

METHODS OF SEQUENCING USING 3' ALLYL BLOCKED NUCLEOTIDES

BACKGROUND

Field

[0001] The present disclosure generally relates to polynucleotide sequencing methods, compositions, and kits for sequencing. The present disclosure also relates to methods to remove 3' blocking groups.

Description of the Related Art

[0002] Advances in the study of molecules have been led, in part, by improvement in technologies used to characterize the molecules or their biological reactions. In particular, the study of the nucleic acids DNA and RNA has benefited from developing technologies used for sequence analysis and the study of hybridization events.

[0003] An example of the technologies that have improved the study of nucleic acids is the development of fabricated arrays of immobilized nucleic acids. These arrays consist typically of a high-density matrix of polynucleotides immobilized onto a solid support material. See, e.g., Fodor et al., *Trends Biotech.* 12: 19-26, 1994, which describes ways of assembling the nucleic acids using a chemically sensitized glass surface protected by a mask, but exposed at defined areas to allow attachment of suitably modified nucleotide phosphoramidites. Fabricated arrays can also be manufactured by the technique of "spotting" known polynucleotides onto a solid support at predetermined positions (e.g., Stimpson et al., *Proc. Natl. Acad. Sci.* 92: 6379-6383, 1995).

[0004] One way of determining the nucleotide sequence of a nucleic acid bound to an array is called "sequencing by synthesis" or "SBS." This technique for determining the sequence of DNA ideally requires the controlled (i.e., one at a time) incorporation of the correct complementary nucleotide opposite the nucleic acid being sequenced. This allows for accurate sequencing by adding nucleotides in multiple cycles as each nucleotide residue is sequenced one at a time, thus preventing an uncontrolled series of incorporations from occurring. The incorporated nucleotide is read using an appropriate label attached thereto before removal of the label moiety and the subsequent next round of sequencing.

[0005] In order to ensure that only a single incorporation occurs, a structural modification ("protecting group" or "blocking group") is included in each labeled nucleotide that is added to the growing chain to ensure that only one nucleotide is incorporated. After the nucleotide with the protecting group has been added, the protecting group is then removed, under reaction conditions which do not interfere with the integrity of the DNA being sequenced. The

sequencing cycle can then continue with the incorporation of the next protected, labeled nucleotide. To be useful in DNA sequencing, nucleotides, which are usually nucleotide triphosphates, generally require a 3' hydroxy blocking group so as to prevent the polymerase used to incorporate it into a polynucleotide chain from continuing to replicate once the base on the nucleotide is added.

[0006] Reversible protecting groups have been described previously. For example, Metzker *et al.*, (*Nucleic Acids Research*, 22 (20): 4259-4267, 1994) discloses the synthesis and use of eight 3'-modified 2-deoxyribonucleoside 5'-triphosphates (3'-modified dNTPs) and testing in two DNA template assays for incorporation activity. WO 2002/029003 describes a sequencing method which may include the use of an allyl protecting group to cap the 3'-OH group on a growing strand of DNA in a polymerase reaction. In addition, the development of a number of reversible protecting groups and methods of deprotecting them under DNA compatible conditions was previously reported in International Application Publication Nos. WO 2004/018497 and WO 2014/139596, each of which is hereby incorporated by reference in its entirety. Although the use of 3' allyl blocked nucleotides has been reported over 20 years ago, SBS using such nucleotides was never successfully applied in the commercial setting for long DNA reads, due to various factors including the low efficiency of the 3' blocking group cleavage chemistry, as well as the rapid loss of signal intensity during sequencing. For example, Ju *et al.* reported four-color DNA sequencing by synthesis using 3' allyl blocked nucleotides and only 25 cycles of SBS were reported with substantial loss of fluorescent intensity only after five cycles. *See* Ju *et al.*, PNAS, 2006, 103(52):19635-19640. As such, there remain a need to improve the sequencing conditions (e.g., the cleavage condition of the 3' allyl blocking group) and the sequencing metrics of SBS using nucleotides with 3' allyl blocking group.

SUMMARY

[0007] One aspect of the present disclosure relates to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

- (i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;
- (ii) the solid support comprises a plurality of different target polynucleotides; and
- (iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein each nucleotide comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;

(c) imaging the solid support to determine the identity of incorporated nucleotides;

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove 3' allyl blocking groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support;

(e) contacting said solid support with an aqueous wash solution; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences.

[0008] Another aspect of the present disclosure relates to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

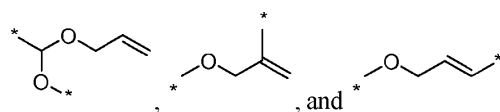
(i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;

(ii) the solid support comprises a plurality of different target polynucleotides; and

(iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C, and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein:

(i) each of at least three types of nucleotides independently comprises a base that is attached to a detectable label via a cleavable linker, and the cleavable linker comprises a moiety selected from the group consisting of:



is connected to the remainder of the nucleotide;

(ii) each nucleotide comprises a 2' deoxyribose moiety with a 3' allyl group  attached to the 3' oxygen atom; and

(iii) the DNA polymerase is an altered archaeal DNA polymerase;

(c) contacting the solid support with a solution comprising one or more radical scavengers and imaging the solid support to determine the identity of incorporated nucleotides;

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl blocking groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers;

(e) contacting said solid support with an aqueous wash solution comprising a palladium scavenger; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences.

[0009] Another aspect of the present disclosure relates to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

(i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;

(ii) the solid support comprises a plurality of different target polynucleotides; and

(iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C, and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein:

(i) at least one of the nucleotides comprise a base that is attached to a detectable label via a cleavable linker; and

(ii) the nucleotides each comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;

(c) contacting the solid support with a solution comprising one or more radical scavengers and imaging the solid support to determine the identity of incorporated nucleotides;

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers;

(e) contacting said solid support with an aqueous wash solution; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences.

[0010] Another aspect of the present disclosure relates to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

(i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;

(ii) the solid support comprises a plurality of different target polynucleotides; and

(iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous incorporation mixture comprising DNA polymerase and one or more of four types of nucleotides A, G, C, and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein:

(i) the nucleotides each comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;

(ii) at least one type of nucleotide is unlabeled; and

(iii) the first type of unlabeled nucleotides comprises a first functional moiety;

(c) contacting the extended copy polynucleotides with an aqueous labeling mixture comprising a first labeling reagent, wherein the first labeling reagent comprises one or more first detectable labels and a first binding moiety that is capable of specific binding to the first functional moiety of the first type of unlabeled nucleotide;

(d) imaging the solid support and performing one or more fluorescent measurements to determine the identity of incorporated nucleotides;

(e) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support;

(f) contacting said solid support with an aqueous wash solution; and

(g) repeating steps (b)-(f) to determine target polynucleotide sequences.

[0011] A further aspect of the present disclosure relates to a sequencing kit comprising:

(a) an incorporation mixture comprising DNA polymerase and nucleotides A, G, C, and T or U, wherein:

- (i) the nucleotides comprise a 2' deoxyribose moiety with a 3' allyl group  attached to the 3' carbon atom; and
- (ii) the DNA polymerase is an altered archaeal DNA polymerase;
- (b) an aqueous deblocking solution comprising a palladium catalyst, tris(hydroxyalkyl)phosphine, and one or more buffer reagents that is suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers; and
- (c) an aqueous wash solution comprising a Pd(II) scavenger;
- wherein said kit is configured for performing at least about 100 cycles of sequencing-by-synthesis.

[0012] In some embodiments of the method or kit described herein, the tris(hydroxyalkyl)phosphine comprises or is tris(hydroxypropyl)phosphine (THPP).

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] **FIG. 1** is a line chart of percent peak area of nucleoside with 3' allyl blocking group in comparison to nucleoside with 3' AOM blocking group as a functional of time in a Pd cleavage solution kinetic assay.

[0014] **FIG. 2** is a line chart of percent 3'-OH nucleoside monophosphate formation as a functional of time in a Pd cleavage solution kinetic assay of three different 3' blocked nucleoside monophosphate.

[0015] **FIG. 3A** is a line chart of percent error rate of standard sequence by synthesis runs on Illumina's iSeq100™ instrument, using three different fully functionalized T nucleotide (ffT) either having a 3' allyl blocking group or a 3' AOM blocking group.

[0016] **FIG. 3B** is a line chart of percent phasing of standard sequence by synthesis runs on Illumina's iSeq100™ instrument, using three different fully functionalized T nucleotide (ffT) either having a 3' allyl blocking group or a 3' AOM blocking group.

[0017] **FIG. 4A** is a line chart of percent error rate of sequence by synthesis runs on Illumina's Miseq® instrument in a post-incorporation labeling workflow, using two different fully functionalized T nucleotide (ffT) either having a 3' allyl blocking group or a 3' AOM blocking group.

[0018] **FIG. 4B** is a line chart of percent phasing of sequence by synthesis runs on Illumina's Miseq® instrument in a post-incorporation labeling workflow, using two different fully functionalized T nucleotide (ffT) either having a 3' allyl blocking group or a 3' AOM blocking group.

DETAILED DESCRIPTION

[0019] Some aspects of the present disclosure relate to methods of nucleic acid sequencing using nucleotide with 3' allyl blocking group. In particular, the sequencing method described herein involves the use of a Pd(0) catalyst formed from a Pd(II) complex and a water soluble tris(hydroxyalkyl)phosphine to cleave the 3' allyl blocking group after each cycle of the sequencing by synthesis. In some embodiments, the water-soluble phosphine is THPP. The method also involves the use of Pd(0) and/or Pd(II) scavengers. In further embodiments, the nucleotides used in the incorporation step of the sequencing method are labeled nucleotides where the detectable label is covalently attached to the nucleobase via a cleavable linker containing an allyl moiety. The 3' allyl blocking nucleotides can be used in both standard SBS setting in which the nucleotides are labeled, or in a post-incorporation labeling workflow in which the nucleotides are unlabeled.

Definitions

[0020] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of ordinary skill in the art. The use of the term "including" as well as other forms, such as "include," "includes," and "included," is not limiting. The use of the term "having" as well as other forms, such as "have," "has," and "had," is not limiting. As used in this specification, whether in a transitional phrase or in the body of the claim, the terms "comprise(s)" and "comprising" are to be interpreted as having an open-ended meaning. That is, the above terms are to be interpreted synonymously with the phrases "having at least" or "including at least." For example, when used in the context of a process, the term "comprising" means that the process includes at least the recited steps, but may include additional steps. When used in the context of a compound, composition, or device, the term "comprising" means that the compound, composition, or device includes at least the recited features or components, but may also include additional features or components.

[0021] Where a range of values is provided, it is understood that the upper and lower limit, and each intervening value between the upper and lower limit of the range is encompassed within the embodiments.

[0022] As used herein, common organic abbreviations are defined as follows:

°C	Temperature in degrees Centigrade
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
dTTP	Deoxythymidine triphosphate

ddNTP	Dideoxynucleotide triphosphate
ffN	Fully functionalized nucleotide
ffA	Fully functionalized “A” nucleotide
ffC	Fully functionalized “C” nucleotide
ffT	Fully functionalized “T” nucleotide
ffG	Fully functionalized “G” nucleotide
IMX	Incorporation mix or Incorporation mixture
Pd	Palladium
RT	Room temperature
SBS	Sequencing by Synthesis
TCEP	tris(2-carboxethyl)phosphine
THPP	tris(hydroxypropyl)phosphine

[0023] As used herein, the term “array” refers to a population of different probe molecules that are attached to one or more substrates such that the different probe molecules can be differentiated from each other according to relative location. An array can include different probe molecules that are each located at a different addressable location on a substrate. Alternatively, or additionally, an array can include separate substrates each bearing a different probe molecule, wherein the different probe molecules can be identified according to the locations of the substrates on a surface to which the substrates are attached or according to the locations of the substrates in a liquid. Exemplary arrays in which separate substrates are located on a surface include, without limitation, those including beads in wells as described, for example, in U.S. Patent No. 6,355,431 B1, US 2002/0102578 and PCT Publication No. WO 00/63437. Exemplary formats that can be used in the invention to distinguish beads in a liquid array, for example, using a microfluidic device, such as a fluorescent activated cell sorter (FACS), are described, for example, in US Pat. No. 6,524,793. Further examples of arrays that can be used in the invention include, without limitation, those described in U.S. Pat Nos. 5,429,807; 5,436,327; 5,561,071; 5,583,211; 5,658,734; 5,837,858; 5,874,219; 5,919,523; 6,136,269; 6,287,768; 6,287,776; 6,288,220; 6,297,006; 6,291,193; 6,346,413; 6,416,949; 6,482,591; 6,514,751 and 6,610,482; and WO 93/17126; WO 95/11995; WO 95/35505; EP 742 287; and EP 799 897.

[0024] As used herein, the term “covalently attached” or “covalently bonded” refers to the forming of a chemical bonding that is characterized by the sharing of pairs of electrons between atoms. For example, a covalently attached polymer coating refers to a polymer coating that forms chemical bonds with a functionalized surface of a substrate, as compared to attachment to the surface via other means, for example, adhesion or electrostatic interaction. It will be

appreciated that polymers that are attached covalently to a surface can also be bonded via means in addition to covalent attachment.

[0025] As used herein, any “R” group(s) represent substituents that can be attached to the indicated atom. An R group may be substituted or unsubstituted.

[0026] It is to be understood that certain radical naming conventions can include either a mono-radical or a di-radical, depending on the context. For example, where a substituent requires two points of attachment to the rest of the molecule, it is understood that the substituent is a di-radical. For example, a substituent identified as alkyl that requires two points of attachment includes di-radicals such as $-\text{CH}_2-$, $-\text{CH}_2\text{CH}_2-$, $-\text{CH}_2\text{CH}(\text{CH}_3)\text{CH}_2-$, and the like. Other radical naming conventions clearly indicate that the radical is a di-radical such as “alkylene” or “alkenylene.”

[0027] The term “halogen” or “halo,” as used herein, means any one of the radio-stable atoms of column 7 of the Periodic Table of the Elements, *e.g.*, fluorine, chlorine, bromine, or iodine, with fluorine and chlorine being preferred.

[0028] As used herein, “C_a to C_b,” “C_a-C_b,” or “C_{a-b}” in which “a” and “b” are integers refer to the number of carbon atoms in an alkyl, alkenyl or alkynyl group, or the number of ring atoms of a cycloalkyl or aryl group. That is, the alkyl, the alkenyl, the alkynyl, the ring of the cycloalkyl, and ring of the aryl can contain from “a” to “b,” inclusive, carbon atoms. For example, a “C₁ to C₄ alkyl” group refers to all alkyl groups having from 1 to 4 carbons, that is, CH_3- , CH_3CH_2- , $\text{CH}_3\text{CH}_2\text{CH}_2-$, $(\text{CH}_3)_2\text{CH}-$, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_2-$, $\text{CH}_3\text{CH}_2\text{CH}(\text{CH}_3)-$ and $(\text{CH}_3)_3\text{C}-$; a C₃ to C₄ cycloalkyl group refers to all cycloalkyl groups having from 3 to 4 carbon atoms, that is, cyclopropyl and cyclobutyl. Similarly, a “4 to 6 membered heterocyclyl” group refers to all heterocyclyl groups with 4 to 6 total ring atoms, for example, azetidine, oxetane, oxazoline, pyrrolidine, piperidine, piperazine, morpholine, and the like. If no “a” and “b” are designated with regard to an alkyl, alkenyl, alkynyl, cycloalkyl, or aryl group, the broadest range described in these definitions is to be assumed. As used herein, the term “C₁-C₆” includes C₁, C₂, C₃, C₄, C₅ and C₆, and a range defined by any of the two numbers. For example, C₁-C₆ alkyl includes C₁, C₂, C₃, C₄, C₅ and C₆ alkyl, C₂-C₆ alkyl, C₁-C₃ alkyl, etc. Similarly, C₂-C₆ alkenyl includes C₂, C₃, C₄, C₅ and C₆ alkenyl, C₂-C₅ alkenyl, C₃-C₄ alkenyl, etc.; and C₂-C₆ alkynyl includes C₂, C₃, C₄, C₅ and C₆ alkynyl, C₂-C₅ alkynyl, C₃-C₄ alkynyl, etc. C₃-C₈ cycloalkyl each includes hydrocarbon ring containing 3, 4, 5, 6, 7 and 8 carbon atoms, or a range defined by any of the two numbers, such as C₃-C₇ cycloalkyl or C₅-C₆ cycloalkyl.

[0029] As used herein, “alkyl” refers to a straight or branched hydrocarbon chain that is fully saturated (*i.e.*, contains no double or triple bonds). The alkyl group may have 1 to 20 carbon atoms (whenever it appears herein, a numerical range such as “1 to 20” refers to each

integer in the given range; *e.g.*, “1 to 20 carbon atoms” means that the alkyl group may consist of 1 carbon atom, 2 carbon atoms, 3 carbon atoms, *etc.*, up to and including 20 carbon atoms, although the present definition also covers the occurrence of the term “alkyl” where no numerical range is designated). The alkyl group may also be a medium size alkyl having 1 to 9 carbon atoms. The alkyl group could also be a lower alkyl having 1 to 6 carbon atoms. The alkyl group may be designated as “C₁-C₄ alkyl” or similar designations. By way of example only, “C₁-C₆ alkyl” indicates that there are one to six carbon atoms in the alkyl chain, *i.e.*, the alkyl chain is selected from the group consisting of methyl, ethyl, propyl, iso-propyl, n-butyl, iso-butyl, sec-butyl, and t-butyl. Typical alkyl groups include, but are in no way limited to, methyl, ethyl, propyl, isopropyl, butyl, isobutyl, tertiary butyl, pentyl, hexyl, and the like.

[0030] As used herein, “alkoxy” refers to the formula –OR wherein R is an alkyl as is defined above, such as “C₁-C₉ alkoxy,” including but not limited to methoxy, ethoxy, n-propoxy, 1-methylethoxy (isopropoxy), n-butoxy, iso-butoxy, sec-butoxy, and tert-butoxy, and the like.

[0031] As used herein, “alkenyl” refers to a straight or branched hydrocarbon chain containing one or more double bonds. The alkenyl group may have 2 to 20 carbon atoms, although the present definition also covers the occurrence of the term “alkenyl” where no numerical range is designated. The alkenyl group may also be a medium size alkenyl having 2 to 9 carbon atoms. The alkenyl group could also be a lower alkenyl having 2 to 6 carbon atoms. The alkenyl group may be designated as “C₂-C₆ alkenyl” or similar designations. By way of example only, “C₂-C₆ alkenyl” indicates that there are two to six carbon atoms in the alkenyl chain, *i.e.*, the alkenyl chain is selected from the group consisting of ethenyl, propen-1-yl, propen-2-yl, propen-3-yl, buten-1-yl, buten-2-yl, buten-3-yl, buten-4-yl, 1-methyl-propen-1-yl, 2-methyl-propen-1-yl, 1-ethyl-ethen-1-yl, 2-methyl-propen-3-yl, buta-1,3-dienyl, buta-1,2,-dienyl, and buta-1,2-dien-4-yl. Typical alkenyl groups include, but are in no way limited to, ethenyl, propenyl, butenyl, pentenyl, and hexenyl, and the like.

[0032] The term “aromatic” refers to a ring or ring system having a conjugated pi electron system and includes both carbocyclic aromatic (*e.g.*, phenyl) and heterocyclic aromatic groups (*e.g.*, pyridine). The term includes monocyclic or fused-ring polycyclic (*i.e.*, rings which share adjacent pairs of atoms) groups provided that the entire ring system is aromatic.

[0033] As used herein, “aryl” refers to an aromatic ring or ring system (*i.e.*, two or more fused rings that share two adjacent carbon atoms) containing only carbon in the ring backbone. When the aryl is a ring system, every ring in the system is aromatic. The aryl group may have 6 to 18 carbon atoms, although the present definition also covers the occurrence of the term “aryl” where no numerical range is designated. In some embodiments, the aryl group has 6 to 10 carbon atoms. The aryl group may be designated as “C₆-C₁₀ aryl,” “C₆ or C₁₀ aryl,” or

similar designations. Examples of aryl groups include, but are not limited to, phenyl, naphthyl, azulenyl, and anthracenyl.

[0034] An “aralkyl” or “arylalkyl” is an aryl group connected, as a substituent, via an alkylene group, such as “C₇₋₁₄ aralkyl” and the like, including but not limited to benzyl, 2-phenylethyl, 3-phenylpropyl, and naphthylalkyl. In some cases, the alkylene group is a lower alkylene group (i.e., a C₁-C₆ alkylene group).

[0035] As used herein, “aryloxy” refers to RO- in which R is an aryl, as defined above, such as but not limited to phenyl.

[0036] As used herein, “heteroaryl” refers to an aromatic ring or ring system (i.e., two or more fused rings that share two adjacent atoms) that contain(s) one or more heteroatoms, that is, an element other than carbon, including but not limited to, nitrogen, oxygen and sulfur, in the ring backbone. When the heteroaryl is a ring system, every ring in the system is aromatic. The heteroaryl group may have 5-18 ring members (i.e., the number of atoms making up the ring backbone, including carbon atoms and heteroatoms), although the present definition also covers the occurrence of the term “heteroaryl” where no numerical range is designated. In some embodiments, the heteroaryl group has 5 to 10 ring members or 5 to 7 ring members. The heteroaryl group may be designated as “5-7 membered heteroaryl,” “5-10 membered heteroaryl,” or similar designations. Examples of heteroaryl rings include, but are not limited to, furyl, thienyl, phthalazinyl, pyrrolyl, oxazolyl, thiazolyl, imidazolyl, pyrazolyl, isoxazolyl, isothiazolyl, triazolyl, thiadiazolyl, pyridinyl, pyridazinyl, pyrimidinyl, pyrazinyl, triazinyl, quinolinyl, isoquinolinyl, benzoimidazolyl, benzoxazolyl, benzothiazolyl, indolyl, isoindolyl, and benzothienyl.

[0037] A “heteroaralkyl” or “heteroarylalkyl” is heteroaryl group connected, as a substituent, via an alkylene group. Examples include but are not limited to 2-thienylmethyl, 3-thienylmethyl, furylmethyl, thienylethyl, pyrrolylalkyl, pyridylalkyl, isoxazolylalkyl, and imidazolylalkyl. In some cases, the alkylene group is a lower alkylene group (i.e., a C₁-C₆ alkylene group).

[0038] As used herein, “carbocyclyl” means a non-aromatic cyclic ring or ring system containing only carbon atoms in the ring system backbone. When the carbocyclyl is a ring system, two or more rings may be joined together in a fused, bridged or spiro-connected fashion. Carbocyclyls may have any degree of saturation provided that at least one ring in a ring system is not aromatic. Thus, carbocyclyls include cycloalkyls, cycloalkenyls, and cycloalkynyls. The carbocyclyl group may have 3 to 20 carbon atoms, although the present definition also covers the occurrence of the term “carbocyclyl” where no numerical range is designated. The carbocyclyl group may also be a medium size carbocyclyl having 3 to 10 carbon atoms. The carbocyclyl

group could also be a carbocyclyl having 3 to 6 carbon atoms. The carbocyclyl group may be designated as “C₃-C₆ carbocyclyl” or similar designations. Examples of carbocyclyl rings include, but are not limited to, cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cyclohexenyl, 2,3-dihydro-indene, bicycle[2.2.2]octanyl, adamantyl, and spiro[4.4]nonanyl.

[0039] As used herein, “cycloalkyl” means a fully saturated carbocyclyl ring or ring system. Examples include cyclopropyl, cyclobutyl, cyclopentyl, and cyclohexyl.

[0040] As used herein, “heterocyclyl” means a non-aromatic cyclic ring or ring system containing at least one heteroatom in the ring backbone. Heterocyclyls may be joined together in a fused, bridged or spiro-connected fashion. Heterocyclyls may have any degree of saturation provided that at least one ring in the ring system is not aromatic. The heteroatom(s) may be present in either a non-aromatic or aromatic ring in the ring system. The heterocyclyl group may have 3 to 20 ring members (i.e., the number of atoms making up the ring backbone, including carbon atoms and heteroatoms), although the present definition also covers the occurrence of the term “heterocyclyl” where no numerical range is designated. The heterocyclyl group may also be a medium size heterocyclyl having 3 to 10 ring members. The heterocyclyl group could also be a heterocyclyl having 3 to 6 ring members. The heterocyclyl group may be designated as “3-6 membered heterocyclyl” or similar designations. In preferred six membered monocyclic heterocyclyls, the heteroatom(s) are selected from one up to three of O, N or S, and in preferred five membered monocyclic heterocyclyls, the heteroatom(s) are selected from one or two heteroatoms selected from O, N, or S. Examples of heterocyclyl rings include, but are not limited to, azepinyl, acridinyl, carbazolyl, cinnolinyl, dioxolanyl, imidazoliny, imidazolidinyl, morpholinyl, oxiranyl, oxepanyl, thiepanyl, piperidinyl, piperazinyl, dioxopiperazinyl, pyrrolidinyl, pyrrolidonyl, pyrrolidionyl, 4-piperidonyl, pyrazolinyl, pyrazolidinyl, 1,3-dioxinyl, 1,3-dioxanyl, 1,4-dioxinyl, 1,4-dioxanyl, 1,3-oxathianyl, 1,4-oxathiinyl, 1,4-oxathianyl, 2*H*-1,2-oxazinyl, trioxanyl, hexahydro-1,3,5-triazinyl, 1,3-dioxolyl, 1,3-dioxolanyl, 1,3-dithiolyl, 1,3-dithiolanyl, isoxazoliny, isoxazolidinyl, oxazoliny, oxazolidinyl, oxazolidinonyl, thiazoliny, thiazolidinyl, 1,3-oxathiolanyl, indoliny, isoindoliny, tetrahydrofuranyl, tetrahydropyranyl, tetrahydrothiophenyl, tetrahydrothiopyranyl, tetrahydro-1,4-thiazinyl, thiamorpholiny, dihydrobenzofuranyl, benzimidazolidinyl, and tetrahydroquinoline.

