucleotide from bonding thereto and misincorporating into said primer in said nucleic acid synthesis reaction.

9. The method of claim 1 wherein said nucleic acid duplex is bound to a solid support.

10. The method of claim 1 comprising a plurality of said duplex, each member of said plurality being individually optically resolvable.

11. The method of claim 1 wherein said first nucleotide is detectably labeled.

12. A method for sequencing a nucleic acid, the method comprising the steps of:

(a) exposing a nucleic acid duplex comprising a template and a primer to a nucleic acid polymerase, a first labeled nucleotide, and at least one nucleotide derivative corresponding to a nucleotide species different from said first nucleotide;

wherein said nucleotide derivative comprises a modification that inhibits formation of a phosphodiester bond between the nucleotide derivative and a 3' hydroxyl group of the primer, and

(b) removing nucleotide derivatives and labeled nucleotides that are not incorporated into the primer;

(c) determining incorporation of the labeled nucleotide and thereby identifying a nucleotide that is complementary to a nucleotide at a position on said template opposite the incorporated labeled nucleotide; and

(d) repeating steps (a), (b) and (c) with a second labeled nucleotide and at least one nucleotide derivative corresponding to a nucleotide species different from said second labeled nucleotide.

13. The method of claim 12, wherein said nucleotide derivative comprises a triphosphate group having a molecule other than oxygen between an  $\alpha$ -phosphate and a  $\beta$ -phosphate.

14. The method of claim 13 wherein said derivative comprises substitution of an oxygen molecule with a carbon, nitrogen or sulfur molecule.

15. The method of claim 13 wherein said derivative comprises a substitution of an oxygen molecule with a methylene group.

16. The method of claim 2 wherein said nucleotide derivative comprises a triphosphate group having a modified  $\alpha$ - or  $\beta$ -phosphate.
17. The method of claim 16 wherein said modification comprises a substitution of an oxygen molecule with a thiol, alkyl, carbonyl, amine, alcohol, aryl or an amino acid group.
18. The method of claim 12 wherein said exposing step is performed in the presence of at least two nucleotide derivatives, each derivative corresponding to a nucleotide species different from said first nucleotide.
19. The method of claim 12 wherein said exposing step is performed in the presence of three nucleotide derivatives, each derivative corresponding to a nucleotide species different from said first nucleotide.
20. The method of claim 12, further comprising the step of compiling a sequence of nucleotides incorporated into said primer.
21. The method of claim 12 wherein said labeled nucleotide comprises a labeled dNTP.
22. The method of claim 12 wherein said first nucleotide is a deoxyadenosine triphosphate and said at least one nucleotide derivative corresponds to a nucleotide species selected from the group consisting of deoxyguanosine triphosphate, deoxycytidine triphosphate, and deoxyuridine triphosphate.
23. The method of claim 12 wherein said first nucleotide is a deoxyguanosine triphosphate thereof and said at least one nucleotide derivative corresponds to a nucleotide species selected from the group consisting of deoxyadenosine triphosphate, deoxycytidine triphosphate and deoxyuridine triphosphate.
24. The method of claim 12 wherein said first nucleotide is a deoxycytidine triphosphate thereof and said three nucleotide derivatives corresponds to the nucleotide species of deoxyadenosine triphosphate, deoxyguanosine triphosphate and deoxyuridine triphosphate.
25. The method of claim 12 wherein said first nucleotides is a deoxyuridine triphosphate thereof and said at least one nucleotide derivative corresponds to a nucleotide species selected from the group consisting of deoxyadenosine triphosphate, deoxyguanosine triphosphate and deoxycytidine triphosphate.



26. The method of claim 19 wherein said first nucleotide is a deoxyadenosine triphosphate thereof and said three nucleotide derivatives correspond to deoxyguanosine triphosphate, deoxycytidine triphosphate, and deoxyuridine triphosphate.
27. The method of claim 19 wherein said first nucleotide is a deoxyguanosine triphosphate and said three nucleotide derivatives correspond to deoxyadenosine triphosphate, deoxycytidine triphosphate and deoxyuridine triphosphate.
28. The method of claim 19 wherein said first nucleotide is a deoxycytidine triphosphate and said three nucleotide derivatives correspond to deoxyadenosine triphosphate, deoxyguanosine triphosphate and deoxyuridine triphosphate.
29. The method of claim 19 wherein said first nucleotide is a deoxyuridine triphosphate and said three nucleotide derivatives correspond to deoxyadenosine triphosphate, deoxyguanosine triphosphate and deoxycytidine triphosphate.
30. The method of claim 12, wherein said exposing step comprises first exposing said nucleic acid duplex to said at least one nucleotide derivative, then exposing said nucleic acid duplex to said nucleic acid polymerase and said labeled nucleotide.
31. A composition for use in a nucleic acid synthesis reaction on a primer/template duplex, the composition comprising:
- a labeled nucleotide corresponding to a first nucleotide species, and at least one nucleotide derivative of a different nucleotide species,
  - wherein said nucleotide derivative comprises a modification that prevents the formation of a phosphodiester bond between the nucleotide derivative and a 3' hydroxyl group of the primer.