[0041] As used herein, “(aryl)alkyl” refer to an aryl group, as defined above, connected, as a substituent, via an alkylene group, as described above. The alkylene and aryl group of an aralkyl may be substituted or unsubstituted. Examples include but are not limited to benzyl, 2-phenylalkyl, 3-phenylalkyl, and naphthylalkyl. In some embodiments, the alkylene is an unsubstituted straight chain containing 1, 2, 3, 4, 5, or 6 methylene unit(s).

[0042] As used herein, “(heteroaryl)alkyl” refer to a heteroaryl group, as defined above, connected, as a substituent, via an alkylene group, as defined above. The alkylene and heteroaryl group of heteroarylalkyl may be substituted or unsubstituted. Examples include but are not limited to 2-thienylalkyl, 3-thienylalkyl, furylalkyl, thienylalkyl, pyrrolylalkyl, pyridylalkyl, isoxazolylalkyl, and imidazolylalkyl, and their benzo-fused analogs. In some embodiments, the alkylene is an unsubstituted straight chain containing 1, 2, 3, 4, 5, or 6 methylene unit(s).

[0043] As used herein, “(heterocycl)alkyl” refer to a heterocyclic or a heterocycl group, as defined above, connected, as a substituent, via an alkylene group, as defined above. The alkylene and heterocycl groups of a (heterocycl)alkyl may be substituted or unsubstituted. Examples include but are not limited to (tetrahydro-2H-pyran-4-yl)methyl, (piperidin-4-yl)ethyl, (piperidin-4-yl)propyl, (tetrahydro-2H-thiopyran-4-yl)methyl, and (1,3-thiazinan-4-yl)methyl. In some embodiments, the alkylene is an unsubstituted straight chain containing 1, 2, 3, 4, 5, or 6 methylene unit(s).

[0044] As used herein, “(carbocycl)alkyl” refer to a carbocycl group (as defined herein) connected, as a substituent, via an alkylene group. Examples include but are not limited to cyclopropylmethyl, cyclobutylmethyl, cyclopentylethyl, and cyclohexylpropyl. In some embodiments, the alkylene is an unsubstituted straight chain containing 1, 2, 3, 4, 5, or 6 methylene unit(s).

[0045] As used herein, “alkoxyalkyl” or “(alkoxy)alkyl” refers to an alkoxy group connected via an alkylene group, such as C₂-C₈ alkoxyalkyl, or (C₁-C₆ alkoxy)C₁-C₆ alkyl, for example, -(CH₂)₁₋₃-OCH₃.

[0046] As used herein, “-O-alkoxyalkyl” or “-O-(alkoxy)alkyl” refers to an alkoxy group connected via an -O-(alkylene) group, such as -O-(C₁-C₆ alkoxy)C₁-C₆ alkyl, for example, -O-(CH₂)₁₋₃-OCH₃.

[0047] As used herein, “haloalkyl” refers to an alkyl group in which one or more of the hydrogen atoms are replaced by a halogen (*e.g.*, mono-haloalkyl, di-haloalkyl, and tri-haloalkyl). Such groups include but are not limited to, chloromethyl, fluoromethyl, difluoromethyl, trifluoromethyl and 1-chloro-2-fluoromethyl, 2-fluoroisobutyl. A haloalkyl may be substituted or unsubstituted.

[0048] As used herein, “haloalkoxy” refers to an alkoxy group in which one or more of the hydrogen atoms are replaced by a halogen (*e.g.*, mono-haloalkoxy, di-haloalkoxy and tri-haloalkoxy). Such groups include but are not limited to, chloromethoxy, fluoromethoxy, difluoromethoxy, trifluoromethoxy and 1-chloro-2-fluoromethoxy, 2-fluoroisobutoxy. A haloalkoxy may be substituted or unsubstituted.

[0049] An “amino” group refers to a $-NH_2$ group. The term “mono-substituted amino group” as used herein refers to an amino ($-NH_2$) group where one of the hydrogen atom is replaced by a substituent. The term “di-substituted amino group” as used herein refers to an amino ($-NH_2$) group where each of the two hydrogen atoms is replaced by a substituent. The term “optionally substituted amino,” as used herein refer to a $-NR_A R_B$ group where R_A and R_B are independently hydrogen, alkyl, cycloalkyl, aryl, heteroaryl, heterocyclyl, aralkyl, or heterocyclyl(alkyl), as defined herein.

[0050] An “O-carboxy” group refers to a “ $-OC(=O)R$ ” group in which R is selected from hydrogen, C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₇ carbocyclyl, C₆-C₁₀ aryl, 5-10 membered heteroaryl, and 3-10 membered heterocyclyl, as defined herein.

[0051] A “C-carboxy” group refers to a “ $-C(=O)OR$ ” group in which R is selected from the group consisting of hydrogen, C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₇ carbocyclyl, C₆-C₁₀ aryl, 5-10 membered heteroaryl, and 3-10 membered heterocyclyl, as defined herein. A non-limiting example includes carboxyl (i.e., $-C(=O)OH$).

[0052] A “sulfonyl” group refers to an “ $-SO_2R$ ” group in which R is selected from hydrogen, C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₇ carbocyclyl, C₆-C₁₀ aryl, 5-10 membered heteroaryl, and 3-10 membered heterocyclyl, as defined herein.

[0053] A “sulfinio” group refers to a “ $-S(=O)OH$ ” group.

[0054] A “sulfo” group refers to a “ $-S(=O)_2OH$ ” or “ $-SO_3H$ ” group.

[0055] A “sulfonate” group refers to a “ $-SO_3^-$ ” group.

[0056] A “sulfate” group refers to “ $-SO_4^-$ ” group.

[0057] A “S-sulfonamido” group refers to a “ $-SO_2NR_A R_B$ ” group in which R_A and R_B are each independently selected from hydrogen, C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₇ carbocyclyl, C₆-C₁₀ aryl, 5-10 membered heteroaryl, and 3-10 membered heterocyclyl, as defined herein.

[0058] An “N-sulfonamido” group refers to a “ $-N(R_A)SO_2R_B$ ” group in which R_A and R_B are each independently selected from hydrogen, C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₇ carbocyclyl, C₆-C₁₀ aryl, 5-10 membered heteroaryl, and 3-10 membered heterocyclyl, as defined herein.

[0059] A “C-amido” group refers to a “ $-C(=O)NR_A R_B$ ” group in which R_A and R_B are each independently selected from hydrogen, C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₇ carbocyclyl, C₆-C₁₀ aryl, 5-10 membered heteroaryl, and 3-10 membered heterocyclyl, as defined herein.

[0060] An “N-amido” group refers to a “ $-N(R_A)C(=O)R_B$ ” group in which R_A and R_B are each independently selected from hydrogen, C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₇

carbocyclyl, C₆-C₁₀ aryl, 5-10 membered heteroaryl, and 3-10 membered heterocyclyl, as defined herein.

[0061] An “O-carbamyl” group refers to a “-OC(=O)N(R_AR_B)” group in which R_A and R_B can be the same as defined with respect to S-sulfonamido. An O-carbamyl may be substituted or unsubstituted.

[0062] An “N-carbamyl” group refers to an “-ROC(=O)N(R_A)-” group in which R and R_A can be the same as defined with respect to N-sulfonamido. An N-carbamyl may be substituted or unsubstituted.

[0063] An “O-thiocarbamyl” group refers to a “-OC(=S)-N(R_AR_B)” group in which R_A and R_B can be the same as defined with respect to S-sulfonamido. An O-thiocarbamyl may be substituted or unsubstituted.

[0064] An “N-thiocarbamyl” group refers to an “-ROC(=S)N(R_A)-” group in which R and R_A can be the same as defined with respect to N-sulfonamido. An N-thiocarbamyl may be substituted or unsubstituted.

[0065] The term “alkylamino” or “(alkyl)amino” refers to an amino group wherein one or both hydrogen is replaced by an alkyl group.

[0066] An “(alkoxy)alkyl” group refers to an alkoxy group connected via an alkylene group, such as a “(C₁-C₆ alkoxy) C₁-C₆ alkyl” and the like.

[0067] The term “hydroxy” as used herein refers to a -OH group.

[0068] The term “cyano” group as used herein refers to a “-CN” group.

[0069] The term “azido” as used herein refers to a -N₃ group.

[0070] When a group is described as “optionally substituted” it may be either unsubstituted or substituted. Likewise, when a group is described as being “substituted,” the substituent may be selected from one or more of the indicated substituents. As used herein, a substituted group is derived from the unsubstituted parent group in which there has been an exchange of one or more hydrogen atoms for another atom or group. Unless otherwise indicated, when a group is deemed to be “substituted,” it is meant that the group is substituted with one or more substituents independently selected from C₁-C₆ alkyl, C₁-C₆ alkenyl, C₁-C₆ alkynyl, C₁-C₆ heteroalkyl, C₃-C₇ carbocyclyl (optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), C₃-C₇carbocyclyl-C₁-C₆-alkyl (optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), 3-10 membered heterocyclyl (optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), 3-10 membered heterocyclyl-C₁-C₆-alkyl (optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), aryl (optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), (aryl)C₁-C₆ alkyl

(optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), 5-10 membered heteroaryl (optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), (5-10 membered heteroaryl)C₁-C₆ alkyl (optionally substituted with halo, C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkyl, and C₁-C₆ haloalkoxy), halo, -CN, hydroxy, C₁-C₆ alkoxy, (C₁-C₆ alkoxy)C₁-C₆ alkyl, -O(C₁-C₆ alkoxy)C₁-C₆ alkyl; (C₁-C₆ haloalkoxy)C₁-C₆ alkyl; -O(C₁-C₆ haloalkoxy)C₁-C₆ alkyl; aryloxy, sulfhydryl (mercapto), halo(C₁-C₆)alkyl (e.g., -CF₃), halo(C₁-C₆)alkoxy (e.g., -OCF₃), C₁-C₆ alkylthio, arylthio, amino, amino(C₁-C₆)alkyl, nitro, O-carbamyl, N-carbamyl, O-thiocarbamyl, N-thiocarbamyl, C-amido, N-amido, S-sulfonamido, N-sulfonamido, C-carboxy, O-carboxy, acyl, cyanato, isocyanato, thiocyanato, isothiocyanato, sulfinyl, sulfonyl, -SO₃H, sulfonate, sulfate, sulfinio, -OSO₂C₁₋₄alkyl, monophosphate, diphosphate, triphosphate, and oxo (=O). Wherever a group is described as “optionally substituted” that group can be substituted with the above substituents.

[0071] As understood by one of ordinary skill in the art, a compound described herein may exist in ionized form, e.g., -CO₂⁻, -SO₃⁻ or -O-SO₃⁻. If a compound contains a positively or negatively charged substituent group, for example, -SO₃⁻, it may also contain a negatively or positively charged counterion such that the compound as a whole is neutral. In other aspects, the compound may exist in a salt form, where the counterion is provided by a conjugate acid or base.

[0072] As used herein, a “nucleotide” includes a nitrogen containing heterocyclic base, a sugar, and one or more phosphate groups. They are monomeric units of a nucleic acid sequence. In RNA, the sugar is a ribose, and in DNA a deoxyribose, *i.e.* a sugar lacking a hydroxy group that is present in ribose. The nitrogen containing heterocyclic base can be purine or pyrimidine base. Purine bases include adenine (A) and guanine (G), and modified derivatives or analogs thereof, such as 7-deaza adenine or 7-deaza guanine. Pyrimidine bases include cytosine (C), thymine (T), and uracil (U), and modified derivatives or analogs thereof. The C-1 atom of deoxyribose is bonded to N-1 of a pyrimidine or N-9 of a purine.

[0073] As used herein, a “nucleoside” is structurally similar to a nucleotide, but is missing the phosphate moieties. An example of a nucleoside analogue would be one in which the label is linked to the base and there is no phosphate group attached to the sugar molecule. The term “nucleoside” is used herein in its ordinary sense as understood by those skilled in the art. Examples include, but are not limited to, a ribonucleoside comprising a ribose moiety and a deoxyribonucleoside comprising a deoxyribose moiety. A modified pentose moiety is a pentose moiety in which an oxygen atom has been replaced with a carbon and/or a carbon has been replaced with a sulfur or an oxygen atom. A “nucleoside” is a monomer that can have a substituted

base and/or sugar moiety. Additionally, a nucleoside can be incorporated into larger DNA and/or RNA polymers and oligomers.

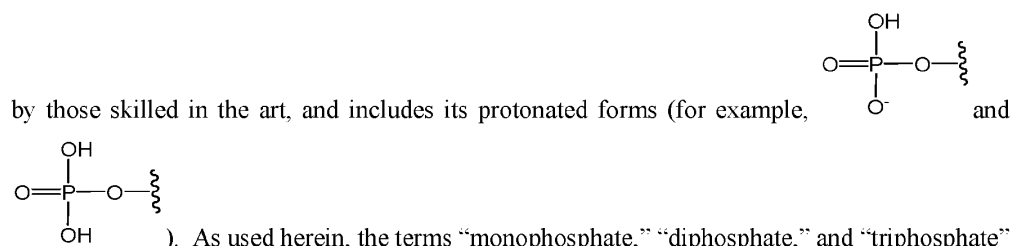
[0074] The term “purine base” is used herein in its ordinary sense as understood by those skilled in the art, and includes its tautomers. Similarly, the term “pyrimidine base” is used herein in its ordinary sense as understood by those skilled in the art, and includes its tautomers. A non-limiting list of optionally substituted purine-bases includes purine, adenine, guanine, deazapurine, 7-deaza adenine, 7-deaza guanine, hypoxanthine, xanthine, alloxanthine, 7-alkylguanine (e.g., 7-methylguanine), theobromine, caffeine, uric acid and isoguanine. Examples of pyrimidine bases include, but are not limited to, cytosine, thymine, uracil, 5,6-dihydrouracil and 5-alkylcytosine (e.g., 5-methylcytosine).

[0075] As used herein, when an oligonucleotide or polynucleotide is described as “comprising” or “incorporating” a nucleoside or nucleotide described herein, it means that the nucleoside or nucleotide described herein forms a covalent bond with the oligonucleotide or polynucleotide. Similarly, when a nucleoside or nucleotide is described as part of an oligonucleotide or polynucleotide, such as “incorporated into” an oligonucleotide or polynucleotide, it means that the nucleoside or nucleotide described herein forms a covalent bond with the oligonucleotide or polynucleotide. In some such embodiments, the covalent bond is formed between a 3' hydroxy group of the oligonucleotide or polynucleotide with the 5' phosphate group of a nucleotide described herein as a phosphodiester bond between the 3' carbon atom of the oligonucleotide or polynucleotide and the 5' carbon atom of the nucleotide.

[0076] As used herein, the term “cleavable linker” is not meant to imply that the whole linker is required to be removed. The cleavage site can be located at a position on the linker that ensures that part of the linker remains attached to the detectable label and/or nucleoside or nucleotide moiety after cleavage.

[0077] As used herein, “derivative” or “analog” means a synthetic nucleotide or nucleoside derivative having modified base moieties and/or modified sugar moieties. Such derivatives and analogs are discussed in, e.g., Scheit, *Nucleotide Analogs* (John Wiley & Son, 1980) and Uhlman *et al.*, *Chemical Reviews* 90:543-584, 1990. Nucleotide analogs can also comprise modified phosphodiester linkages, including phosphorothioate, phosphorodithioate, alkyl-phosphonate, phosphoranilidate and phosphoramidate linkages. “Derivative,” “analog” and “modified” as used herein, may be used interchangeably, and are encompassed by the terms “nucleotide” and “nucleoside” defined herein.

[0078] As used herein, the term “phosphate” is used in its ordinary sense as understood



As used herein, the terms “monophosphate,” “diphosphate,” and “triphosphate” are used in their ordinary sense as understood by those skilled in the art, and include protonated forms.

[0079] The terms “protecting group” and “protecting groups” as used herein refer to any atom or group of atoms that is added to a molecule in order to prevent existing groups in the molecule from undergoing unwanted chemical reactions. Sometimes, “protecting group” and “blocking group” can be used interchangeably.

[0080] As used herein, the term “phasing” refers to a phenomenon in SBS that is caused by incomplete removal of the 3' terminators and fluorophores, and failure to complete the incorporation of a portion of DNA strands within clusters by polymerases at a given sequencing cycle. Pre-phasing is caused by the incorporation of nucleotides without effective 3' terminators, wherein the incorporation event goes 1 cycle ahead due to a termination failure. Phasing and pre-phasing cause the measured signal intensities for a specific cycle to consist of the signal from the current cycle as well as noise from the preceding and following cycles. As the number of cycles increases, the fraction of sequences per cluster affected by phasing and pre-phasing increases, hampering the identification of the correct base. Pre-phasing can be caused by the presence of a trace amount of unprotected or unblocked 3'-OH nucleotides during sequencing by synthesis (SBS). The unprotected 3'-OH nucleotides could be generated during the manufacturing processes or possibly during the storage and reagent handling processes. Accordingly, the discovery of nucleotide analogues which decrease the incidence of pre-phasing is surprising and provides a great advantage in SBS applications over existing nucleotide analogues. For example, the nucleotide analogues provided can result in faster SBS cycle time, lower phasing and pre-phasing values, and longer sequencing read lengths.

Sequencing Methods Utilizing Nucleotides with 3' Allyl Blocking Group

[0081] Some embodiments of the present disclosure relate to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

- (i) the solid support comprises a plurality of different target polynucleotides; and
- (ii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C and T or U (e.g., dATP, dCTP, dGTP and dTTP or dUTP) and under conditions suitable for DNA polymerase-mediated primer extension, wherein each nucleotide comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;

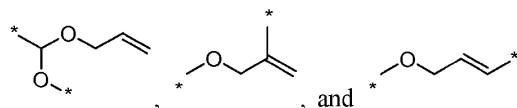
(c) imaging the solid support to determine the identity of incorporated nucleotides (e.g., by performing one or more fluorescent measurements of the extended copy polynucleotides);

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove 3' allyl blocking groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support;

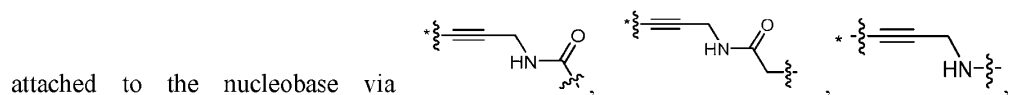
(e) contacting said solid support with an aqueous wash solution; and

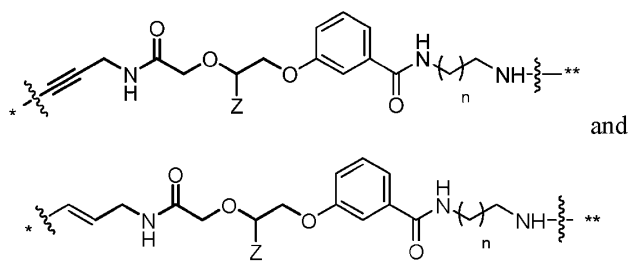
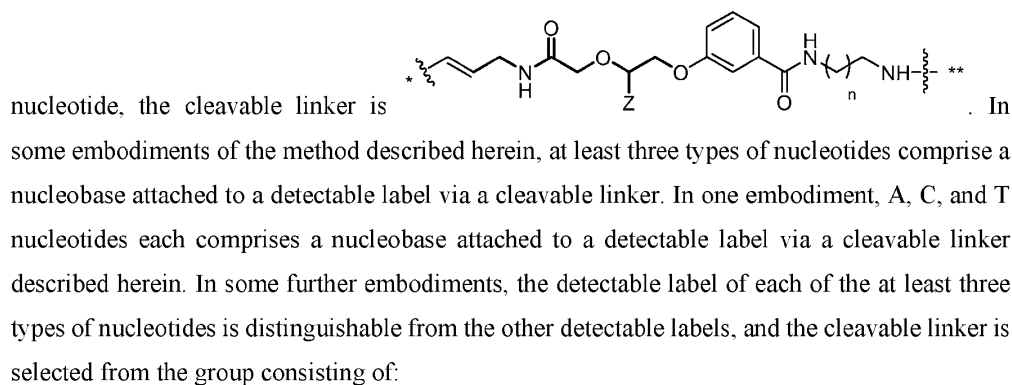
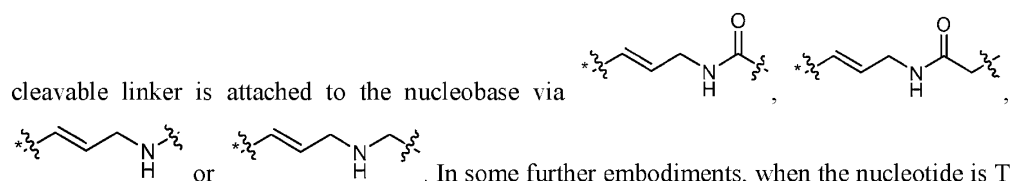
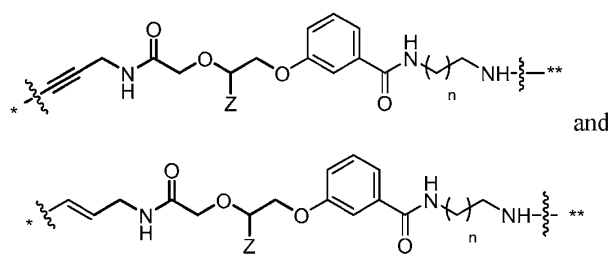
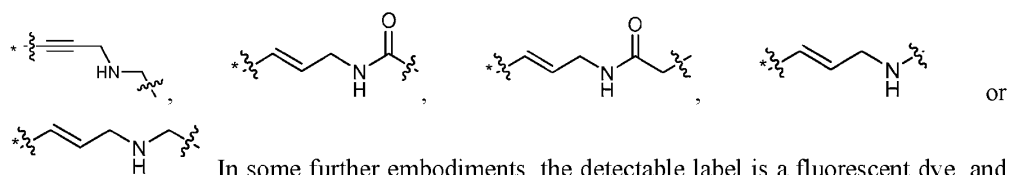
(f) repeating steps (b)-(e) to determine target polynucleotide sequences. In some embodiments of the method described herein, the aqueous solution comprising DNA polymerase in step (b) further comprises a palladium scavenger. In further embodiments, the palladium scavenger in step (b) is a Pd(0) scavenger described herein. In some embodiments, the aqueous wash solution in step (e) further comprises a palladium scavenger. In some further embodiments, the palladium scavenger in step (e) is a Pd(II) scavenger described herein. In some further embodiments, the aqueous deblocking solution further comprises ascorbate, such as sodium ascorbate. In some embodiments, the solid support comprises at least 500,000, 1,000,000, 2,000,000, 3,000,000, 4,000,000, or 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides.

[0082] In some embodiments of the method described herein, at least one type of nucleotide comprises a base attached to a detectable label via a cleavable linker. In some further embodiments, the cleavable linker comprises a moiety selected from the group consisting of:



, and * indicates where the moiety is connected to the remainder of the nucleotide. In some further embodiments, the cleavable linker is





wherein Z is $-O-CH_2-CH=CH_2$; n is an integer of 1, 2, 3, 4 or 5; * indicates the attachment point of the cleavable linker to the base; and ** indicates the attachment point of the cleavable linker to the detectable label. In other embodiments, each of the four types of nucleotides is unlabeled.

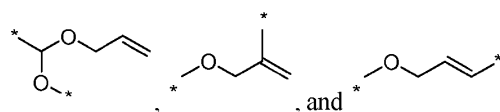
[0083] Some other embodiments of the present disclosure relate to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

- (i) the solid support comprises a plurality of different target polynucleotides; and
- (ii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C, and T or U (e.g., dATP, dCTP, dGTP and dTTP or dUTP) under conditions suitable for DNA polymerase-mediated primer extension, wherein:

- (i) each of at least three types of nucleotides independently comprises a base that is attached to a detectable label via a cleavable linker, and the cleavable linker comprises a moiety selected from the group consisting of:



is connected to the remainder of the nucleotide;

- (ii) each nucleotide comprises a 2' deoxyribose moiety with a 3' allyl group attached to the 3' oxygen atom; and

- (iii) the DNA polymerase is an altered archaeal DNA polymerase;

(c) contacting the solid support with a solution comprising one or more radical scavengers and imaging the solid support to determine the identity of incorporated nucleotides (e.g., by performing one or more fluorescent measurements of the extended copy polynucleotides);

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl blocking groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers;

(e) contacting said solid support with an aqueous wash solution comprising a palladium scavenger; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences. In some embodiments, the palladium scavenger in the aqueous wash solution of step (e) is a Pd(II) scavenger as described herein. In some further embodiments, the aqueous solution comprising DNA polymerase in step (b) further comprises a palladium scavenger. In some such embodiments, the palladium scavenger in step (b) is a Pd(0) scavenger described herein. In some embodiments, the solid support comprises at least 500,000, 1,000,000, 2,000,000, 3,000,000, 4,000,000, or 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides.

[0084] Some additional embodiments of the present disclosure relate to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

- (i) the solid support comprises a plurality of different target polynucleotides; and
- (ii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C, and T or U (e.g., dATP, dCTP, dGTP and dTTP or dUTP) under conditions suitable for DNA polymerase-mediated primer extension, wherein:

- (i) at least one of the nucleotides comprise a base that is attached to a detectable label via a cleavable linker; and
- (ii) the nucleotides each comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;

(c) contacting the solid support with a solution comprising one or more radical scavengers and imaging the solid support to determine the identity of incorporated nucleotides (e.g., by performing one or more fluorescent measurements of the extended copy polynucleotides);

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers;

(e) contacting said solid support with an aqueous wash solution; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences. In some embodiments, the aqueous wash solution in step (e) further comprises a palladium scavenger. In some embodiments, the solid support comprises at least 500,000, 1,000,000, 2,000,000,

3,000,000, 4,000,000, or 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides. In some such embodiments, the palladium scavenger in the aqueous wash solution of step (e) is a Pd(II) scavenger as described herein. In some further embodiments, the aqueous solution comprising DNA polymerase in step (b) further comprises a palladium scavenger. In some such embodiments, the palladium scavenger in step (b) is a Pd(0) scavenger described herein.

Post-Incorporation Labeling SBS Workflow

[0085] Another aspect of the present disclosure relates to an alternative sequencing by synthesis method in which at least one labeling reagent is introduced after the incorporation of unlabeled nucleotides. In particular, the disclosure relates to a method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

- (i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;
- (ii) the solid support comprises a plurality of different target polynucleotides; and
- (iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous incorporation mixture comprising DNA polymerase and one or more of four types of nucleotides A, G, C, and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein:

- (i) the nucleotides each comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;
- (ii) at least one type of nucleotide is unlabeled; and
- (iii) the first type of unlabeled nucleotides comprises a first functional moiety;

(c) contacting the extended copy polynucleotides with an aqueous labeling mixture comprising a first labeling reagent, wherein the first labeling reagent comprises one or more first detectable labels and a first binding moiety that is capable of specific binding to the first functional moiety of the first type of unlabeled nucleotide;

(d) imaging the solid support and performing one or more fluorescent measurements (i.e., to determine the identity of incorporated nucleotides);

(e) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support;

(f) contacting said solid support with an aqueous wash solution; and

(g) repeating steps (b)-(f) to determine target polynucleotide sequences. In some embodiments, at least two types of nucleotides are unlabeled.

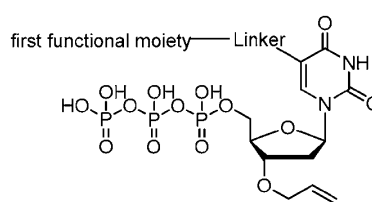
[0086] In some embodiments, step (c) is performed under conditions where the first labeling reagent binds specifically to the incorporated unlabeled first type of nucleotides to provide labeled extended copy polynucleotides. In some embodiments, the first functional moiety of the first type of unlabeled nucleotide is bound to the first labeling reagent by covalent bonding, optionally via the cleavable linker as described herein. In some other embodiments, the first functional moiety of the first type of unlabeled nucleotide is bound to the first labeling reagent by noncovalent interaction, optionally via the cleavable linker as described herein.

[0087] In some embodiments, each of the four types of nucleotides in the aqueous incorporation mixture is unlabeled, the second type of unlabeled nucleotides comprises a second functional moiety, wherein the aqueous labeling mixture comprises a second labeling reagent, and the second labeling reagent comprises one or more second detectable labels and a second binding moiety that is capable of specific binding to the second functional moiety of the second type of unlabeled nucleotides. In some such embodiments, step (c) is performed under conditions where the first labeling reagent binds specifically to the incorporated unlabeled first type of nucleotides, and the second labeling reagent binds specifically to the incorporated unlabeled second type of nucleotides to provide labeled extended copy polynucleotides. In some such embodiments, the first functional moiety of the first type of unlabeled nucleotide is bound to the first labeling reagent by covalent bonding, optionally via a cleavable linker as described herein. In some other embodiments, the second functional moiety of the first type of unlabeled nucleotide is bound to the second labeling reagent by noncovalent interaction, optionally via the cleavable linker as described herein. In some such embodiments, the third type of unlabeled nucleotides comprises a third functional moiety, wherein the aqueous labeling mixture comprises a third labeling reagent, and the third labeling reagent comprises one or more third detectable labels and a third binding moiety that is capable of specific binding to the third functional moiety of the third type of unlabeled nucleotides. In some such embodiments, step (c) is performed under conditions where the first labeling reagent binds specifically to the incorporated unlabeled first type of nucleotides, the second labeling reagent binds specifically to the incorporated unlabeled second type of nucleotides, and the third labeling reagent binds specifically to the incorporated unlabeled third

type of nucleotides to provide labeled extended copy polynucleotides. In some other embodiments, the third type of unlabeled nucleotide comprises a mixture of the third type of unlabeled nucleotides comprising the first functional moiety and the third type of unlabeled nucleotides comprising the second functional moiety, and wherein both the first labeling reagent and the second labeling reagent are capable of specific binding to the third type of unlabeled nucleotides.

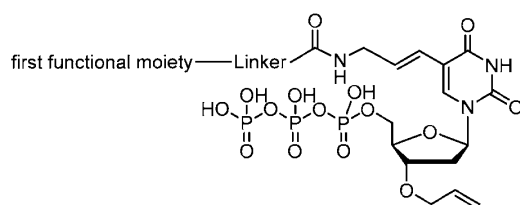
[0088] In some embodiments of the post-incorporation labeling method described herein, the fourth type of unlabeled nucleotides is not capable of specific binding with any of the first, second, or third labeling reagent.

[0089] In some embodiments of the post-incorporation labeling method described

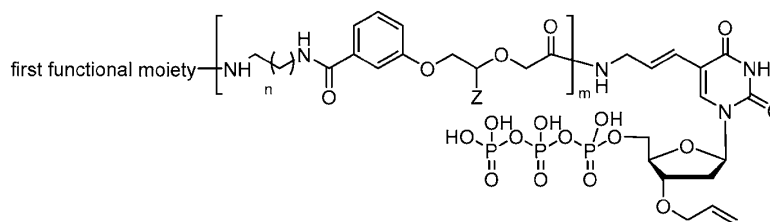


herein, the T nucleotide has the structure:

embodiments, the T nucleotide has the structure:



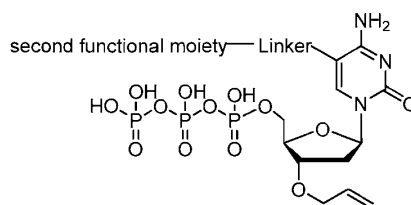
. In further embodiments, the T nucleotide has



the structure:

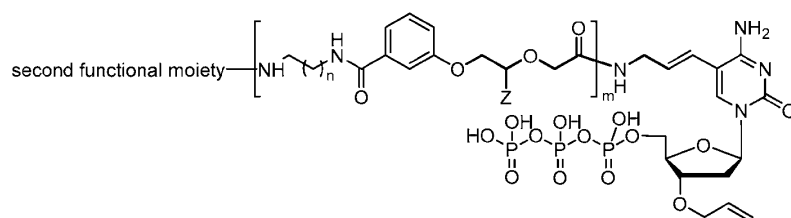
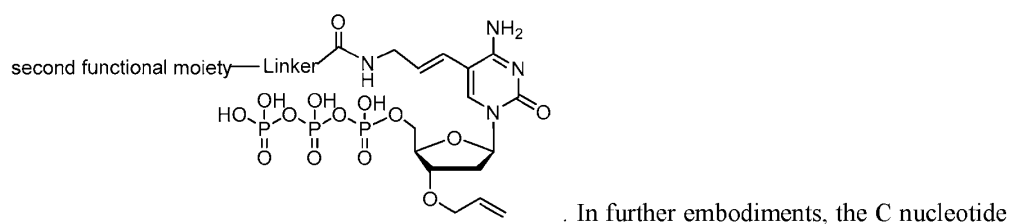
wherein Z is $-O-CH_2-CH=CH_2$, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

[0090] In some embodiments of the post-incorporation labeling method described



herein, the C nucleotide has the structure:

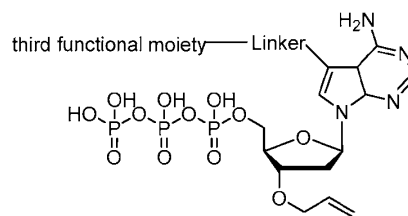
further embodiments, the C nucleotide has the structure:



has the structure:

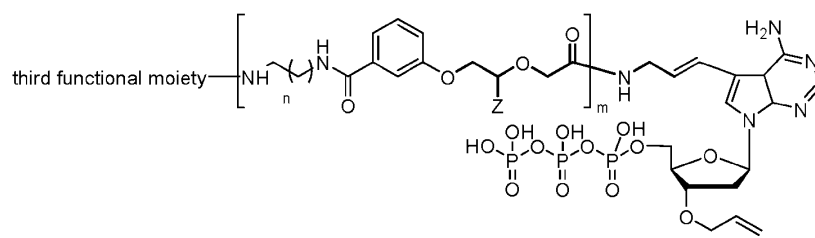
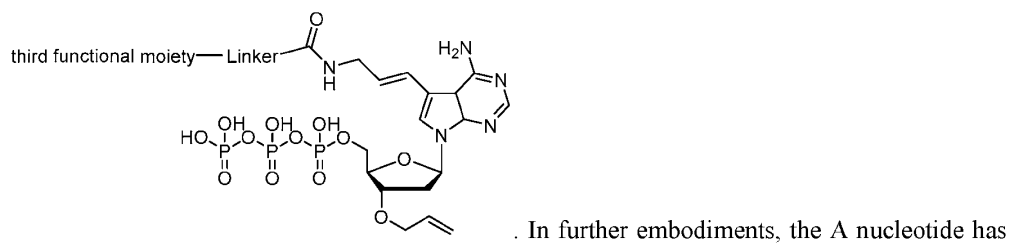
wherein Z is -O-CH₂-CH=CH₂, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

[0091] In some embodiments of the post-incorporation labeling method described



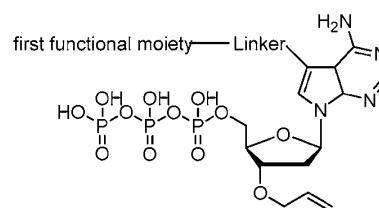
herein, the A nucleotide has the structure:

further embodiments, the A nucleotide has the structure:

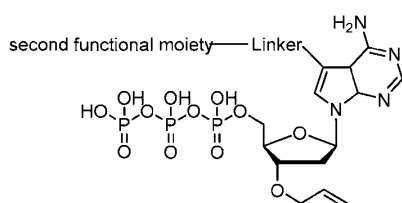


the structure:

wherein Z is -O-CH₂-CH=CH₂, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

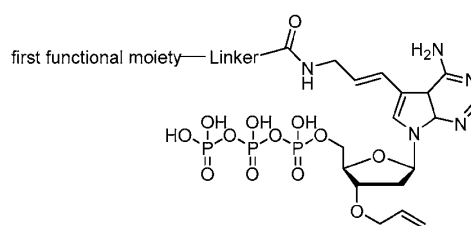


In other embodiments, the A nucleotide has the structure:



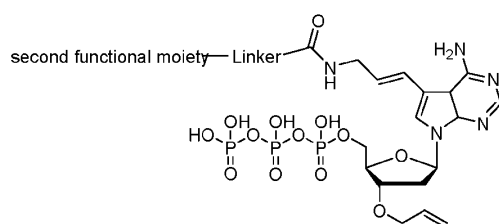
and

. In further embodiments, the A nucleotide has the

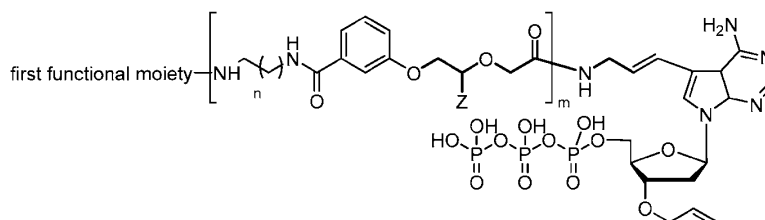


structure:

and

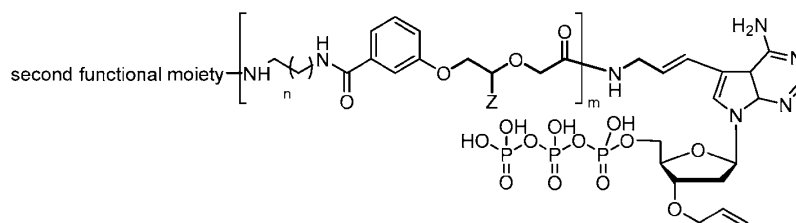


. In further embodiments, the A nucleotide has



the structure:

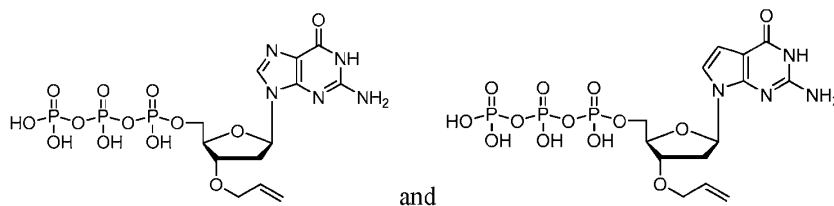
and



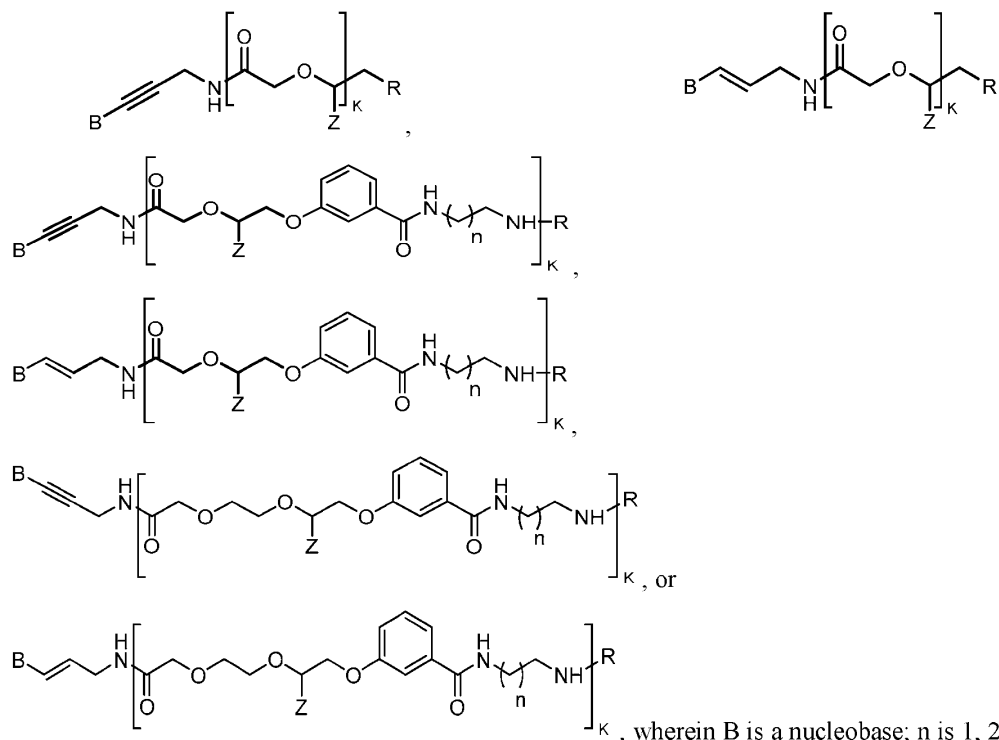
, wherein Z is -O-

CH₂-CH=CH₂, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

[0092] In some embodiments, the G nucleotide is unlabeled. In some further embodiments, G nucleotide has a structure selected from the group consisting of:



[0093] Additional examples of linkers are disclosed in U.S. Publication No. 2020/0216891 A1, which is incorporated by reference in its entirety:



, wherein B is a nucleobase; n is 1, 2, 3, 4, 5; k is 1; Z is $-N_3$ (azido), $-O-C_1-C_6$ alkyl, $-O-C_2-C_6$ alkenyl, or $-O-C_2-C_6$ alkynyl; and R comprises a detectable label or a binding moiety for post-incorporation labeling, which may contain additional linker and/or spacer structure. One of ordinary skill in the art understands that the detectable label or the binding moiety described herein is covalently bound to the linker by reacting a functional group of the binding moiety containing compound (e.g., carboxyl) with a functional group of the linker (e.g., amino) to form an amide bond. In one embodiment, the cleavable linker comprises ("AOL" linker moiety) where Z is $-O$ -allyl. For the purpose of the present disclosure, the nucleotide may contain multiple cleavable linkers repeating units (e.g., k is 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10).

[0094] In particular embodiments, the unlabeled nucleotide may be enzymatically incorporable and enzymatically extendable. Accordingly, a linker moiety may be of sufficient length to connect the nucleotide to the compound such that the compound does not significantly interfere with the overall binding and recognition of the nucleotide by a nucleic acid replication enzyme. Thus, the linker can also comprise a spacer unit, such as one or more PEG unit(s) $(-\text{OCH}_2\text{CH}_2-)_n$, where n is an integer of 1-20, for example 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15. The spacer distances, for example, the nucleotide base from a cleavage site or label.

[0095] In some embodiments of any of the methods described herein, said contacting the solid support with a deblocking solution is performed for about 1 to 20 seconds, about 2 to 10 seconds, about 3 to 8 seconds, or about 4 to 5 seconds. For example, the contacting the solid support with a deblocking solution of step (d) or step (e) is performed for about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 seconds, or a range defined by any two of the preceding values. In some further embodiments, said contacting the solid support with a deblocking solution is performed via continuous flow without pausing to incubate.

[0096] In some embodiments of any of the method described herein, the sequencing cycles (e.g., steps (b)-(e) or steps (b)-(f)) are repeated at least about 20 times, 30 times, or 50 times. In other embodiments, steps (b)-(e) or steps (b)-(f) are repeated at least about 100 times, 150 times, 200 times, 250 times, 300 times, 350 times, 400 times, 450 times or 500 times. In some embodiments, after about 50 repeated cycles the pre-phasing value is less than 0.18. In some embodiments, after about 50 repeated cycles the phasing value is less than 0.18. In further embodiments, after about 50 repeated cycles the pre-phasing or phasing value is less than 0.07. In some further embodiments, after about 100 repeated cycles the pre-phasing value is less than 0.10 and the phasing value is less than 0.10. In some further embodiments, after about 150 repeated cycles the pre-phasing value is less than 0.25 and the phasing value is less than 0.25. In some further embodiments, after about 150 repeated cycles the pre-phasing value is less than 0.2 and the phasing value is less than 0.2. In some further embodiments, after about 150 repeated cycles the pre-phasing value is less than 0.15 and the phasing value is less than 0.15. In some further embodiments, after about 150 repeated cycles the pre-phasing value is less than 0.1 and the phasing value is less than 0.1.

[0097] In some embodiments of any of the method described herein, the DNA polymerase is an altered family B archaeal DNA polymerase comprising a 3-amino acid region that is functionally equivalent or homologous to amino acids 408-410 in 9° N DNA polymerase, wherein the first amino acid of the 3-amino acid region is an amino acid selected from the group consisting of isoleucine (I), alanine (A), valine (V), and serine (S); the second amino acid of the 3-amino acid region is an amino acid selected from the group consisting of alanine (A) and glycine

(G); and the third amino acid of the 3-amino acid region is an amino acid selected from the group consisting of alanine (A), isoleucine (I), valine (V), leucine (L), threonine (T), and proline (P).

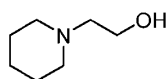
Palladium Catalysts

[0098] In any embodiments of any of the method described herein, the Pd catalyst used for removing or cleaving the 3' allyl blocking group described herein is water soluble. In some such embodiments, the Pd catalyst is the active Pd(0) form. In some instances, the Pd(0) catalyst may be generated *in situ* from reduction of a Pd complex or Pd precatalyst (e.g., a Pd(II) complex) by reagents such as alkenes, alcohols, amines, phosphines, or metal hydrides. Suitable Pd sources include Pd(CH₃CN)₂Cl₂, [Pd(Allyl)Cl]₂, [Pd(Allyl)(THPP)]Cl, [Pd(Allyl)(THPP)₂]Cl, Na₂PdCl₄, K₂PdCl₄, Li₂PdCl₄, Pd(OAc)₂, Pd(PPh₃)₄, Pd(dba)₂, Pd(Acac)₂, PdCl₂(COD), Pd(TFA)₂, Na₂PdBr₄, K₂PdBr₄, PdCl₂, PdBr₂, and Pd(NO₃)₂. In one such embodiment, the Pd(0) complex is generated *in situ* from an organic or inorganic salt of palladate (II), for example, Na₂PdCl₄ or K₂PdCl₄. In another embodiment, the palladium source is allyl Pd(II) chloride dimer [(Allyl)PdCl]₂ or [PdCl(C₃H₅)]₂. In some embodiments, the Pd(0) catalyst is generated in an aqueous solution by mixing a Pd(II) complex with a water soluble phosphine, such as tris(hydroxyalkyl)phosphine. Suitable tris(hydroxyalkyl)phosphines include tris(hydroxypropyl)phosphine (THPP), tris(hydroxymethyl)phosphine (THMP), or tris(hydroxyethyl)phosphine (THEP), or combinations thereof. In one embodiment, the water soluble phosphine is THPP.

[0099] In some embodiments, the palladium catalyst is prepared by mixing [(Allyl)PdCl]₂ with THP *in situ*. The molar ratio of [(Allyl)PdCl]₂ and the THP may be about 1:1, 1:1.5, 1:2, 1:2.5, 1:3, 1:3.5, 1:4, 1:4.5, 1:5, 1:5.5, 1:6, 1:6.5, 1:7, 1:7.5, 1:8, 1:8.5, 1:9, 1:9.5 or 1:10. In one embodiment, the molar ratio of [(Allyl)PdCl]₂ to THP is 1:10. In some other embodiment, the palladium catalyst is prepared by mixing a water soluble Pd reagent such as Na₂PdCl₄ or K₂PdCl₄ with THP *in situ*. The molar ratio of Na₂PdCl₄ or K₂PdCl₄ and THP may be about 1:1, 1:1.5, 1:2, 1:2.5, 1:3, 1:3.5, 1:4, 1:4.5, 1:5, 1:5.5, 1:6, 1:6.5, 1:7, 1:7.5, 1:8, 1:8.5, 1:9, 1:9.5 or 1:10. In one embodiment, the molar ratio of Na₂PdCl₄ or K₂PdCl₄ to THP is about 1:3. In another embodiment, the molar ratio of Na₂PdCl₄ or K₂PdCl₄ to THP is about 1:3.5. In yet another embodiment, the molar ratio of Na₂PdCl₄ or K₂PdCl₄ to THP is about 1:2.5.