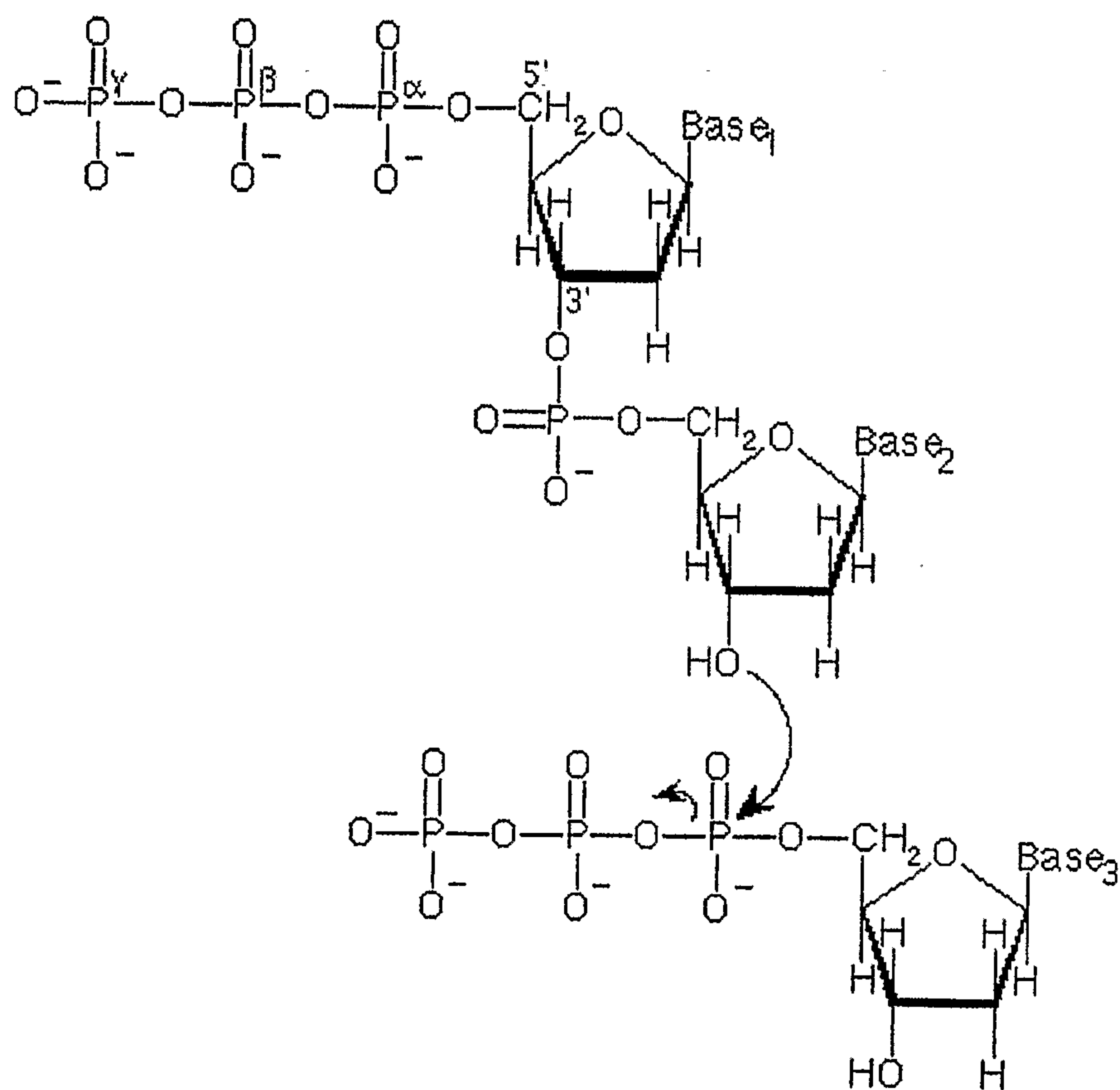


FIG. 1