[0100] The Pd complex and the water soluble phosphine for use in the cleavage step of the method described herein may be in a composition or a mixture, also called cleavage mix. In some further embodiments, the cleavage mix may contain one or more reducing agents, such as ascorbic acid or a salt thereof (e.g., sodium ascorbate), or a boron-containing reducing agent. In some further embodiments, the cleavage mix may contain one or more buffer reagents, such as

a primary amine, a secondary amine, a tertiary amine, a natural amino acid, a non-natural amino acid, a carbonate salt, a phosphate salt, or a borate salt, or combinations thereof. In some further embodiments, the buffer reagent comprises ethanolamine (EA), tris(hydroxymethyl)aminomethane (Tris), glycine, sodium carbonate, sodium phosphate, sodium borate, dimethylethanolamine (DMEA), diethylethanolamine (DEEA), N,N,N',N'-tetramethylethylenediamine (TMEDA), N,N,N',N'-tetraethylethylenediamine (TEEDA), or 2-piperidine ethanol (also known as (2-hydroxyethyl)piperidine, having the structure



), or combinations thereof. In one embodiment, the one or more buffer reagents comprise DEEA. In another embodiment, the one or more buffer reagents comprise (2-hydroxyethyl)piperidine. In another embodiment, the one or more buffer reagents contains one or more inorganic salts such as a carbonate salt, a phosphate salt, or a borate salt, or combinations thereof. In one embodiment, the inorganic salt is a sodium salt.

Reducing Agents

[0101] A reducing agent may be used to facilitate the formation of Pd(0) active species for a 3' deblocking reaction and/or cleavage of the linker in accordance with the present disclosure. In some embodiments, the reducing agent contains boron. For example, the reducing agent may be a borane, a borinic acid, a boronic acid, a boronic ester, or a borohydride. In some embodiments, the reducing agent is a borane or a boronic acid. Suitable boranes include diborane, triborane, tetraborane, pentaborane, hexaborane, heptaborane, octaborane, nonaborane, decaborane, borane tetrahydrofuran, borane dimethyl sulfide, catecholborane, dichlorophenylborane, borane diphenylphosphine complex, dicyclohexylidoborane, bromodimethylborane, diethylmethoxyborane, dichloromethyldiisopropoxyborane, bromodimethylborane, and monobromoborane methyl sulfide. Suitable boranes include amino-boranes. Suitable amine-boranes include pyridine-boranes, N-based heteroaromatic-borane complexes, aminetrihydridoboron (NH_3BH_3), borane-ammonia complex, borane dimethylamine, borane tert-butylamine, borane trimethylamine, borane isopropylamine, dichloro(diisopropylamino)borane, and borazine. Suitable boronic acids include boric acid, 4-hydroxyphenylboronic acid, tetrahydroxidoboron ($\text{B}_2(\text{OH})_4$), phenylboronic acid, 2-thienylboronic acid, methylboronic acid, *cis*-propenylboronic acid, and *trans*-propenylboronic acid. Suitable boronic esters include allylboronic acid pinacol ester, phenyl boronic acid trimethylene glycol ester, and diisopropoxymethylborane. Suitable borohydrides include sodium borohydride (Na_2BH_4), lithium borohydride, calcium borohydride, calcium borohydride bis(tetrahydrofuran), magnesium borohydride, potassium borohydride, potassium triethylborohydride, sodium triacetoxyborohydride, tetraethylammonium borohydride,

and tetrabutylammonium borohydride. In some embodiments, the concentration of the reducing agent (such as boron-containing reducing agent) in the first aqueous solution (also called cleavage mixture) is about 0.1 mM, 0.2 mM, 0.3 mM, 0.4 mM, 0.5 mM, 0.6 mM, 0.7 mM, 0.8 mM, 0.9 mM, 1 mM, 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM, 4 mM, 4.5 mM, 5 mM, 5.5 mM, 6 mM, 6.5 mM, 7 mM, 7.5 mM, 8 mM, 8.5 mM, 9 mM, 9.5 mM, 10 mM, 12.5 mM, 15 mM, 17.5 mM, 20 mM, 22.5 mM, 25 mM, 37.5 mM, 30 mM, 35 mM, 40 mM, 45 mM, 50 mM, 55 mM, 60 mM, 65 mM, 70 mM, or 75 mM, or a range defined by any two of the preceding values.

[0102] In some embodiments, the cleavage condition for the 3' blocking group is the same as the condition for cleaving the cleavable linker of the nucleotide. For example, the nucleotide may comprise a linker moiety that is the same as the 3' blocking group. In other embodiments, the cleavage condition for the 3' blocking group is different from the condition for cleaving the cleavable linker of the nucleotide.

Palladium Scavengers

Pd(0) Scavengers

[0103] Certain aspects of the present disclosure relate to employing alternative palladium scavengers in several steps of sequencing by synthesis, where at least one palladium scavenger comprises one or more allyl moieties. For example, -O-allyl, -S-allyl, -NR-allyl, or -N⁺RR'-allyl, or combinations thereof, wherein R is H, unsubstituted or substituted C₁-C₆ alkyl, unsubstituted or substituted C₂-C₆ alkenyl, unsubstituted or substituted C₂-C₆ alkynyl, unsubstituted or substituted C₆-C₁₀ aryl, unsubstituted or substituted 5 to 10 membered heteroaryl, unsubstituted or substituted C₃-C₁₀ carbocyclyl, or unsubstituted or substituted 5 to 10 membered heterocyclyl; and R' is H, unsubstituted C₁-C₆ alkyl or substituted C₁-C₆ alkyl.

[0104] The allyl containing Pd scavenger acts as a competitive substrate to consume any residual Pd(0) sticking on the nucleic acid (i.e., a Pd(0) scavenger). These palladium scavengers are described in WO 2022/243480, which is incorporated by reference in its entirety. The sequencing methods described herein substantially improve the sequencing metrics (e.g., reduce phasing and prephasing values) and may also reduce the sequencing time for each cycle by certain eliminating post-cleavage treatment step.

[0105] In some embodiments of any of the methods described herein, the Pd(0) scavenger comprises one or more allyl moieties is in the aqueous solution containing DNA polymerase and the nucleotides. In some instances, the aqueous solution in step (b) is also known as the incorporation mix (IMX). In some such embodiments, such palladium scavenger is compatible with the other sequencing reagents in the incorporation mix, which may also include a polymerase (such as DNA polymerase), in addition to the one or more different types of

nucleotides. In some such embodiments, the polymerase is a DNA polymerase, such as a mutant of 9°N polymerase (e.g., those disclosed in WO 2005/024010, which is incorporated by reference), for example, Pol 812, Pol 1901, Pol 1558 or Pol 963. The amino acid sequences of Pol 812, Pol 1901, Pol 1558 or Pol 963 DNA polymerases are described, for example, in U.S. Patent Publication Nos. 2020/0131484 A1 and 2020/0181587 A1, both of which are incorporated by reference herein. In some embodiments, the incorporation mix further comprises one or more buffering agents. The buffering agents may comprise a primary amine, a secondary amine, a tertiary amine, a natural amino acid, or a non-natural amino acid, or combinations thereof. In further embodiments, the buffering agents comprise ethanolamine or glycine, or a combination thereof. In one embodiment, the buffer agent comprises or is glycine. In further embodiments, the palladium scavenger comprises one or more allyl moieties does not require a separate washing step prior to the next incorporation cycle. In further embodiments, the palladium scavenger in the incorporation mix is a Pd(0) scavenger described herein. In some embodiments, the Pd(0) scavenger is premixed with the DNA polymerase and/or the one or more of four types of nucleotides (e.g., dATP, dCTP, dGTP, and dTTP or dUTP). In other embodiments, the Pd(0) scavenger is stored separately from the DNA polymerase and/or the one or more of four types of nucleotides and is mixed with these components shortly before sequencing run starts.

[0106] In some embodiments of any of the methods described herein, the concentration of the Pd(0) scavenger comprising one or more allyl moieties in the aqueous solution containing the DNA polymerase and nucleotides (e.g., A, G, C, and T or U; or dATP, dCTP, dGTP, and dTTP or dUTP) is from about 0.1 mM to about 100 mM, from 0.2 mM to about 75 mM, from about 0.5 mM to about 50 mM, from about 1 mM to about 20 mM, or from about 2 mM to about 10 mM. In further embodiments, the concentration of the palladium scavenger (e.g., the Pd(0) scavenger) is about 0.1 mM, 0.2 mM, 0.3 mM, 0.4 mM, 0.5 mM, 0.6 mM, 0.7 mM, 0.8 mM, 0.9 mM, 1 mM, 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM, 4 mM, 4.5 mM, 5 mM, 5.5 mM, 6 mM, 6.5 mM, 7 mM, 7.5 mM, 8 mM, 8.5 mM, 9 mM, 9.5 mM, 10 mM, 12.5 mM, 15 mM, 17.5 mM or 20 mM, or a range defined by any two of the preceding values. In further embodiments, the concentration of such palladium scavenger is the concentration in the incorporation mix. In further embodiments, the pH of the incorporation mix is about 9.

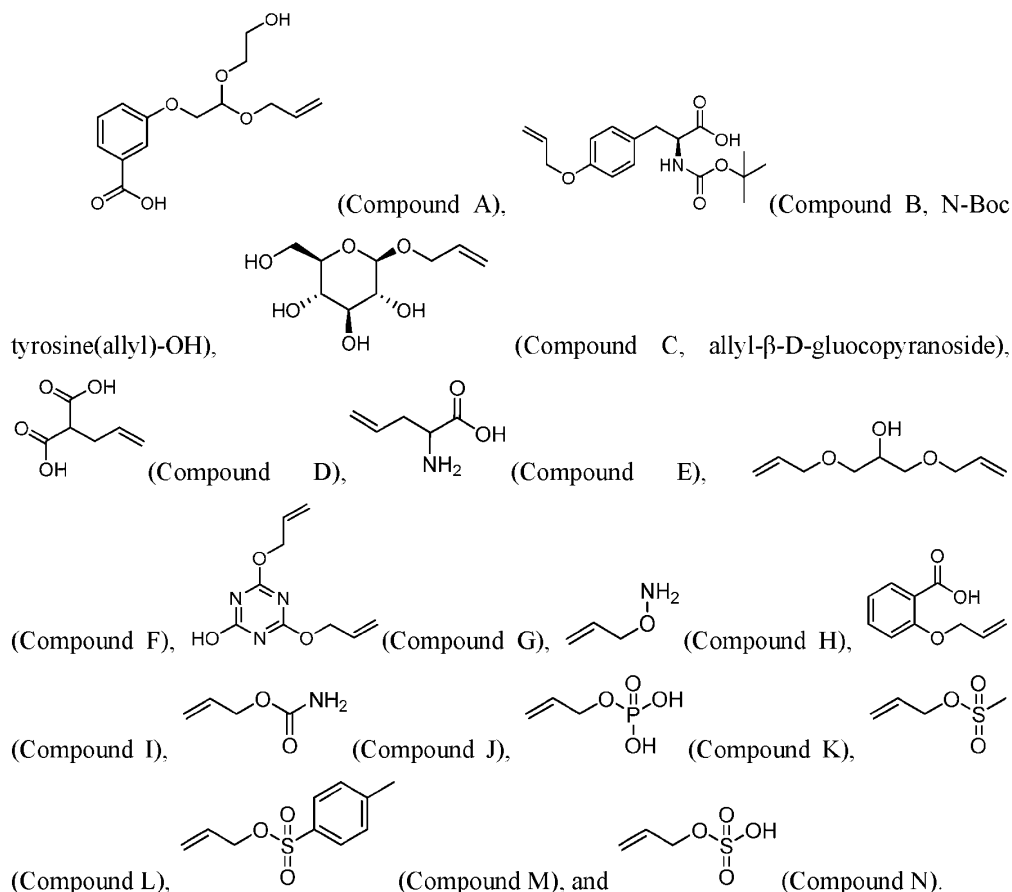
[0107] In some other embodiments of any of the methods described herein, the Pd(0) scavenger comprises one or more allyl moieties is in a solution when performing one or more fluorescent measurements. In such embodiment, such palladium scavenger is compatible with the sequencing reagents of the scanning solution (also known as the scan mix). In further embodiments, the one or more palladium scavengers does not require a separate washing step

prior to the next incorporation cycle. In further embodiments, the palladium scavenger in the scan solution is a Pd(0) scavenger described herein.

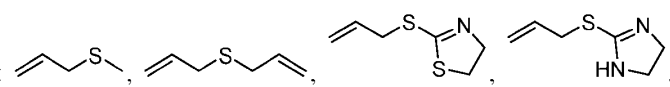
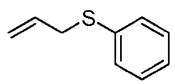
[0108] In other embodiments of the methods described herein, the Pd(0) scavenger comprises one or more allyl moieties in the post cleavage wash solution. In further embodiments, the palladium scavenger in the post cleavage wash solution is a Pd(0) scavenger described herein. In some such embodiment, the post cleavage wash solution does not comprise lipoic acid or 3,3'-dithiodipropionic acid (DDPA).

[0109] In still other embodiments of the method described herein, the Pd(0) scavenger comprises one or more allyl moieties may be present both in the incorporation mix and the post cleavage wash solution, or present in both the incorporation mix and the scan mix. In some such embodiment, the post cleavage wash solution does not comprise lipoic acid or DDPA.

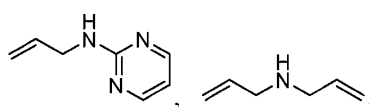
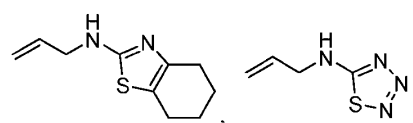
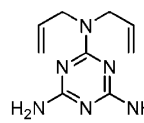
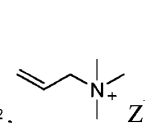
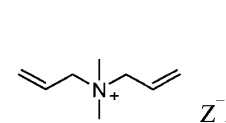
[0110] Non-limiting examples of the Pd(0) scavenger comprising one or more -O-allyl or allyl moieties include the following:



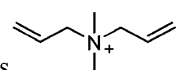
[0111] Non-limiting examples of the Pd(0) scavenger comprising one or more –S-allyl

moieties include the following:  and .

[0112] Non-limiting examples of the Pd(0) scavenger comprising one or more –NR-

allyl or –N⁺RR'-allyl moieties include the following: , , ,  or ,

where Z[–] is an anion (e.g., a halide anion such as F[–] or Cl[–]). In one embodiment, the palladium

scavenger is  Cl[–] (Compound O, diallyldimethylammonium chloride, also known as DADMAC).

Pd(II) Scavengers

[0113] In some embodiments of the methods described herein, the method may further use additional palladium scavenger(s), such as Pd(II) scavenger(s). In some such embodiments, the use of additional Pd(II) scavenger(s) may improve the phasing value of the sequencing metrics. For example, the Pd(II) scavenger(s) may comprise an isocyanoacetate (ICNA) salt, ethyl isocyanoacetate, methyl isocyanoacetate, cysteine (e.g., L-cysteine) or a salt thereof (e.g., *N*-acetyl-L-cysteine), potassium ethylxanthogenate, potassium isopropyl xanthate, glutathione, ethylenediaminetetraacetic acid (EDTA), iminodiacetic acid, nitrilotriacetic acid, trimercapto-S-triazine, dimethyldithiocarbamate, dithiothreitol, mercaptoethanol, allyl alcohol, propargyl alcohol, thiol, thiosulfate salt (e.g., sodium thiosulfate or potassium thiosulfate), tertiary amine and/or tertiary phosphine, or combinations thereof. In one embodiment, the method also includes the use of L-cysteine or a salt thereof. In another embodiment, the method also includes the use of a thiosulfate salt such as sodium thiosulfate (Na₂S₂O₃). In some such embodiments, the Pd(II) scavenger (e.g., L-cysteine or sodium thiosulfate) is in the aqueous solution containing the DNA polymerase and the nucleotides (i.e., incorporation mix). In other embodiments, the Pd(II) scavenger (e.g., L-cysteine or sodium thiosulfate) is in the post cleavage wash solution. In other embodiments, the Pd(II) scavenger (e.g., L-cysteine or sodium thiosulfate) may be both present in the incorporation mixture. In other embodiments, the Pd(II) scavenger (e.g., L-cysteine or

sodium thiosulfate) may be present in the scan mixture (i.e., the solution in which one or more fluorescent measurements of the incorporated nucleotide are performed). In other embodiments, the Pd(II) scavenger may be present in one or more of incorporation mixture (i.e., the aqueous solution of step (b)), the scan mixture, or the post-cleavage wash solution (i.e., the aqueous solution of step (e)). In further embodiments, the concentration of the Pd(II) scavenger such as L-cysteine or sodium thiosulfate in the incorporation mixture or the post cleavage wash solution is from about 0.1 mM to about 100 mM, from 0.2 mM to about 75 mM, from about 0.5 mM to about 50 mM, from about 1 mM to about 20 mM, or from about 2 mM to about 10 mM. In further embodiments, the concentration of the Pd(II) scavenger such as L-cysteine or sodium thiosulfate is about 0.1 mM, 0.5 mM, 1 mM, 2 mM, 3 mM, 4 mM, 5 mM, 6 mM, 6.5 mM, 7 mM, 8 mM, 9 mM, 10 mM, 12.5 mM, 15 mM, 17.5 mM or 20 mM, or a range defined by any two of the preceding values. In further embodiments, the Pd(II) scavenger is in the post cleavage wash solution, and the concentration of the Pd(II) scavenger in the post cleavage wash solution is about 10 mM. In further embodiment, the post cleavage wash solution does not contain lipoic acid or DDPA. In other embodiments, the method does not include a post-cleavage wash step.

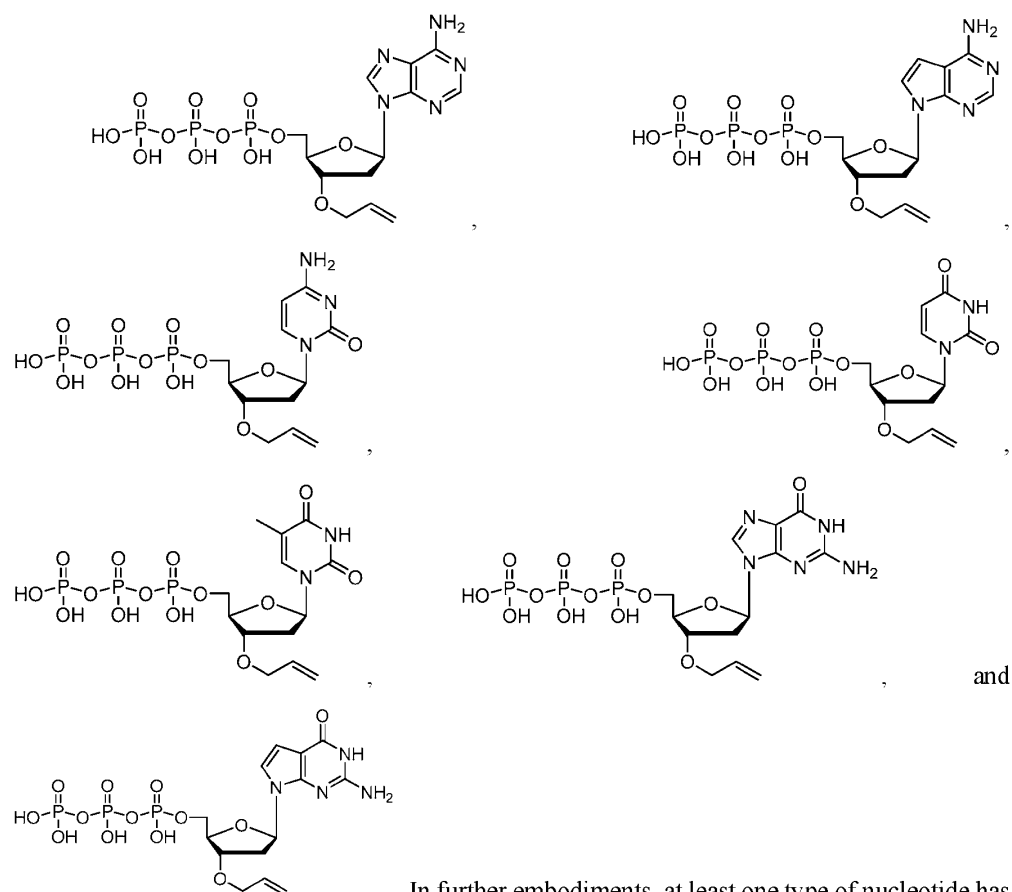
[0114] In some embodiments of the methods described herein, the target polynucleotide is immobilized to a surface of a substrate or a solid support. In some further embodiments, the surface comprises a plurality of immobilized target polynucleotides, for example, an array of different immobilized target polynucleotides. In some such embodiments, the substrate or solid support comprises glass, modified or functionalized glass, plastics, polysaccharides, nylon, nitrocellulose, resins, silica, silicon, modified silicon, carbon, metals, inorganic glasses, or optical fiber bundles, or combinations thereof. In some further embodiments, the substrate is a flowcell, a nanoparticle, or a bead (such as spherical silica beads, inorganic nanoparticles, magnetic nanoparticles, cadmium-based dots, and cadmium free dots, or a bead disclosed in U.S. Publication No. 2021/0187470 A1, which is incorporated by reference). In one embodiment, the substrate or solid support is a flowcell comprising patterned nanowells separated by interstitial regions, and wherein the immobilized target polynucleotides reside inside the patterned nanowells. In some embodiments, the substrate or solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise concatemers comprising said multiple copies of target polynucleotides. In some other embodiments, the substrate or solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise clusters of immobilized nucleic acid molecules comprising said multiple copies of target polynucleotides.

[0115] In some embodiments of any of the methods described herein, the method is performed on an automated sequencing instrument, and wherein the automated sequencing instrument comprises two light sources operating at different wavelengths (e.g., at about 450 nm

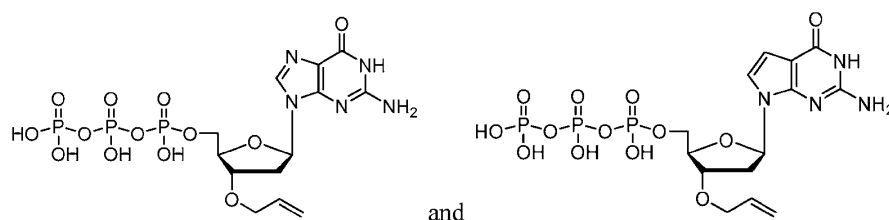
to about 460 nm, and about 520 nm to about 540 nm, in particular at about 460 nm and about 532 nm). In other embodiments, the automated sequencing instrument comprises a single light source operating at one wavelength.

Nucleotides with 3' Allyl Blocking Groups

[0116] Some embodiments of the present disclosure relate to a nucleotide molecule comprising a nucleobase, a ribose or deoxyribose moiety, and a 3' blocking group comprising an unsubstituted or substituted allyl moiety. In one embodiment, the 3' blocking group is an allyl group ($-\text{CH}_2\text{CH}=\text{CH}_2$), attached to the 3' oxygen atom of the deoxyribose moiety. In some embodiments, the nucleotide is unlabeled. In other embodiments, the nucleotide is a labeled nucleotide described herein. In some embodiments, wherein the bases for the A and G nucleotides are deazapurines (e.g., 7-deazapurine such as 7-deaza adenine or 7-deaza guanine). In some further embodiments, at least one type of nucleotide used in any of the sequencing method described herein has a structure selected from the group consisting of:

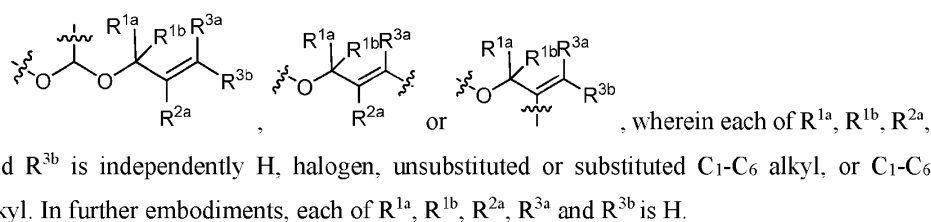


. In further embodiments, at least one type of nucleotide has a structure selected from the group consisting of:



Labeled Nucleotides

[0117] In some embodiments, the 3' blocked nucleotide also comprises a detectable label and such nucleotide is called a labeled nucleotide or a fully functionalized nucleotide (ffN). The label (e.g., a fluorescent dye) is conjugated via a cleavable linker by a variety of means including hydrophobic attraction, ionic attraction, and covalent attachment. In some aspect, the dyes are conjugated to the nucleotide by covalent attachment via the cleavable linker. One of ordinary skill in the art understands that label may be covalently bounded to the linker by reacting a functional group of the label (e.g., carboxyl) with a functional group of the linker (e.g., amino). In some such embodiments, the cleavable linker may comprise a moiety that is the same as the 3' blocking group. As such, the cleavable linker and the 3' blocking group may be cleaved or removed under the same reaction condition. In some such embodiments, the cleavable linker may comprise an allyl moiety, more particularly comprises a moiety of the structure:



[0118] In some embodiments, the detectable label may be covalently attached to oligonucleotides or nucleotides via the nucleotide base. For example, the labeled nucleotide or oligonucleotide may have the label attached to the C5 position of a pyrimidine base or the C7 position of a 7-deaza purine base through a cleavable linker moiety.

[0119] Nucleotides may be labeled at sites on the sugar or nucleobase. As known in the art, a “nucleotide” consists of a nitrogenous base, a sugar, and one or more phosphate groups. In RNA, the sugar is ribose and in DNA is a deoxyribose, i.e., a sugar lacking a hydroxy group that is present in ribose. The nitrogenous base is a derivative of purine (e.g., deazapurine, 7-deazapurine) or pyrimidine. The purines are adenine (A) and guanine (G), and the pyrimidines are cytosine (C) and thymine (T) or in the context of RNA, uracil (U). The C-1 atom of deoxyribose is bonded to N-1 of a pyrimidine or N-9 of a purine. A nucleotide is also a phosphate

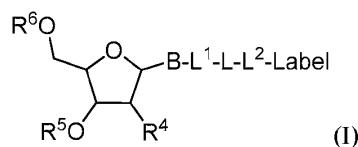
ester of a nucleoside, with esterification occurring on the hydroxy group attached to the C-3 or C-5 of the sugar. Nucleotides are usually mono, di- or triphosphates.

[0120] Although the base is usually referred to as a purine or pyrimidine, the skilled person will appreciate that derivatives and analogues are available which do not alter the capability of the nucleotide or nucleoside to undergo Watson-Crick base pairing. "Derivative" or "analogue" means a compound or molecule whose core structure is the same as, or closely resembles that of a parent compound but which has a chemical or physical modification, such as, for example, a different or additional side group, which allows the derivative nucleotide or nucleoside to be linked to another molecule. For example, the base may be a deazapurine. In particular embodiments, the derivatives should be capable of undergoing Watson-Crick pairing. "Derivative" and "analogue" also include, for example, a synthetic nucleotide or nucleoside derivative having modified base moieties and/or modified sugar moieties. Such derivatives and analogues are discussed in, for example, Scheit, *Nucleotide analogs* (John Wiley & Son, 1980) and Uhlman *et al.*, *Chemical Reviews* 90:543-584, 1990. Nucleotide analogues can also comprise modified phosphodiester linkages including phosphorothioate, phosphorodithioate, alkylphosphonate, phosphoranilidate, phosphoramidite linkages and the like.

[0121] In particular embodiments the labeled nucleotide may be enzymatically incorporable and enzymatically extendable. Accordingly, a linker moiety may be of sufficient length to connect the nucleotide to the compound such that the compound does not significantly interfere with the overall binding and recognition of the nucleotide by a nucleic acid replication enzyme. Thus, the linker can also comprise a spacer unit. The spacer distances, for example, the nucleotide base from a cleavage site or label.

[0122] The disclosure also encompasses polynucleotides incorporating a nucleotide described herein. Such polynucleotides may be DNA or RNA comprised respectively of deoxyribonucleotides or ribonucleotides joined in phosphodiester linkage. Polynucleotides may comprise naturally occurring nucleotides, non-naturally occurring (or modified) nucleotides other than the labeled nucleotides described herein or any combination thereof, in combination with at least one modified nucleotide (e.g., labeled with a dye compound) as set forth herein. Polynucleotides according to the disclosure may also include non-natural backbone linkages and/or non-nucleotide chemical modifications. Chimeric structures comprised of mixtures of ribonucleotides and deoxyribonucleotides comprising at least one labeled nucleotide are also contemplated.

[0123] In some embodiments, the labeled nucleotide described herein comprises or has the structure of Formula (I):

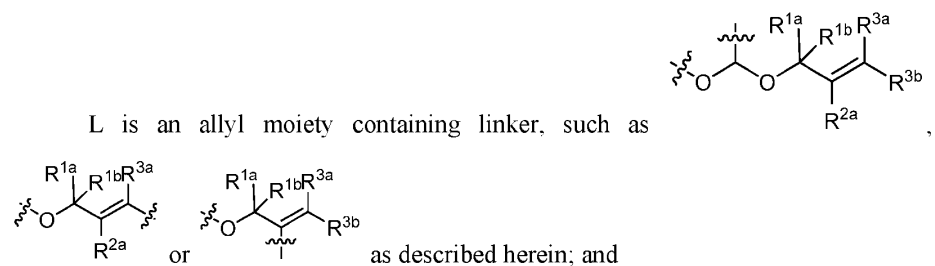


wherein B is the nucleobase;

R⁴ is H or OH;

R⁵ is an unsubstituted or substituted allyl group, such as ; wherein each of R^a, R^b, R^c, R^d and R^e is independently H, halogen, unsubstituted or substituted C₁-C₆ alkyl, or C₁-C₆ haloalkyl;

R⁶ is H, monophosphate, diphosphate, triphosphate, thiophosphate, a phosphate ester analog, a reactive phosphorous containing group, or a hydroxy protecting group;



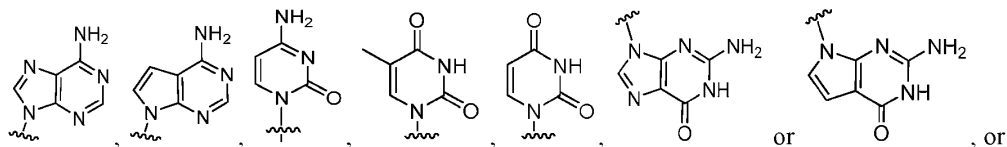
each of L¹ and L² is independently an optionally present linker moiety.

[0124] In some embodiments of the nucleotide described herein, each of R^a, R^b, R^c, R^d and R^e is H. In other embodiments, each of R^a and R^b is H and at least one of R^c, R^d and R^e is independently halogen (e.g., fluoro, chloro) or unsubstituted C₁-C₆ alkyl (e.g., methyl, ethyl, isopropyl, isobutyl, or t-butyl). For example, R^c is unsubstituted C₁-C₆ alkyl and each of R^d and R^e is H. In another example, R^c is H and one or both of R^d and R^e is halogen or unsubstituted C₁-C₆ alkyl. In some embodiments of the nucleotide described herein, each of R^{1a}, R^{1b}, R^{2a}, R^{3a} and R^{3b} is H. In other embodiments, at least one of R^{1a}, R^{1b}, R^{2a}, R^{3a} and R^{3b} is halogen (e.g., fluoro, chloro) or unsubstituted C₁-C₆ alkyl (e.g., methyl, ethyl, isopropyl, isobutyl, or t-butyl). In some such instances, each of R^{1a} and R^{1b} is H and at least one of R^{2a}, R^{3a} and R^{3b} is unsubstituted C₁-C₆ alkyl or halogen (for example, R^{2a} is unsubstituted C₁-C₆ alkyl and each of R^{3a} and R^{3b} is H; or R^{2a} is H and one or both of R^{3a} and R^{3b} is halogen or unsubstituted C₁-C₆ alkyl). In one

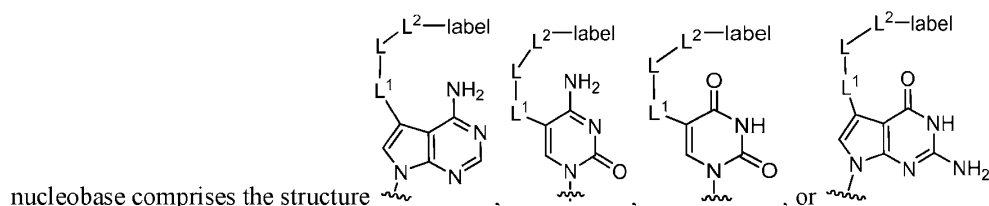
embodiment, the cleavable linker or L comprises (“AOL” linker moiety).

[0125] In some embodiments of the nucleotide of Formula (I) as described herein, the nucleobase (“B” in Formula (I)) is purine (adenine or guanine), a deaza purine, or a pyrimidine

(e.g., cytosine, thymine or uracil). In some further embodiments, the deaza purine is 7-deaza purine (e.g., 7-deaza adenine or 7-deaza guanine). Non-limiting examples of B comprises

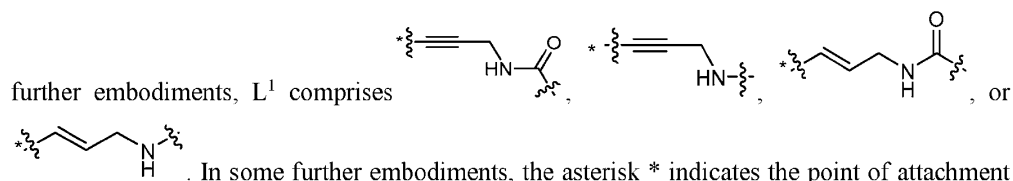


optionally substituted derivatives and analogs thereof. In some further embodiments, the labeled



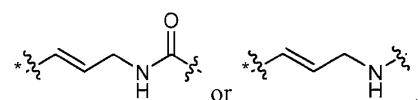
nucleobase comprises the structure

[0126] In some embodiments of the nucleotide of Formula (I) as described herein, L¹ is present and L¹ comprises a moiety selected from the group consisting of a propargylamine, a propargylamide, an allylamine, an allylamide, and optionally substituted variants thereof. In some



further embodiments, L¹ comprises

In some further embodiments, the asterisk * indicates the point of attachment of L¹ to the nucleobase (e.g., C5 position of a pyrimidine base or the C7 position of a 7-deaza purine base). In further embodiments, when the nucleobase is T, L¹ comprises or is

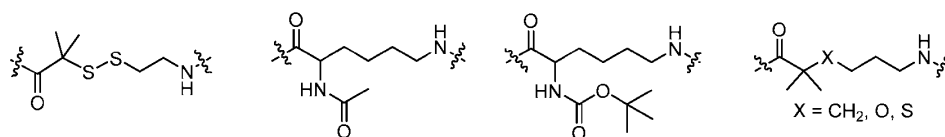


[0127] In some further embodiments of the nucleotide of Formula (I) as described

herein, L² is present and L² comprises

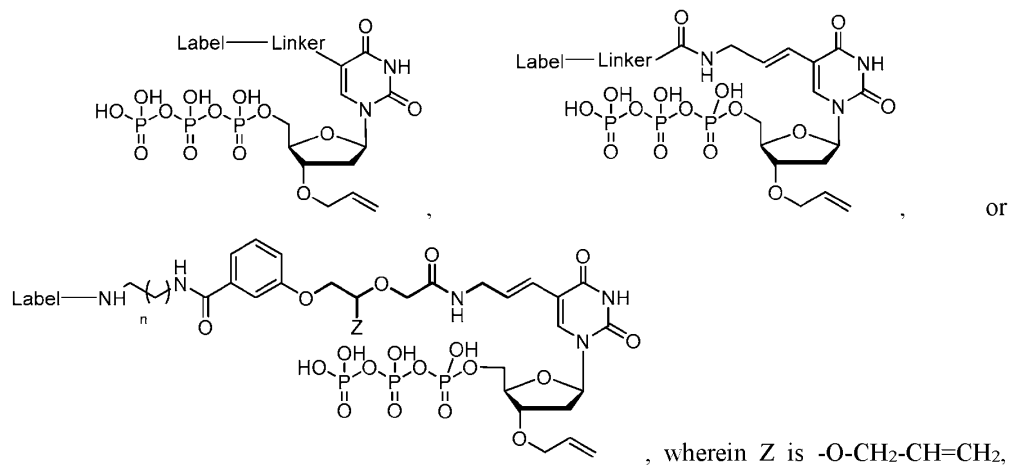
wherein n is an integer of 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 and the phenyl moiety is optionally substituted. In some such embodiments, n is 5 and the phenyl moiety of L² is unsubstituted.

[0128] Additional non-limiting examples of a linker moiety may be incorporated into L¹ or L² include:

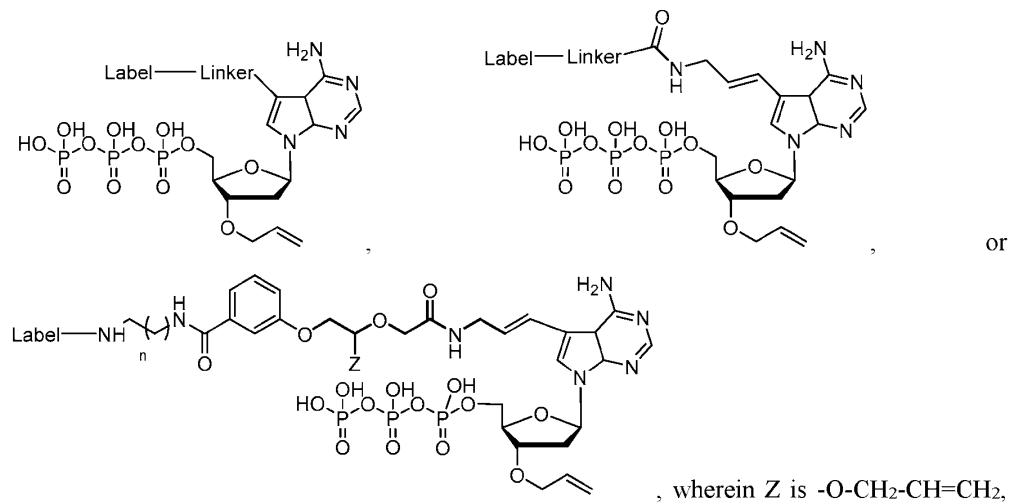


Additional linker moieties are disclosed in WO 2004/018493 and U.S. Publication Nos. 2016/0040225 and 2021/0403500, which are herein incorporated by references.

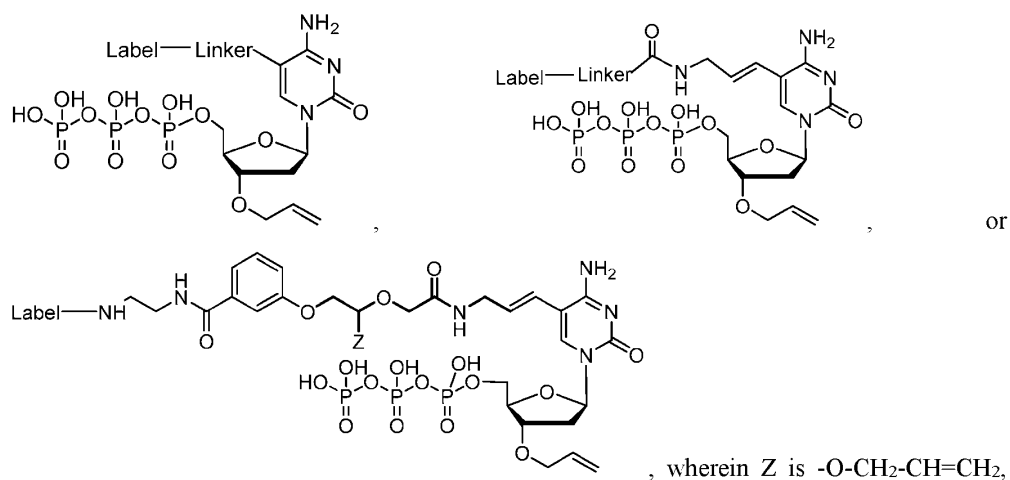
[0129] Some further embodiments of T nucleotide described herein include:



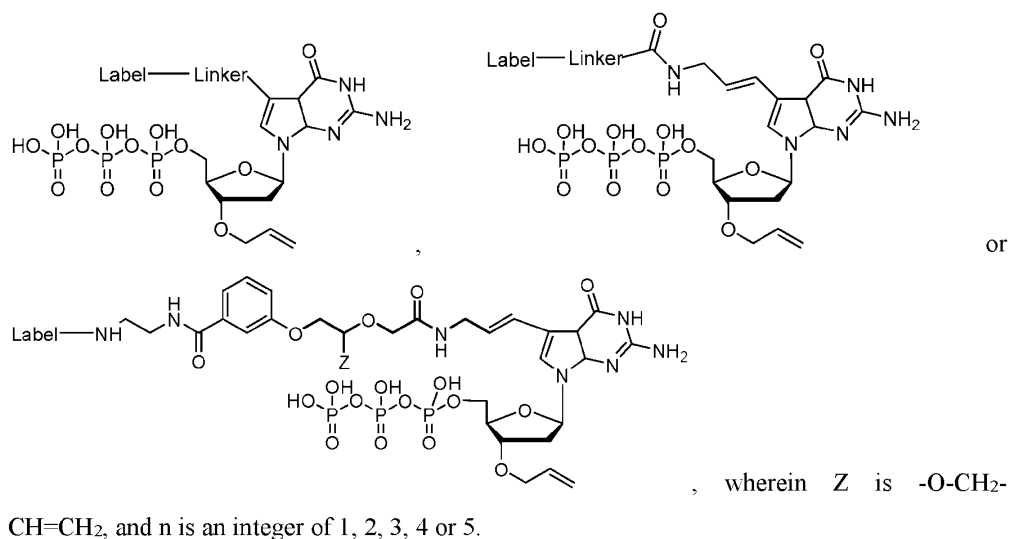
[0130] Some further embodiments of the A nucleotide described herein include:



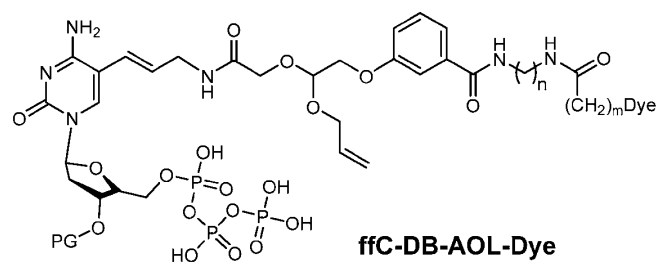
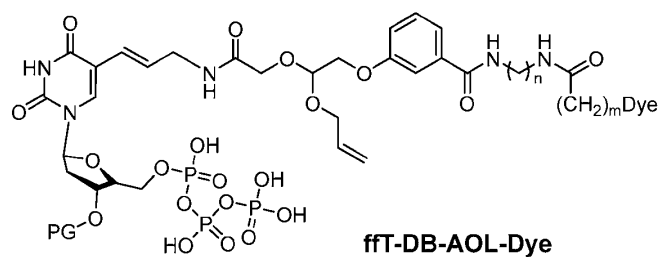
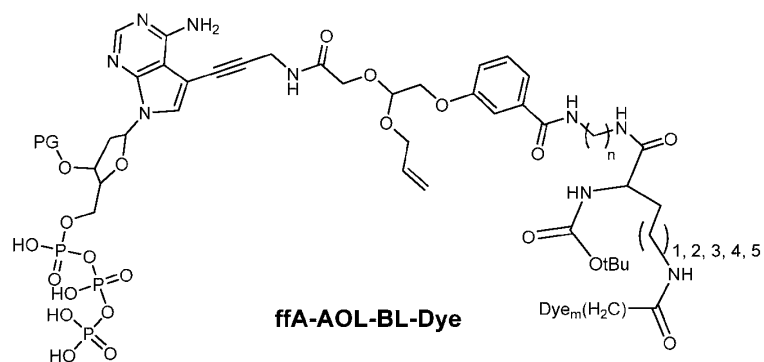
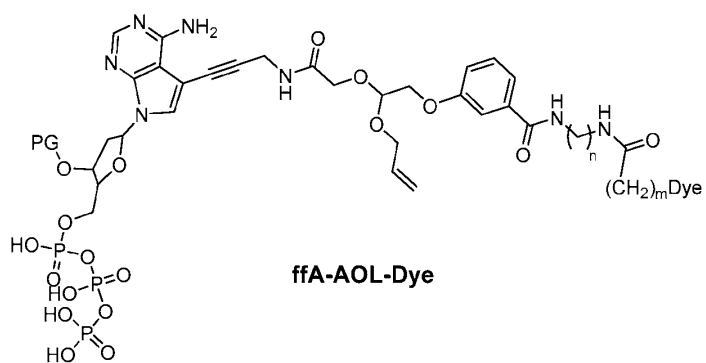
[0131] Some further embodiments of the C nucleotide described herein include:

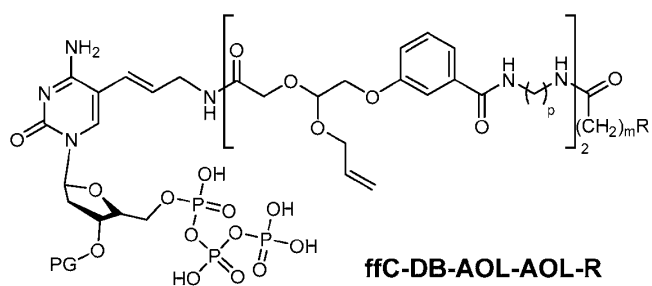
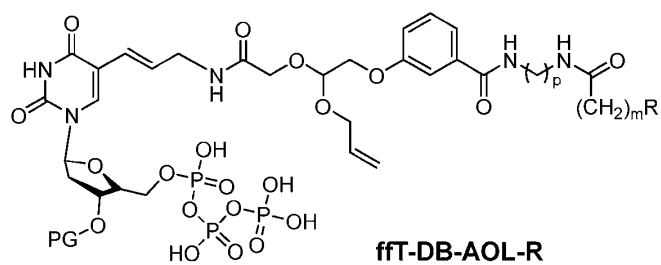
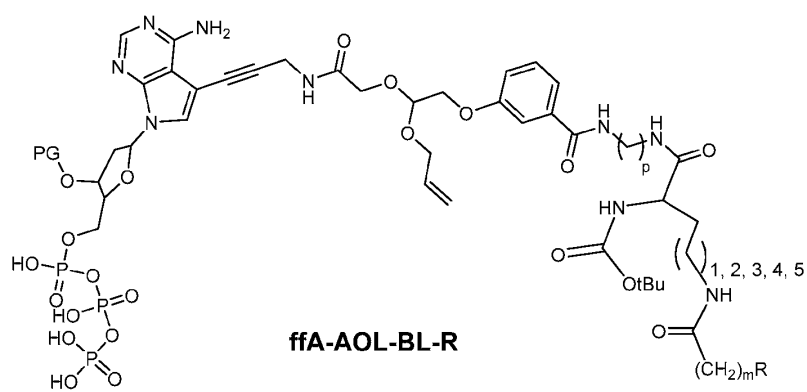
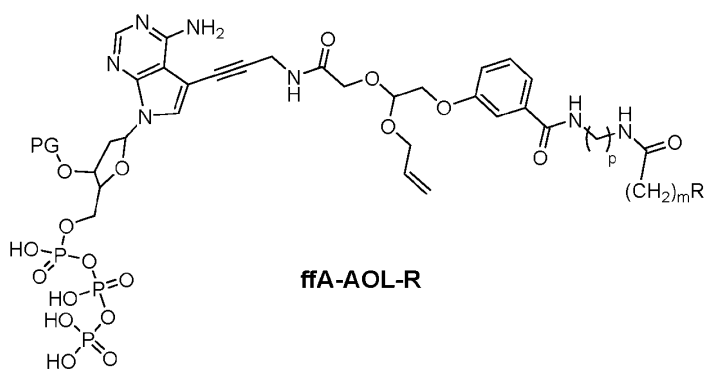


[0132] Some further embodiments of the G nucleotide described herein include:

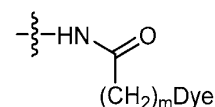


[0133] Non-limiting exemplary labeled nucleotides as described herein include:



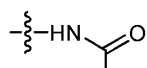


wherein PG stands for the 3' allyl blocking group described herein; n is an integer of 1, 2,



3, 4, 5, 6, 7, 8, 9, or 10; and m is 0, 1, 2, 3, 4, or 5. In one embodiment, n is 5.

refers to the connection point of the Dye with the cleavable linker as a result of a reaction between an amino group of the linker moiety and the carboxyl group of the Dye. In one embodiment, p is



5. $(\text{CH}_2)_m \text{R}$ refers to the connection point of the first/second/third functional moiety with the cleavable linker as a result of a reaction between an amino group of the linker moiety and the carboxyl group of the first/second/third functional moiety. In further embodiments, the nucleotide may be attached to the first/second/third functional moiety via more than one of the same cleavable linkers (such as AOL-AOL). In addition, the linker may further include additional PEG spacers as described herein, for example, between R and $(\text{CH}_2)_m$.

[0134] Various fluorescent dyes may be used in the present disclosure as detectable labels, in particularly those dyes that may be excitation by a blue light (e.g., about 450 nm to about 460 nm) or a green light (e.g., about 520 nm to about 540 nm). These dyes may also be referred to as “blue dyes” and “green dyes” respectively. Examples of various type of blue dyes, including but not limited to coumarin dyes, chromenoquinoline dyes, and bisboron containing heterocycles are disclosed in U.S. Publication Nos. 2018/0094140, 2018/0201981, 2020/0277529, 2020/0277670, 2021/0188832, 2022/0033900, 2022/0195517, 2022/0380389, and 2023/0313292, and U.S. Ser. No. 18/342064, each of which is incorporated by reference in its entirety. Examples of green dyes including cyanine or polymethine dyes disclosed in International Publication Nos. WO2013/041117, WO2014/135221, WO 2016/189287, WO2017/051201 and WO2018/060482A1, each of which is incorporated by reference in its entirety.

[0135] In any embodiments of nucleotide as described herein, the nucleotide comprises a 2' deoxyribose moiety (i.e., R^4 is Formula (I) is H). In some further respect, the 2' deoxyribose contains one, two or three phosphate groups at the 5' position of the sugar ring. In some further aspect, the nucleotides described herein are nucleotide triphosphate (i.e., $-\text{OR}^6$ in Formula (I) forms triphosphate).

[0136] Additional embodiments of the present disclosure relate to an oligonucleotide or a polynucleotide comprising a nucleoside or nucleotide described herein. In some such embodiments, the oligonucleotide or polynucleotide is hybridized to a template or target polynucleotide. In some such embodiments, the template polynucleotide is immobilized on a solid support.

[0137] Additional embodiments of the present disclosure relate to a solid support comprises an array of a plurality of immobilized template or target polynucleotides and at least a portion of such immobilized template or target polynucleotides is hybridized to an oligonucleotide or a polynucleotide comprising a nucleoside or nucleotide described herein.

Cleavage Condition of the Cleavable Linker

[0138] In any embodiments of the nucleotides or nucleosides described herein, the 3' blocking group and the cleavable linker (and the attached label) may be removable under the same or substantially same chemical reaction conditions, for example, the 3' blocking group and the detectable label may be removed in a single chemical reaction. In other embodiments, the 3' blocking group and the detectable labeled are removed in two separate steps.

[0139] The cleavable linker described herein may be removed or cleaved under various chemical conditions. Non-limiting cleaving condition includes a palladium catalyst, such as a Pd(II) complex (e.g., Pd(OAc)₂, allylPd(II) chloride dimer [(Allyl)PdCl]₂ or Na₂PdCl₄) in the presence of a water soluble phosphine ligand, for example tris(hydroxylpropyl)phosphine (THPP or THP), tris(hydroxymethyl)phosphine, and/or tris(2-carboxyethyl)phosphine (TCEP), with or without the presence of a reducing agent. In some embodiments, the 3' blocking group may be cleaved under the same or substantially the same cleavage condition as that for the cleavable linker.

Compatibility with Linearization

[0140] In order to maximize the throughput of nucleic acid sequencing reactions it is advantageous to be able to sequence multiple template molecules in parallel. Parallel processing of multiple templates can be achieved with the use of nucleic acid array technology. These arrays typically consist of a high-density matrix of polynucleotides immobilized onto a solid support material.

[0141] WO 98/44151 and WO 00/18957 both describe methods of nucleic acid amplification which allow amplification products to be immobilized on a solid support in order to form arrays comprised of clusters or "colonies" formed from a plurality of identical immobilized polynucleotide strands and a plurality of identical immobilized complementary strands. Arrays of this type are referred to herein as "clustered arrays." The nucleic acid molecules present in DNA colonies on the clustered arrays prepared according to these methods can provide templates for sequencing reactions, for example as described in WO 98/44152. The products of solid-phase amplification reactions such as those described in WO 98/44151 and WO 00/18957 are so-called "bridged" structures formed by annealing of pairs of immobilized polynucleotide strands and

immobilized complementary strands, both strands being attached to the solid support at the 5' end. In order to provide more suitable templates for nucleic acid sequencing, it is preferred to remove substantially all or at least a portion of one of the immobilized strands in the "bridged" structure in order to generate a template which is at least partially single-stranded. The portion of the template which is single-stranded will thus be available for hybridization to a sequencing primer. The process of removing all or a portion of one immobilized strand in a "bridged" double-stranded nucleic acid structure is referred to as "linearization." There are various ways for linearization, including but not limited to enzymatic cleavage, photo-chemical cleavage, or chemical cleavage. Non-limiting examples of linearization methods are disclosed in PCT Publication No. WO 2007/010251, U.S. Patent Publication No. 2009/0088327, U.S. Patent Publication No. 2009/0118128, and U.S. Publication No. 2019/0352327, which are incorporated by reference in their entireties.

[0142] In some embodiments, the condition for the removal of the 3' blocking group and/or the cleavable linker is also compatible with the linearization processes, for example, a chemical linearization process which comprises the use of a Pd complex and a phosphine. In some embodiments, the Pd complex is a Pd(II) complex (e.g., Pd(OAc)₂, [(Allyl)PdCl]₂ or Na₂PdCl₄), which generates Pd(0) in situ in the presence of a water soluble phosphine described herein, without or without the presence of a reducing agent.

Embodiments and Alternatives of Sequencing-By-Synthesis

[0143] Alternatively, the sequencing methods described herein may also be carried out using unlabeled nucleotides and affinity reagents containing a fluorescent dye described herein. For example, one, two, three or each of the four different types of nucleotides (e.g., dATP, dCTP, dGTP and dTTP or dUTP) in the incorporation mixture of step (a) may be unlabeled. Each of the four types of nucleotides (e.g., dNTPs) has a 3' blocking group to ensure that only a single base can be added by a polymerase to the 3' end of the primer polynucleotide. After incorporation of an unlabeled nucleotide in step (b), the remaining unincorporated nucleotides are washed away. An affinity reagent is then introduced that specifically recognizes and binds to the incorporated dNTP to provide a labeled extension product comprising the incorporated dNTP. Uses of unlabeled nucleotides and affinity reagents in sequencing by synthesis have been disclosed in WO 2018/129214 and WO 2020/097607. A modified sequencing method of the present disclosure using unlabeled nucleotides may include the following steps:

(a') contacting a solid support with sequencing primers under hybridization conditions, wherein the solid support comprises a plurality of target polynucleotides

immobilized thereon; and the sequencing primers are complementary to at least a portion of the target polynucleotides;

(b') contacting the solid support with an aqueous solution comprising DNA polymerase and one or more of four different types of unlabeled nucleotides (A, G, C and T or U) under conditions suitable for DNA polymerase-mediated primer extension, wherein each of the nucleotides comprises a 3' blocking group comprising an unsubstituted or substituted allyl moiety as described herein (e.g., each nucleotide comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom);

(c') contacting the extended copy polynucleotides with a set of affinity reagents under conditions wherein one affinity reagent binds specifically to the incorporated unlabeled nucleotides to provide labeled extended copy polynucleotides;

(d') imaging the solid support to determine the identity of the incorporated nucleotides (e.g., by performing one or more fluorescent measurements of the extended copy polynucleotides); and

(e') contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove 3' allyl blocking groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support; and

repeating steps (b')-(e') to determine target polynucleotide sequences.

[0144] In some embodiments of the modified sequencing method described herein, the method further comprises removing the affinity reagents from the incorporated nucleotides. In still further embodiments, the 3' blocking group and the affinity reagent are removed in the same reaction. In some embodiments, the method further comprises a step (f') washing the solid support with a third aqueous wash solution. In further embodiments, steps (b') through (f') are repeated at least 50, 100, 150, 200, 250 or 300 cycles to determine the target polynucleotide sequences. In some embodiments, the set of affinity reagents may comprise a first affinity reagent that binds specifically to the first type of nucleotide, a second affinity reagent that binds specifically to the second type of nucleotide, and a third affinity reagent that binds specifically to the third type of nucleotide. In some further embodiments, each of the first, second and the third affinity reagents comprises a detectable label that is spectrally distinguishable. In some embodiments, the affinity reagents may include protein tags, antibodies (including but not limited to binding fragments of antibodies, single chain antibodies, bispecific antibodies, and the like), aptamers, knottins, affimers, or any other known agent that binds an incorporated nucleotide with a suitable specificity and affinity. In one embodiment, at least one affinity reagent is an antibody or a protein tag. In another embodiment, at least one of the first type, the second type, and the third type of affinity

reagents is an antibody or a protein tag comprising one or more detectable labels (e.g., multiple copies of the same detectable label), wherein the detectable label is or comprises a bis-boron dye moiety described herein.

[0145] Some embodiments include pyrosequencing techniques. Pyrosequencing detects the release of inorganic pyrophosphate (PPi) as particular nucleotides are incorporated into the nascent strand (Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. and Nyren, P. (1996) "Real-time DNA sequencing using detection of pyrophosphate release." *Analytical Biochemistry* 242(1), 84-9; Ronaghi, M. (2001) "Pyrosequencing sheds light on DNA sequencing." *Genome Res.* 11(1), 3-11; Ronaghi, M., Uhlen, M. and Nyren, P. (1998) "A sequencing method based on real-time pyrophosphate." *Science* 281(5375), 363; U.S. Pat. Nos. 6,210,891; 6,258,568 and 6,274,320, the disclosures of which are incorporated herein by reference in their entireties). In pyrosequencing, released PPi can be detected by being immediately converted to adenosine triphosphate (ATP) by ATP sulfurase, and the level of ATP generated is detected via luciferase-produced photons. The nucleic acids to be sequenced can be attached to features in an array and the array can be imaged to capture the chemiluminescent signals that are produced due to incorporation of a nucleotides at the features of the array. An image can be obtained after the array is treated with a particular nucleotide type (e.g., A, T, C or G). Images obtained after addition of each nucleotide type will differ with regard to which features in the array are detected. These differences in the image reflect the different sequence content of the features on the array. However, the relative locations of each feature will remain unchanged in the images. The images can be stored, processed and analyzed using the methods set forth herein. For example, images obtained after treatment of the array with each different nucleotide type can be handled in the same way as exemplified herein for images obtained from different detection channels for reversible terminator-based sequencing methods.

[0146] In another exemplary type of SBS, cycle sequencing is accomplished by stepwise addition of reversible terminator nucleotides containing, for example, a cleavable or photobleachable dye label as described, for example, in WO 04/018497 and U.S. Pat. No. 7,057,026, the disclosures of which are incorporated herein by reference. This approach is being commercialized by Solexa (now Illumina, Inc.), and is also described in WO 91/06678 and WO 07/123,744, each of which is incorporated herein by reference. The availability of fluorescently-labeled terminators in which both the termination can be reversed, and the fluorescent label cleaved facilitates efficient cyclic reversible termination (CRT) sequencing. Polymerases can also be co-engineered to efficiently incorporate and extend from these modified nucleotides.

[0147] Preferably in reversible terminator-based sequencing embodiments, the labels do not substantially inhibit extension under SBS reaction conditions. However, the detection

labels can be removable, for example, by cleavage or degradation. Images can be captured following incorporation of labels into arrayed nucleic acid features. In particular embodiments, each cycle involves simultaneous delivery of four different nucleotide types to the array and each nucleotide type has a spectrally distinct label. Four images can then be obtained, each using a detection channel that is selective for one of the four different labels. Alternatively, different nucleotide types can be added sequentially, and an image of the array can be obtained between each addition step. In such embodiments each image will show nucleic acid features that have incorporated nucleotides of a particular type. Different features will be present or absent in the different images due to the different sequence content of each feature. However, the relative position of the features will remain unchanged in the images. Images obtained from such reversible terminator-SBS methods can be stored, processed and analyzed as set forth herein. Following the image capture step, labels can be removed, and reversible terminator moieties can be removed for subsequent cycles of nucleotide addition and detection. Removal of the labels after they have been detected in a particular cycle and prior to a subsequent cycle can provide the advantage of reducing background signal and crosstalk between cycles. Examples of useful labels and removal methods are set forth below.

[0148] Some embodiments can utilize detection of four different nucleotides using fewer than four different labels. For example, SBS can be performed utilizing methods and systems described in the incorporated materials of U.S. Pub. No. 2013/0079232. As a first example, a pair of nucleotide types can be detected at the same wavelength, but distinguished based on a difference in intensity for one member of the pair compared to the other, or based on a change to one member of the pair (e.g. via chemical modification, photochemical modification or physical modification) that causes apparent signal to appear or disappear compared to the signal detected for the other member of the pair. As a second example, three of four different nucleotide types can be detected under particular conditions while a fourth nucleotide type lacks a label that is detectable under those conditions, or is minimally detected under those conditions (e.g., minimal detection due to background fluorescence, etc.). Incorporation of the first three nucleotide types into a nucleic acid can be determined based on presence of their respective signals and incorporation of the fourth nucleotide type into the nucleic acid can be determined based on absence or minimal detection of any signal. As a third example, one nucleotide type can include label(s) that are detected in two different channels, whereas other nucleotide types are detected in no more than one of the channels. The aforementioned three exemplary configurations are not considered mutually exclusive and can be used in various combinations. An exemplary embodiment that combines all three examples, is a fluorescent-based SBS method that uses a first nucleotide type that is detected in a first channel (e.g. dATP having a label that is detected in the

first channel when excited by a first excitation wavelength), a second nucleotide type that is detected in a second channel (e.g. dCTP having a label that is detected in the second channel when excited by a second excitation wavelength), a third nucleotide type that is detected in both the first and the second channel (e.g. dTTP having at least one label that is detected in both channels when excited by the first and/or second excitation wavelength) and a fourth nucleotide type that lacks a label that is not, or minimally, detected in either channel (e.g. dGTP having no label).

[0149] Further, as described in the incorporated materials of U.S. Pub. No. 2013/0079232, sequencing data can be obtained using a single channel. In such so-called one-dye sequencing approaches, the first nucleotide type is labeled but the label is removed after the first image is generated, and the second nucleotide type is labeled only after a first image is generated. The third nucleotide type retains its label in both the first and second images, and the fourth nucleotide type remains unlabeled in both images.

[0150] Some embodiments can utilize sequencing by ligation techniques. Such techniques utilize DNA ligase to incorporate oligonucleotides and identify the incorporation of such oligonucleotides. The oligonucleotides typically have different labels that are correlated with the identity of a particular nucleotide in a sequence to which the oligonucleotides hybridize. As with other SBS methods, images can be obtained following treatment of an array of nucleic acid features with the labeled sequencing reagents. Each image will show nucleic acid features that have incorporated labels of a particular type. Different features will be present or absent in the different images due the different sequence content of each feature, but the relative position of the features will remain unchanged in the images. Images obtained from ligation-based sequencing methods can be stored, processed and analyzed as set forth herein. Exemplary SBS systems and methods which can be utilized with the methods and systems described herein are described in U.S. Pat. Nos. 6,969,488, 6,172,218, and 6,306,597, the disclosures of which are incorporated herein by reference in their entireties.

[0151] Some embodiments can utilize nanopore sequencing (Deamer, D. W. & Akeson, M. "Nanopores and nucleic acids: prospects for ultrarapid sequencing." *Trends Biotechnol.* 18, 147-151 (2000); Deamer, D. and D. Branton, "Characterization of nucleic acids by nanopore analysis," *Acc. Chem. Res.* 35:817-825 (2002); Li, J., M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko, "DNA molecules and configurations in a solid-state nanopore microscope," *Nat. Mater.* 2:611-615 (2003), the disclosures of which are incorporated herein by reference in their entireties). In such embodiments, the target nucleic acid passes through a nanopore. The nanopore can be a synthetic pore or biological membrane protein, such as α -hemolysin. As the target nucleic acid passes through the nanopore, each base-pair can be identified by measuring fluctuations in the electrical conductance of the pore. (U.S. Pat. No. 7,001,792; Soni,

G. V. & Meller, "A. Progress toward ultrafast DNA sequencing using solid-state nanopores," *Clin. Chem.* 53, 1996-2001 (2007); Healy, K. "Nanopore-based single-molecule DNA analysis," *Nanomed.* 2, 459-481 (2007); Cockroft, S. L., Chu, J., Amorin, M. & Ghadiri, M. R. "A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution," *J. Am. Chem. Soc.* 130, 818-820 (2008), the disclosures of which are incorporated herein by reference in their entireties). Data obtained from nanopore sequencing can be stored, processed and analyzed as set forth herein. In particular, the data can be treated as an image in accordance with the exemplary treatment of optical images and other images that is set forth herein.

[0152] Some other embodiments of sequencing method involve the use the 3' blocked nucleotide described herein in nanoball sequencing technique, such as those described in U.S. Patent No. 9,222,132, the disclosure of which is incorporated by reference. Through the process of rolling circle amplification (RCA), a large number of discrete DNA nanoballs may be generated. The nanoball mixture is then distributed onto a patterned slide surface containing features that allow a single nanoball to associate with each location. In DNA nanoball generation, DNA is fragmented and ligated to the first of four adapter sequences. The template is amplified, circularized and cleaved with a type II endonuclease. A second set of adapters is added, followed by amplification, circularization and cleavage. This process is repeated for the remaining two adapters. The final product is a circular template with four adapters, each separated by a template sequence. Library molecules undergo a rolling circle amplification step, generating a large mass of concatemers called DNA nanoballs, which are then deposited on a flow cell. Goodwin *et al.*, "Coming of age: ten years of next-generation sequencing technologies," *Nat Rev Genet.* 2016;17(6):333-51.

[0153] Some embodiments can utilize methods involving the real-time monitoring of DNA polymerase activity. Nucleotide incorporations can be detected through fluorescence resonance energy transfer (FRET) interactions between a fluorophore-bearing polymerase and γ -phosphate-labeled nucleotides as described, for example, in U.S. Pat. Nos. 7,329,492 and 7,211,414, both of which are incorporated herein by reference, or nucleotide incorporations can be detected with zero-mode waveguides as described, for example, in U.S. Pat. No. 7,315,019, which is incorporated herein by reference, and using fluorescent nucleotide analogs and engineered polymerases as described, for example, in U.S. Pat. No. 7,405,281 and U.S. Pub. No. 2008/0108082, both of which are incorporated herein by reference. The illumination can be restricted to a zeptoliter-scale volume around a surface-tethered polymerase such that incorporation of fluorescently labeled nucleotides can be observed with low background (Levene, M. J. *et al.* "Zero-mode waveguides for single-molecule analysis at high concentrations," *Science* 299, 682-686 (2003); Lundquist, P. M. *et al.* "Parallel confocal detection of single molecules in

real time,” *Opt. Lett.* 33, 1026-1028 (2008); Korlach, J. *et al.* “Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nano structures,” *Proc. Natl. Acad. Sci. USA* 105, 1176-1181 (2008), the disclosures of which are incorporated herein by reference in their entireties). Images obtained from such methods can be stored, processed and analyzed as set forth herein.

[0154] Some SBS embodiments include detection of a proton released upon incorporation of a nucleotide into an extension product. For example, sequencing based on detection of released protons can use an electrical detector and associated techniques that are commercially available from Ion Torrent (Guilford, CT, a Life Technologies subsidiary) or sequencing methods and systems described in U.S. Pub. Nos. 2009/0026082; 2009/0127589; 2010/0137143; and 2010/0282617, all of which are incorporated herein by reference. Methods set forth herein for amplifying target nucleic acids using kinetic exclusion can be readily applied to substrates used for detecting protons. More specifically, methods set forth herein can be used to produce clonal populations of amplicons that are used to detect protons.

[0155] The above SBS methods can be advantageously carried out in multiplex formats such that multiple different target nucleic acids are manipulated simultaneously. In particular embodiments, different target nucleic acids can be treated in a common reaction vessel or on a surface of a particular substrate. This allows convenient delivery of sequencing reagents, removal of unreacted reagents and detection of incorporation events in a multiplex manner. In embodiments using surface-bound target nucleic acids, the target nucleic acids can be in an array format. In an array format, the target nucleic acids can be typically bound to a surface in a spatially distinguishable manner. The target nucleic acids can be bound by direct covalent attachment, attachment to a bead or other particle or binding to a polymerase or other molecule that is attached to the surface. The array can include a single copy of a target nucleic acid at each site (also referred to as a feature) or multiple copies having the same sequence can be present at each site or feature. Multiple copies can be produced by amplification methods such as, bridge amplification or emulsion PCR as described in further detail below.

[0156] The methods set forth herein can use arrays having features at any of a variety of densities including, for example, at least about 10 features/cm², 100 features/cm², 500 features/cm², 1,000 features/cm², 5,000 features/cm², 10,000 features/cm², 50,000 features/cm², 100,000 features/cm², 1,000,000 features/cm², 5,000,000 features/cm², or higher.

[0157] An advantage of the methods set forth herein is that they provide for rapid and efficient detection of a plurality of target nucleic acid in parallel. Accordingly, the present disclosure provides integrated systems capable of preparing and detecting nucleic acids using techniques known in the art such as those exemplified above. Thus, an integrated system of the

present disclosure can include fluidic components capable of delivering amplification reagents and/or sequencing reagents to one or more immobilized DNA fragments, the system comprising components such as pumps, valves, reservoirs, fluidic lines and the like. A flow cell can be configured and/or used in an integrated system for detection of target nucleic acids. Exemplary flow cells are described, for example, in U.S. Pub. No. 2010/0111768 and U.S. Patent Appl. No. 13/273,666, each of which is incorporated herein by reference. As exemplified for flow cells, one or more of the fluidic components of an integrated system can be used for an amplification method and for a detection method. Taking a nucleic acid sequencing embodiment as an example, one or more of the fluidic components of an integrated system can be used for an amplification method set forth herein and for the delivery of sequencing reagents in a sequencing method such as those exemplified above. Alternatively, an integrated system can include separate fluidic systems to carry out amplification methods and to carry out detection methods. Examples of integrated sequencing systems that are capable of creating amplified nucleic acids and also determining the sequence of the nucleic acids include, without limitation, the MiSeq™ platform (Illumina, Inc., San Diego, CA) and devices described in U.S. Patent Appl. No. 13/273,666, which is incorporated herein by reference.

[0158] Arrays in which polynucleotides have been directly attached to silica-based supports are those for example disclosed in WO 00/06770 (incorporated herein by reference), wherein polynucleotides are immobilized on a glass support by reaction between a pendant epoxide group on the glass with an internal amino group on the polynucleotide. In addition, polynucleotides can be attached to a solid support by reaction of a sulfur-based nucleophile with the solid support, for example, as described in WO 2005/047301 (incorporated herein by reference). A still further example of solid-supported template polynucleotides is where the template polynucleotides are attached to hydrogel supported upon silica-based or other solid supports, for example, as described in WO 00/31148, WO 01/01143, WO 02/12566, WO 03/014392, U.S. Pat. No. 6,465,178 and WO 00/53812, each of which is incorporated herein by reference.

[0159] A particular surface to which template polynucleotides may be immobilized is a polyacrylamide hydrogel. Polyacrylamide hydrogels are described in the references cited above and in WO 2005/065814, which is incorporated herein by reference. Specific hydrogels that may be used include those described in WO 2005/065814 and U.S. Pub. No. 2014/0079923. In one embodiment, the hydrogel is PAZAM (poly(N-(5-azidoacetamidyl)pentyl) acrylamide-co-acrylamide)).

[0160] DNA template molecules can be attached to beads or microparticles, for example, as described in U.S. Pat. No. 6,172,218 (which is incorporated herein by reference).

Attachment to beads or microparticles can be useful for sequencing applications. Bead libraries can be prepared where each bead contains different DNA sequences. Exemplary libraries and methods for their creation are described in *Nature*, 437, 376-380 (2005); *Science*, 309, 5741, 1728-1732 (2005), each of which is incorporated herein by reference. Sequencing of arrays of such beads using nucleotides set forth herein is within the scope of the disclosure.

[0161] Templates that are to be sequenced may form part of an “array” on a solid support, in which case the array may take any convenient form. Thus, the method of the disclosure is applicable to all types of high-density arrays, including single-molecule arrays, clustered arrays, and bead arrays. Labeled nucleotides of the present disclosure may be used for sequencing templates on essentially any type of array, including but not limited to those formed by immobilization of nucleic acid molecules on a solid support.

[0162] However, labeled nucleotides of the disclosure are particularly advantageous in the context of sequencing of clustered arrays. In clustered arrays, distinct regions on the array (often referred to as sites, or features) comprise multiple polynucleotide template molecules. Generally, the multiple polynucleotide molecules are not individually resolvable by optical means and are instead detected as an ensemble. Depending on how the array is formed, each site on the array may comprise multiple copies of one individual polynucleotide molecule (e.g., the site is homogenous for a particular single- or double-stranded nucleic acid species) or even multiple copies of a small number of different polynucleotide molecules (e.g., multiple copies of two different nucleic acid species). Clustered arrays of nucleic acid molecules may be produced using techniques generally known in the art. By way of example, WO 98/44151 and WO 00/18957, each of which is incorporated herein, describe methods of amplification of nucleic acids wherein both the template and amplification products remain immobilized on a solid support in order to form arrays comprised of clusters or “colonies” of immobilized nucleic acid molecules. The nucleic acid molecules present on the clustered arrays prepared according to these methods are suitable templates for sequencing using the nucleotides labeled with dye compounds of the disclosure.

[0163] The labeled nucleotides of the present disclosure are also useful in sequencing of templates on single molecule arrays. The term “single molecule array” (“SMA”) as used herein refers to a population of polynucleotide molecules, distributed (or arrayed) over a solid support, wherein the spacing of any individual polynucleotide from all others of the population is such that it is possible to individually resolve the individual polynucleotide molecules. The target nucleic acid molecules immobilized onto the surface of the solid support can thus be capable of being resolved by optical means in some embodiments. This means that one or more distinct signals,

each representing one polynucleotide, will occur within the resolvable area of the particular imaging device used.

[0164] Single molecule detection may be achieved wherein the spacing between adjacent polynucleotide molecules on an array is at least 100 nm, more particularly at least 250 nm, still more particularly at least 300 nm, even more particularly at least 350 nm. Thus, each molecule is individually resolvable and detectable as a single molecule fluorescent point, and fluorescence from said single molecule fluorescent point also exhibits single step photobleaching.

[0165] The terms “individually resolved” and “individual resolution” are used herein to specify that, when visualized, it is possible to distinguish one molecule on the array from its neighboring molecules. Separation between individual molecules on the array will be determined, in part, by the particular technique used to resolve the individual molecules. The general features of single molecule arrays will be understood by reference to published applications WO 00/06770 and WO 01/57248, each of which is incorporated herein by reference. Although one use of the nucleotides of the disclosure is in sequencing-by-synthesis reactions, the utility of the nucleotides is not limited to such methods. In fact, the nucleotides may be used advantageously in any sequencing methodology which requires detection of fluorescent labels attached to nucleotides incorporated into a polynucleotide.

[0166] In particular, the labeled nucleotides of the disclosure may be used in automated fluorescent sequencing protocols, particularly fluorescent dye-terminator cycle sequencing based on the chain termination sequencing method of Sanger and co-workers. Such methods generally use enzymes and cycle sequencing to incorporate fluorescently labeled dideoxynucleotides in a primer extension sequencing reaction. So-called Sanger sequencing methods, and related protocols (Sanger-type), utilize randomized chain termination with labeled dideoxynucleotides.

[0167] Thus, the present disclosure also encompasses labeled nucleotides which are dideoxynucleotides lacking hydroxy groups at both of the 3' and 2' positions, such dideoxynucleotides being suitable for use in Sanger type sequencing methods and the like.

[0168] Labeled nucleotides of the present disclosure incorporating 3' blocking groups, it will be recognized, may also be of utility in Sanger methods and related protocols since the same effect achieved by using dideoxy nucleotides may be achieved by using nucleotides having 3'-OH blocking groups: both prevent incorporation of subsequent nucleotides. Where nucleotides according to the present disclosure, and having a 3' blocking group are to be used in Sanger-type sequencing methods it will be appreciated that the dye compounds or detectable labels attached to the nucleotides need not be connected via cleavable linkers, since in each instance where a labeled

nucleotide of the disclosure is incorporated; no nucleotides need to be subsequently incorporated and thus the label need not be removed from the nucleotide.

Kits

[0169] Some additional embodiments of the present disclosure relate to a sequencing kit comprising:

(a) an incorporation mixture comprising DNA polymerase and nucleotides A, G, C, and T or U, wherein:

(i) the nucleotides comprise a 2' deoxyribose moiety with a 3' unsubstituted or substituted allyl group as described herein (e.g., ) attached to the 3' carbon atom; and

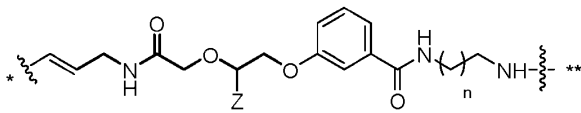
(ii) the DNA polymerase is an altered archaeal DNA polymerase;

(b) an aqueous deblocking solution comprising a palladium catalyst, tris(hydroxyalkyl)phosphine, and one or more buffer reagents that is suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers; and

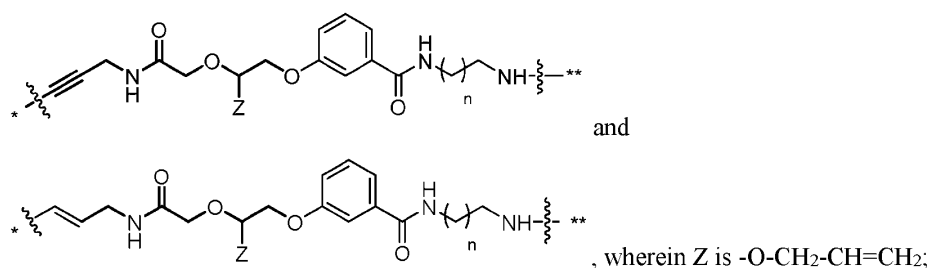
(c) an aqueous wash solution comprising a Pd(II) scavenger;

wherein said kit is configured for performing at least about 50, 100, 150, 200, 250 or 300 cycles of sequencing-by-synthesis.

[0170] In some further embodiment, the kit may comprise four types of labeled nucleotides of fully functionalized nucleotides described herein (A, C, T and G), where each type of nucleotide comprises the 3' allyl blocking group and the AOL linker moiety described herein. In further embodiments, G is unlabeled and does not comprise any cleavable linker. In still further embodiments, one or more the remaining three nucleotides (i.e., A, C and T) comprises an allylamine or allylamide linker moiety described herein. In one embodiment, the kit comprises unlabeled ffG, labeled ffA(s), labeled ffC, and labeled ffT (wherein ffT contains an allylamine or allylamide linker moiety between the nucleobase and the detectable label; also referred to as double bond or DB linker). In further embodiment, the kit comprises unlabeled ffG, labeled ffA(s), labeled ffC-DB, and labeled ffT-DB described herein. In some further embodiments, at least one type of the nucleotides comprises a base that is attached to a detectable label via a cleavable linker,

and the cleavable linker is , wherein Z is -O-CH2-CH=CH2; n is an integer of 1, 2, 3, 4 or 5; * indicates the attachment point of the cleavable

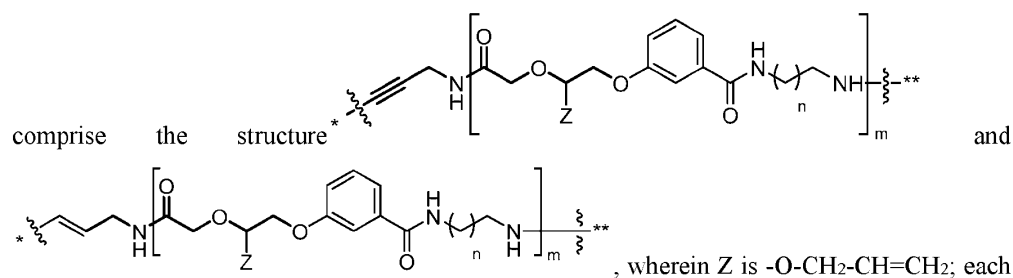
linker to the base; and ** indicates the attachment point of the cleavable linker to the detectable label. In further embodiments, the nucleobase of T nucleotide is attached to the detectable label via the cleavable linker. In some further embodiments, wherein each of at least three of the nucleotides independently comprises a base that is attached to a detectable label via a cleavable linker, and the cleavable linker is selected from the group consisting of:



n is an integer of 1, 2, 3, 4 or 5; * indicates the attachment point of the cleavable linker to the base; and ** indicates the attachment point of the cleavable linker to the detectable label.

[0171] In other embodiments, one or more types of nucleotides A, G, C, and T or U are unlabeled, and wherein the first type of unlabeled nucleotides comprises a first functional moiety, and the kit further comprises a first labeling reagent, wherein the first labeling reagent comprises one or more first detectable labels and a first binding moiety that is capable of specific binding to the first functional moiety of the first type of unlabeled nucleotide. In some such embodiments, the first functional moiety of the first type of unlabeled nucleotide is bound to the first labeling reagent by either covalent bonding or noncovalent interaction via a cleavable linker described herein. In some embodiments, two or more types of nucleotides A, G, C, and T or U are unlabeled. In some further embodiments, each of the four types of nucleotides is unlabeled, and wherein the second type of unlabeled nucleotides comprises a second functional moiety, and the kit further comprises a second labeling reagent, wherein the second labeling reagent comprises one or more second detectable labels and a second binding moiety that is capable of specific binding to the second functional moiety of the second type of unlabeled nucleotide. In some such embodiments, the second functional moiety of the second type of unlabeled nucleotide is bound to the second labeling reagent by either covalent bonding or noncovalent interaction via a cleavable linker described herein. In some further embodiments, the third type of unlabeled nucleotides comprises a third functional moiety, and the kit further comprises a third labeling reagent, wherein the third labeling reagent comprises one or more third detectable labels and a third binding moiety that is capable of specific binding to the third functional moiety of the third type of unlabeled nucleotide. In some such embodiments, the third functional moiety of the third type of unlabeled nucleotide is bound to the third labeling reagent by either covalent bonding or noncovalent interaction via a cleavable linker described herein. In other embodiments, the third

type of unlabeled nucleotides comprises a mixture of the third type of unlabeled nucleotides comprising the first functional moiety and the third type of unlabeled nucleotides comprising the second functional moiety, and wherein both the first labeling reagent and the second labeling reagent are capable of specific binding to the third type of unlabeled nucleotides. In some further embodiments, the fourth type of unlabeled nucleotides is not capable of specific binding with any of the first, second, or third labeling reagent. Post-incorporation labeling kits and methods have been described in U.S. Publication No. 2023/0383342 A1, which is incorporated by reference in its entirety. Non-limiting examples of noncovalent interaction between a functional moiety of the nucleotide and a binding moiety of the labeling reagent include but are not limited to avidin (e.g., streptavidin or neutravidin) and biotin; dinitrophenyl (DNP) moiety and anti-DNP antibody; digoxigenin (DIG) and anti-DIG antibody; β -N-acetyl glucosamine (O-GlcNAc) and WGA (lectin); alkyl guanine moiety and SNAP-Tag®, alkyl chloride moiety and HaloTag®; 3-nitrotyrosine and anti-nitrotyrosine antibody; nickel or cobalt complex such as Ni-nitrilotriacetic acid (NTA) and His-Tag; zinc complex and oligo-aspartate protein. Non-limiting examples of covalent interaction between a functional moiety and a labeling reagent include but are not limited to a biorthogonal reaction selected from the group consisting of a [3+2] dipolar cycloaddition, a Diels-Alder cycloaddition, a [4+1] cycloaddition, a phosphine ligation, or condensation with 2-acylphenyl boronic acid. For example, one of the functional moiety and the binding moiety comprises or is norbornene, transcyclooctene (TCO), dibenzocyclooctyne (DBCO), or bicyclo[6.1.0]nonyne (BCN), and the other one of the functional moiety and the binding moiety comprises or is azido. In some other embodiments, one of the functional moiety and the binding moiety comprises or is TCO, and the other one of the functional moiety and the binding moiety comprises or is an optionally substituted 1,2,4,5-tetrazine moiety. The cleavable linker may



of m and n is independently an integer of 1, 2, 3, 4 or 5; * indicates the attachment point of the cleavable linker to the base; and ** indicates the attachment point of the cleavable linker to the first/second/third functional moiety of the unlabeled nucleotides. In some embodiments m is 2. In some further embodiments, n is 4.

[0172] In a particular embodiment, a kit can include at least one labeled 3' blocked nucleotide or nucleoside together with labeled or unlabeled nucleotides or nucleosides. For

example, nucleotides labeled with dyes may be supplied in combination with unlabeled or native nucleotides, and/or with fluorescently labeled nucleotides or any combination thereof. Combinations of nucleotides may be provided as separate individual components (e.g., one nucleotide type per vessel or tube) or as nucleotide mixtures (e.g., two or more nucleotides mixed in the same vessel or tube).

[0173] Where kits comprise a plurality, particularly two, or three, or more particularly four, 3' blocked nucleotides labeled with a dye compound, the different nucleotides may be labeled with different dye compounds, or one may be dark, with no dye compounds. Where the different nucleotides are labeled with different dye compounds, it is a feature of the kits that the dye compounds are spectrally distinguishable fluorescent dyes. As used herein, the term "spectrally distinguishable fluorescent dyes" refers to fluorescent dyes that emit fluorescent energy at wavelengths that can be distinguished by fluorescent detection equipment (for example, a commercial capillary-based DNA sequencing platform) when two or more such dyes are present in one sample. When two nucleotides labeled with fluorescent dye compounds are supplied in kit form, it is a feature of some embodiments that the spectrally distinguishable fluorescent dyes can be excited at the same wavelength, such as, for example by the same laser. When four 3' blocked nucleotides (A, C, T, and G) labeled with fluorescent dye compounds are supplied in kit form, it is a feature of some embodiments that two of the spectrally distinguishable fluorescent dyes can both be excited at one wavelength and the other two spectrally distinguishable dyes can both be excited at another wavelength. Particular excitation wavelengths are 450 nm to 460 nm, or 520 nm to 532 nm.

[0174] In an alternative embodiment, the kits of the disclosure may contain 3' blocked nucleotides where the same base is labeled with two or more different dyes. A first nucleotide (e.g., 3' allyl blocked T nucleotide triphosphate or 3' blocked G nucleotide triphosphate) may be labeled with a first dye. A second nucleotide (e.g., 3' blocked C nucleotide triphosphate) may be labeled with a second spectrally distinct dye from the first dye, for example a "green" dye absorbing at less than 600 nm, and a "blue" dye absorbs at less than 500 nm, for example 400 nm to 500, in particular 450 nm to 460 nm). A third nucleotide (e.g., 3' blocked A nucleotide triphosphate) may be labeled as a mixture of the first and the second dyes, or a mixture of the first, the second and a third dyes, and the fourth nucleotide (e.g., 3' blocked G nucleotide triphosphate or 3' blocked T nucleotide triphosphate) may be 'dark' and contain no label. In one example, the nucleotides 1-4 may be labeled 'blue', 'green', 'blue/green', and dark. To simplify the instrumentation further, four nucleotides can be labeled with two dyes excited with a single laser, and thus the labeling of nucleotides 1-4 may be 'blue 1', 'blue 2', 'blue 1/blue 2', and dark.

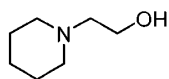
[0175] In particular embodiments, the kits may contain four labeled 3' blocked nucleotides (e.g., A, C, T, G), where each type of nucleotide comprises the same 3' allyl blocking group and a fluorescent label, and wherein each fluorescent label has a distinct fluorescence maximum and each of the fluorescent labels is distinguishable from the other three labels. The kits may be such that two or more of the fluorescent labels have a similar absorbance maximum but different Stokes shift. In some other embodiments, one type of the nucleotide is unlabeled.

[0176] Although kits are exemplified herein in regard to configurations having different nucleotides that are labeled with different dye compounds, it will be understood that kits can include 2, 3, 4 or more different nucleotides that have the same dye compound. In some embodiments, the kit also includes an enzyme and a buffer appropriate for the action of the enzyme. In some such embodiments, the enzyme is a polymerase, a terminal deoxynucleotidyl transferase, or a reverse transcriptase. In particular embodiments, the enzyme is a DNA polymerase, such as DNA polymerase 812 (Pol 812) or DNA polymerase 1901 (Pol 1901). In some further embodiment, the kit may comprise an incorporation mix described herein. In further embodiments, the kit containing the incorporation mix described herein also comprises at least one Pd scavenger (e.g., the Pd(0) scavenger described herein that comprises one or more allyl moieties). In the Pd(0) scavenger comprises one or more allyl moieties each independently selected from the group consisting of -O-allyl, -S-allyl, -NR-allyl, and -N¹RR'-allyl, wherein R is H, unsubstituted or substituted C₁-C₆ alkyl, unsubstituted or substituted C₂-C₆ alkenyl, unsubstituted or substituted C₂-C₆ alkynyl, unsubstituted or substituted C₆-C₁₀ aryl, unsubstituted or substituted 5 to 10 membered heteroaryl, unsubstituted or substituted C₃-C₁₀ carbocyclyl, or unsubstituted or substituted 5 to 10 membered heterocyclyl; and R' is H, unsubstituted C₁-C₆ alkyl or substituted C₁-C₆ alkyl. In some such embodiments, the Pd(0) scavenger in the incorporation solution comprises one or more -O-allyl moieties as described herein in connection with the sequencing methods.

[0177] Other components to be included in such kits may include buffers and the like. The nucleotides of the present disclosure, and other any nucleotide components including mixtures of different nucleotides, may be provided in the kit in a concentrated form to be diluted prior to use. In such embodiments a suitable dilution buffer may also be included. For example, the incorporation mix kit may comprise one or more buffering agents selected from a primary amine, a secondary amine, a tertiary amine, a natural amino acid, or a non-natural amino acid, or combinations thereof. In further embodiments, the buffering agents in the incorporation mix comprise ethanolamine or glycine, or a combination thereof.

[0178] Again, one or more of the components identified in a method set forth herein can be included in a kit of the present disclosure. In some further embodiments, the kit may

comprise a palladium catalyst described herein. In some embodiments, the Pd catalyst is generated by mixing a Pd(II) complex (i.e., a Pd pre-catalyst) with one or more water soluble phosphines described herein. In some such embodiments, the kit containing the Pd catalyst is the cleavage mix kit. In further embodiments, the cleavage mix kit may contain Pd(Allyl)Cl₂ or Na₂PdCl₄ and a water soluble phosphine tris(hydroxyalkyl)phosphine such as THPP to generate the active Pd(0) species. The molar ratio of Pd(II) complex (e.g., Pd(Allyl)Cl₂ or Na₂PdCl₄) to the water soluble phosphine (e.g., THP) may be about 1:1, 1:1.5, 1:2, 1:2.5, 1:3, 1:3.5, 1:4, 1:4.5, 1:5, 1:5.5, 1:6, 1:6.5, 1:7, 1:7.5, 1:8, 1:8.5, 1:9, 1:9.5 or 1:10. In further embodiments, the cleavage mix kit may also contain one or more buffer reagents selected from the group consisting of a primary amine, a secondary amine, a tertiary amine, a carbonate salt, a phosphate salt, and a borate salt, and combinations thereof. Non-limiting example of the buffer reagents in the cleavage mix kit are selected from the group consisting of ethanolamine (EA), tris(hydroxymethyl)aminomethane (Tris), glycine, a carbonate salt, a phosphate salt, a borate salt, dimethylethanolamine (DMEA), diethylethanolamine (DEEA), N,N,N',N'-tetramethylethylenediamine (TEMED), and N,N,N',N'-tetraethylethylenediamine (TEEDA), 2-piperidine ethanol (also known as (2-



hydroxyethyl)piperidine, having the structure), and combinations thereof. In one embodiment, the cleavage mix kit contains DEEA. In other embodiment, the cleavage mix kit contains (2-hydroxyethyl)piperidine.

[0179] In some further embodiments, the kit may comprise one or more palladium scavengers (e.g., a Pd(II) scavenger described herein). In some such embodiments, the Pd(II) scavenger is in the post-cleavage washing buffer. In one embodiment, the post-cleavage washing buffer kit comprises L-cysteine or a salt thereof.

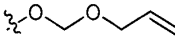
[0180] In any embodiments of the kits described herein, the Pd scavengers (e.g., the Pd(0) or Pd(II) scavengers described herein) are in separate containers/compartments from the Pd catalyst.

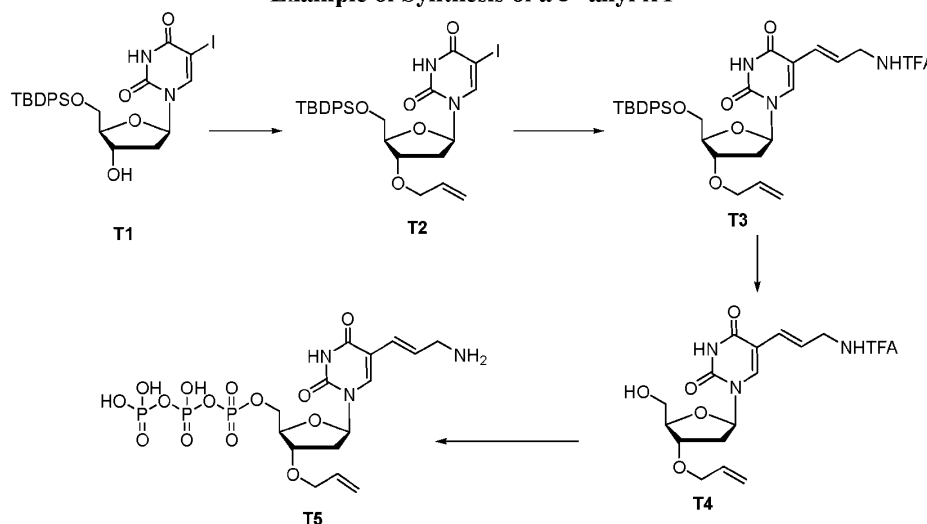
[0181] The present disclosure also provides for a cartridge for use with a sequencing apparatus, comprising a plurality of chambers, where one or more of the plurality of chambers is for use with the kit described herein. For example, the cartridge may contain two or more separate chambers, one chamber contains the aqueous cleavage mixture described herein, and another chamber contains the incorporation mixture described herein.

EXAMPLES

[0182] Additional embodiments are disclosed in further detail in the following examples, which are not in any way intended to limit the scope of the claims.

Example 1. Pd cleavage solution kinetic study

[0183] In this example, the efficiency of a Pd(0) cleavage mixture generated *in situ* from a Na₂PdCl₄ and THPP in a molar ratio of 1:3.5, in a buffered solution containing (2-hydroxyethyl)piperidine at pH of about 9.8. T nucleosides with two different 3' blocking groups were used: 3'-O-Allyl-T-5'-OH and 3'-AOM-T-5'-OH (3'-AOM blocking group has the structure  attached to the 3' carbon atom of the nucleotide). The study compared the effectiveness the Pd cleavage mixture in cleaving the 3' allyl blocking group and the 3'-AOM blocking group. As shown in FIG. 1, the Pd cleavage efficiency over the two blocking groups were very comparable, and the 3' blocking groups cleavage was nearly complete after 30 minutes.

Example 2. Syntheses of 3'-allyl blocked fully functionalized nucleotides (ffNs)**Example of Synthesis of a 3'-allyl ffT**5'-O-(tert-butyldiphenylsilyl)- 3'-O-allyl – 5-iodo-2'-deoxyuridine (T2)

[0184] Sodium hydride (200mg, 5 mmol) was suspended in 5 mL of anhydrous tetrahydrofuran in a flask under nitrogen. Then a solution of 5-Iodo-5'-TBDPS-2'-deoxyuridine **T1** (1 g, 1.68 mmol) in 10 mL of anhydrous THF was added dropwise at 0 °C. The reaction was then brought to room temperature and stirred for 30 minutes, then allyl bromide (430 µL, 5 mmol) was added dropwise at room temperature. The reaction was stirred overnight at room temperature, then quenched with 300 µL of acetic acid, diluted with 100 mL of water and extracted with 100 mL of ethyl acetate. The organic phase was washed with 100 mL of water, 100 mL of brine. The organic phase was dried over MgSO₄, filtered and evaporated to dryness. The residue was purified by flash chromatography on silica gel on a Biotage system (SNAP Sfar 25 g, gradient from 100%

DCM to 75:25 to DCM/EtOAc). Yield: 724 mg (1.14 mmol), 68%. TLC (EtOAc/DCM 2:8): R_f 0.7. ^1H NMR (400 MHz, d_6 -DMSO): δ (ppm) 11.76 (s, 1H, NH), 8.01 (s, 1H, C-6), 7.65 (ddd, J = 7.9, 4.9, 1.7 Hz, 4H, Ph), 7.55 – 7.40 (m, 6H, Ph), 6.04 (dd, J = 8.2, 5.9 Hz, 1H, 1'-CH), 5.87 (ddt, J = 17.2, 10.5, 5.3 Hz, 1H, O-CH₂-CH=), 5.23 (dq, J = 17.3, 1.8 Hz, 1H, CHH=), 5.15 (dq, J = 10.4, 1.5 Hz, 1H, CHH=), 4.11 (dt, J = 5.4, 2.4 Hz, 1H, 3'-CH), 4.06 – 4.00 (m, 1H, 4'-CH), 3.95 (tt, J = 3.9, 1.5 Hz, 2H, CH₂ allyl), 3.85 (dd, J = 11.3, 4.0 Hz, 1H, 5'-CHH), 3.73 (dd, J = 11.3, 4.5 Hz, 1H, 5'-CHH), 2.36 – 2.27 (m, 1H, 2'-CHH), 2.22 (ddd, J = 13.9, 8.3, 6.1 Hz, 1H, 2'-CHH), 1.03 (s, 9H, *t*Bu). LC-MS (ES and CI): (positive ion) m/z 633 ($M+H^+$), 655 ($M+Na^+$); (negative ion) m/z 631 ($M-H^+$).

5'-O-(*tert*-butyldiphenylsilyl)-3'-O-allyl-5-[3-(2,2,2-trifluoroacetamido)-allyl]-2'-deoxyuridine (T3)

[0185] 5'-O-(*tert*-butyldiphenylsilyl)-3'-O-allyl-5-iodo-2'-deoxyuridine (T2) (401 mg, 0.634 mmol), 2-Dicyclohexylphosphino-2',4',6'-triisopropylbiphenyl (24 mg, 0.05 mmol, XPhos), [3-(2,2,2-trifluoroacetamido)]-allylboronic acid pinacol ester (354 mg, 1.27 mmol) and allylpalladium(II) chloride dimer (9 mg, 0.025 mmol) were dissolved in 6 mL of anhydrous dimethylformamide and degassed by bubbling with nitrogen gas for 10 minutes. A solution of Cs₂CO₃ (412 mg, 1.27 mmol) in 0.6 mL of water was then added and the reaction was heated to 45 °C. After 1 hour the reaction was diluted with 100 mL of ethyl acetate and washed with 2x 100 mL of water and 100 mL of brine. The organic phase was dried over MgSO₄, filtered and evaporated to dryness. The residue was purified by flash chromatography on silica gel on a Biotage system (SNAP Sfar 25 g, gradient from 100% DCM to 7:3 to DCM/EtOAc). Yield: 282 mg (0.429 mmol), 68%. TLC (EtOAc/DCM 2:8): R_f 0.5. ^1H NMR (400 MHz, d_6 -DMSO): δ (ppm) 11.52 (s, 1H, NH), 9.62 (t, J = 5.7 Hz, 1H), 7.95 (s, 1H, C-6), 7.70 – 7.59 (m, 4H, Ph), 7.53 – 7.38 (m, 6H, Ph), 6.45 (dt, J = 15.8, 6.3 Hz, 1H, Ar-CH=CH), 6.11 (t, J = 7.1 Hz, 1H, 1'-CH), 5.99 (d, J = 15.9 Hz, 1H, Ar-CH=), 5.89 (ddt, J = 17.3, 10.6, 5.3 Hz, 1H, O-CH₂-CH=), 5.26 (dq, J = 17.3, 1.8 Hz, 1H, CHH=), 5.16 (dq, J = 10.4, 1.5 Hz, 1H, CHH=), 4.17 (dt, J = 6.3, 3.3 Hz, 1H, 3'-CH), 4.08 – 3.94 (m, 3H, 4'-CH, O-CH₂-All), 3.88 (dd, J = 11.1, 4.7 Hz, 1H, 5'-CHH), 3.77 (m, 3H, 5'-CHH, CH₂-NHTFA), 2.34 – 2.25 (m, 2H, 2'-CH₂), 1.01 (s, 9H, *t*Bu). LC-MS (ES and CI): (positive ion) m/z 658 ($M+H^+$); (negative ion) m/z 656 ($M-H^+$).

3'-O-allyl-5-[3-(2,2,2-trifluoroacetamido)-allyl]-2'-deoxyuridine (T4)

[0186] 5'-O-(*tert*-butyldiphenylsilyl)-3'-O-allyl-5-[3-(2,2,2-trifluoroacetamido)-allyl]-2'-deoxyuridine (**T3**) (470 mg, 0.743 mmol) was dissolved in dry THF (6 mL) under N₂ atmosphere, then a solution of 1.0 M TBAF in THF (390 μ L, 0.89 mmol) was added. The solution

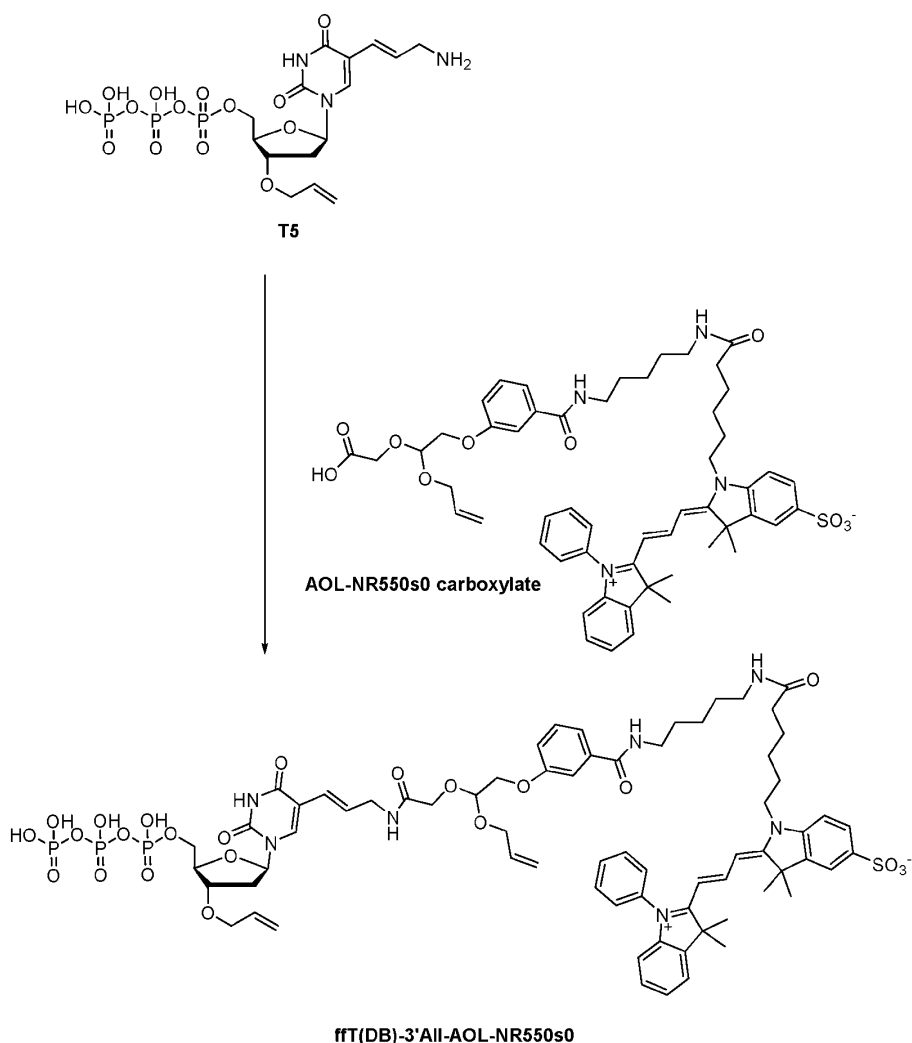
was stirred at room temperature for 18 hours then TLC (EtOAc/DCM 1:1) showed completion. The solution was diluted with 100 mL of EtOAc, then washed with 50 mL of NaH₂PO₄ sat. (pH=3), and with 2x 50 mL of water. The organic phase was dried over MgSO₄, filtered and evaporated to dryness. The residue was purified by flash chromatography on silica gel on a Biotage system (SNAP Sfar 10g, gradient from 100% DCM to 1:1 to DCM/EtOAc). Yield: 218 mg (520 mmol), 70%. TLC (EtOAc/DCM 1:1): R_f 0.5. ¹H NMR (400 MHz, *d*₆-DMSO): δ (ppm) 11.47 (s, 1H, NH), 9.68 (t, *J* = 5.7 Hz, 1H, NHTFA), 8.03 (s, 1H, H-6), 6.47 (dt, *J* = 15.9, 6.2 Hz, 1H, Ar-CH=CH), 6.21 (d, *J* = 16.0 Hz, 1H, Ar-CH=), 6.12 (dd, *J* = 7.9, 6.0 Hz, 1H, 1'-CH) 5.90 (ddt, *J* = 17.3, 10.5, 5.3 Hz, 1H, O-CH₂-CH=), 5.28 (dq, *J* = 17.3, 1.8 Hz, 1H, CHH=), 5.16 (m, 2H, CHH=, 5'-OH), 4.13 (dt, *J* = 5.5, 2.6 Hz, 1H, 3'-CH), 4.03 – 3.94 (m, 3H, 4'-CH, O-CH₂-All), 3.88 (t, *J* = 6.0 Hz, 2H, CH₂-NHTFA), 3.67 – 3.54 (m, 2H, 5'-CH₂), 2.28 (ddd, *J* = 13.6, 6.0, 2.7 Hz, 1H, 2'-CH₂), 2.20 (ddd, *J* = 13.7, 7.9, 5.8 Hz, 1H, 2'-CH₂). LC-MS (ES and CI): (positive ion) *m/z* 420 (M+H⁺), 442 (M+Na⁺); (negative ion) *m/z* 418 (M-H⁺).

5'-O-triphosphate-3'-O-allyloxymethyl - 5-(3-aminoallyl)-2'-deoxyuridine (T5)

[0187] 3'-O-allyl - 5-[3-(2,2,2-trifluoroacetamido)-allyl]-2'-deoxyuridine (**T4**) (210 mg, 0.501 mmol,) was dried under reduced pressure over P₂O₅ for 18 hrs. Anhydrous triethyl phosphate (2 mL) and some freshly activated 4 Å molecular sieves were added to it under nitrogen, then the reaction flask was cooled to -15 °C in an ice/salt bath. Freshly distilled POCl₃ (56 µL, 0.601 mmoles) was added dropwise followed by Proton Sponge® (160 mg, 0.75 mmol). After the addition, the reaction was further stirred at 0°C for 15 minutes. Then, pyrophosphate as bis-tri-*n*-butylammonium salt (1.3 g, 2.50 mmol) as 0.5 M solution in anhydrous DMF was quickly added, followed immediately by tri-*n*-butyl amine (537 µL, 2.25 mmol). The reaction was removed from the bath and stirred at room temperature for 10 minutes, then quenched into 1 M aqueous triethylammonium bicarbonate (TEAB, 20 mL) and further stirred at room temperature for 4 hours. Then the solution was extracted with 30 mL of ethyl acetate. The aqueous phase was evaporated under reduced pressure, then a 35% aqueous solution of ammonia (20 mL) was added to the residue and the mixture was stirred at room temperature for 4 hours. The solvents were then evaporated under reduced pressure, the residue resuspended in 10 mL of 0.1 M TEAB and filtered. The filtrate was purified firstly by ion-exchange chromatography on DEAE-Sephadex A25 (100 g). The column was eluted with a linear gradient of aqueous triethylammonium bicarbonate (TEAB, from 0.1 M to 0.8 M over 1 L). The fractions containing the triphosphate were pooled and the solvent was evaporated to dryness under reduced pressure. The crude material was further purified by preparative scale HPLC using a YMC-Pack-Pro C18 column, eluting with 0.1 M TEAB and acetonitrile. Compound **T5** was obtained as triethylammonium salt. Yield: 195 µmol

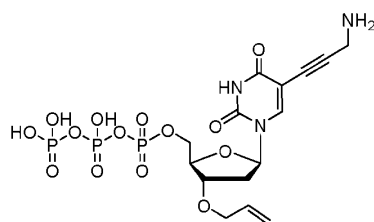
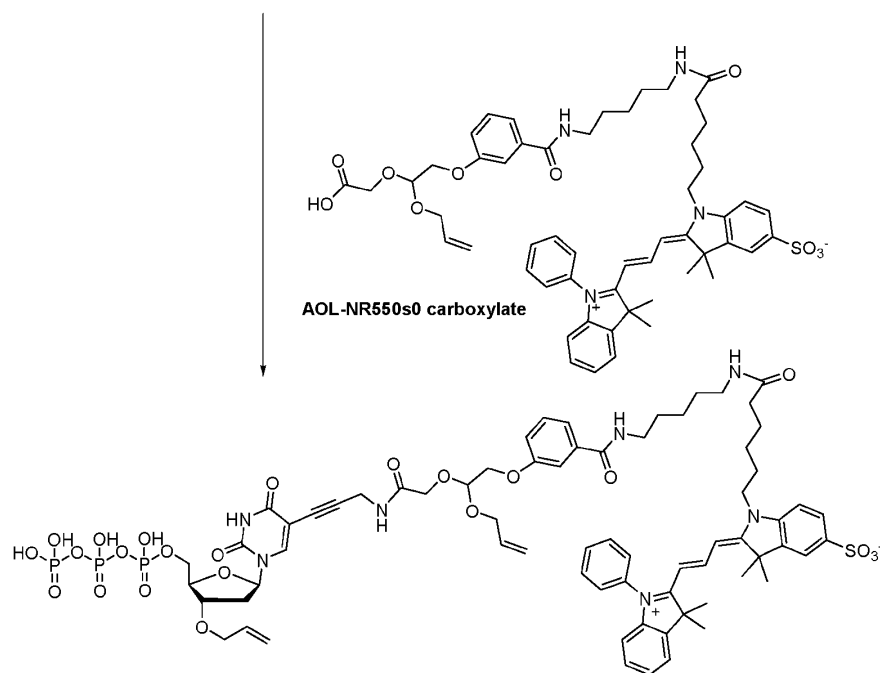
(39%). ^1H NMR (400 MHz, D_2O): δ (ppm) 8.15 (s, 1H, H-6), 6.50 (d, $J = 15.9$ Hz, 1H, Ar-CH=), 6.38 (dt, $J = 16.0$, 6.5 Hz, 1H, Ar-CH=CH-), 6.29 (dd, $J = 8.2$, 5.9 Hz, 1H, 1'-CH), 5.91 (ddt, $J = 17.0$, 10.3, 5.9 Hz, 1H O-CH₂-CH=), 5.30 (dq, $J = 17.3$, 1.5 Hz, 1H, CHH=), 5.24 – 5.16 (m, 1H, CHH=), 4.41 (dt, $J = 5.9$, 2.0 Hz, 1H, 3'-CH), 4.33 (m, 1H, 4'-CH), 4.22 (ddd, $J = 11.8$, 4.1, 2.4 Hz, 1H, 5'-CHH), 4.15 (ddd, $J = 11.9$, 6.2, 2.5 Hz, 1H, 5'-CHH), 4.07 (ddt, $J = 5.8$, 2.3, 1.3 Hz, 2H, =CH-CH₂-O), 3.65 (dd, $J = 6.6$, 1.1 Hz, 2H, CH₂-NH₂), 3.12 (q, $J = 7.3$ Hz, Et₃N⁺), 2.47 (ddd, $J = 14.3$, 6.0, 2.1 Hz, 1H, 2'-CHH), 2.27 (ddd, $J = 14.2$, 8.3, 5.8 Hz, 1H, 2'-CHH), 1.20 (t, $J = 7.3$ Hz, Et₃N⁺). LC-MS (ES and CI): (negative ion) m/z 562 (M-H^+).

Synthesis of ffT(DB)-3'All-AOL-NR550s0



[0188] AOL-NR550s0 carboxylate (0.012 mmol) was coevaporated with 2x2 mL of anhydrous *N,N'*-dimethylformamide (DMF), then dissolved in 1.5 mL of anhydrous *N,N'*-

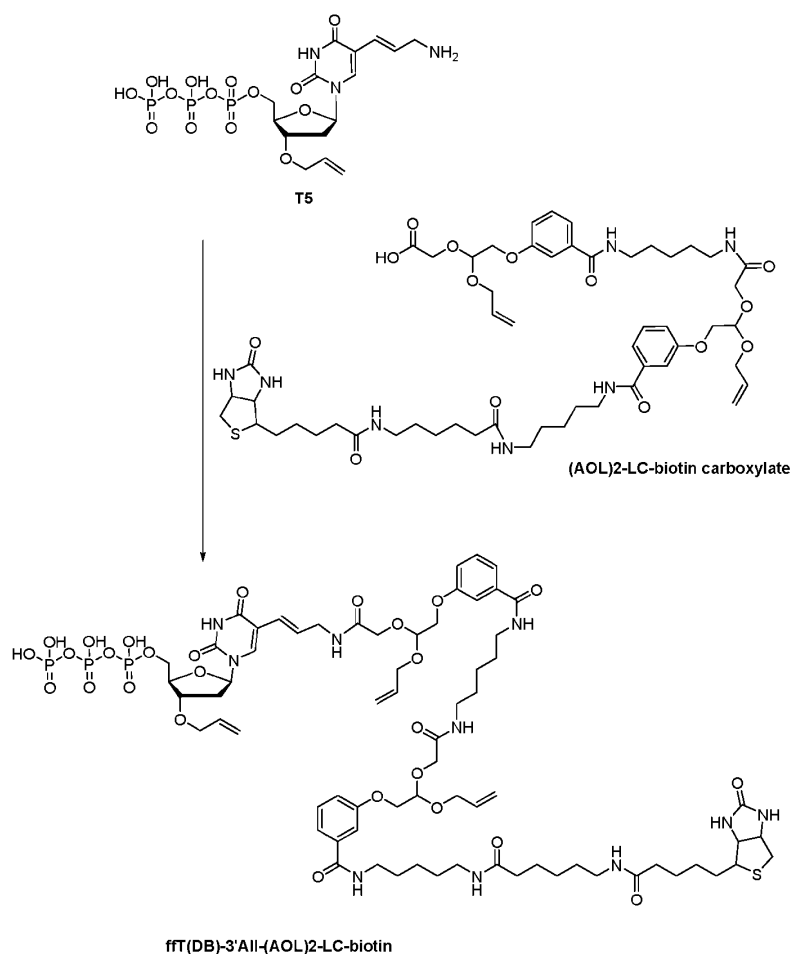
dimethylacetamide (DMA). *N,N*-diisopropylethylamine (17 μ L, 0.1 mmol) was added, followed by *N,N,N',N'*-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate (TSTU, 140 μ L of a 100 mM solution in DMF, 0.014 mmol). The reaction was stirred under nitrogen at room temperature for 1 hour. In the meantime, an aqueous solution of the triphosphate **T5** (0.01 mmol) was evaporated to dryness under reduced pressure and resuspended in 150 μ L of 0.1 M triethylammonium bicarbonate (TEAB) solution in water. The activated AOL-NR550s0 carboxylate solution was added to the triphosphate and the reaction was stirred at room temperature for 18 hours and monitored by RP-HPLC. The crude product was purified firstly by ion-exchange chromatography on DEAE-Sephadex A25 (25 g) eluting with a gradient from 0.1M TEAB/acetonitrile 8:2 to 1M TEAB/acetonitrile 8:2. The fractions containing the triphosphate were pooled and the solvent was evaporated to dryness under reduced pressure. The crude material was further purified by preparative scale RP-HPLC using a YMC-Pack-Pro C18 column, eluting with 0.1 M TEAB and acetonitrile. LC-MS (ES): (negative ion) m/z 1526 ($M-2H^++Na^+$), 752 ($M-2H^+$). RP-HPLC (YMC-Packpro C18, 250x4.6 mm, gradient from 5% to 60% acetonitrile in 0.1 M TEAB, 20 minutes, flow rate 1 mL/min): t_R = 18.3 min.

Synthesis of fT(PA)-3'All-AOL-NR550s0**5'-O-triphosphate-3'-O-allyl - 5-(3-aminopropargyl)-2'-deoxyuridine****fT(PA)-3'All-AOL-NR550s0**

[0189] AOL-NR550s0 carboxylate (0.015 mmol) was coevaporated with 2x 2 mL of anhydrous *N,N'*-dimethylformamide (DMF), then dissolved in 2 mL of anhydrous *N,N'*-dimethylacetamide (DMA). *N,N*-diisopropylethylamine (17 μ L, 0.1 mmol) was added, followed by *N,N,N',N'*-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate (TSTU, 180 μ L of a 100 mM solution in DMF, 0.018 mmol). The reaction was stirred under nitrogen at room temperature for 1 hour. In the meantime, an aqueous solution of the 5'-*O*-triphosphate-3'-*O*-allyl - 5-(3-aminopropargyl)-2'-deoxyuridine (0.01 mmol) was evaporated to dryness under reduced pressure and resuspended in 200 μ L of 0.1 M triethylammonium bicarbonate (TEAB) solution in water. The activated Dye-linker solution was added to the triphosphate and the reaction was stirred at room temperature for 18 hours and monitored by RP-HPLC. The crude product was purified firstly by ion-exchange chromatography on DEAE-Sephadex A25 (25 g) eluting with a gradient from

0.1M TEAB/acetonitrile 8:2 to 1M TEAB/acetonitrile 8:2. The fractions containing the triphosphate were pooled and the solvent was evaporated to dryness under reduced pressure. The crude material was further purified by preparative scale RP-HPLC using a YMC-Pack-Pro C18 column, eluting with 0.1 M TEAB and acetonitrile. Yield: 65 μ mol, (65%). LC-MS (ES): (negative ion) m/z 1502 ($M-H^+$), 750 ($M-2H^+$). RP-HPLC (YMC-Packpro C18, 250x4.6 mm, gradient from 5% to 60% acetonitrile in 0.1M TEAB, 20 minutes, flow rate 1 mL/min): t_R 18.2 min.

Synthesis of fffT(DB)-3'-All-(AOL)2-LC-Biotin

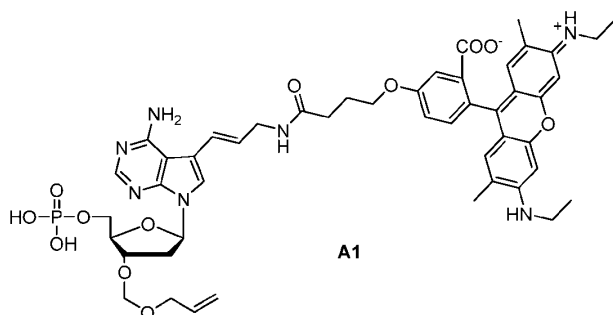


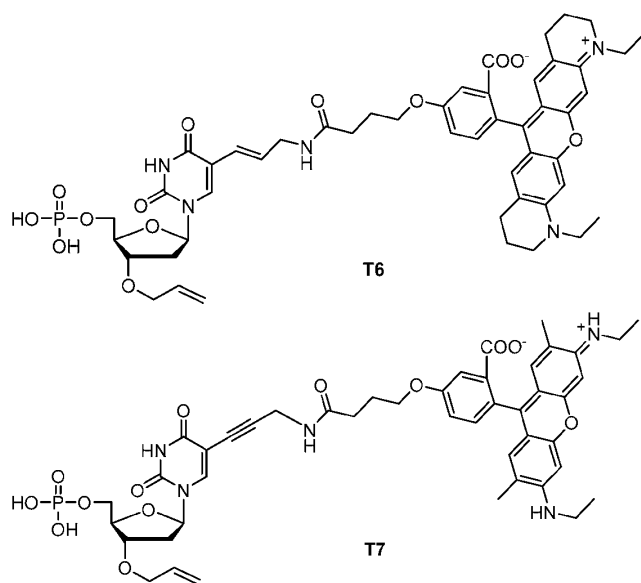
[0190] (AOL)2-LC-biotin carboxylate (0.01 mmol) was coevaporated with 2x2 mL of anhydrous *N,N'*-dimethylformamide (DMF), then dissolved in 1mL of anhydrous *N,N'*-dimethylacetamide (DMA). *N,N*-diisopropylethylamine (8.7 μ L, 0.05 mmol) was added, followed by *N,N,N',N'*-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate (TSTU, 110 μ L of a 100 mM solution in DMF, 0.011 mmol). The reaction was stirred under nitrogen at room temperature for 1 hour. In the meantime, an aqueous solution of the triphosphate **T5** (0.005 mmol) was

evaporated to dryness under reduced pressure and resuspended in 100 μ L of 0.1 M triethylammonium bicarbonate (TEAB) solution in water. The activated (AOL)2-LC-biotin carboxylate solution was added to the triphosphate and the reaction was stirred at room temperature for 18 hours and monitored by RP-HPLC. The crude product was purified by preparative scale RP-HPLC using a YMC-Pack-Pro C18 column, eluting with 0.1 M TEAB and acetonitrile. LC-MS (ES): (negative ion) m/z 1625 ($M-H^+$), 812 ($M-2H^+$). RP-HPLC (YMC-Packpro C18, 250x4.6 mm, gradient from 5% to 50% acetonitrile in 0.1 M TEAB, 20 minutes, flow rate 1 mL/min): t_R 17.9 min.

Example 3. 3'-Allyl cleavage study by palladium catalyst

[0191] The rate of 3'-allyl deprotection by palladium in solution was measured using the model compounds **T6** and **T7** and compared with the rate of deprotection of the 3'-AOM model compound **A1**. In a typical 400 μ L reaction, to a 0.1 mM solution of substrate in 0.1 M DEEA buffer pH 9.8 was added a solution of Pd(0)[tris(3-hydroxypropyl)phosphine]_x complex to a final palladium concentration of 1 mM. The reaction was incubated at room temperature in a closed tube and 20 μ L aliquots of the reaction were taken at set time points, quenched immediately with 10 μ L of 0.25 M H₂O₂/EDTA 1:1, then analyzed by UHPLC for the formation of the respective 3'-OH product. The formation of the 3'-OH product over time is shown in **FIG. 2**. The data shows that the rate of deprotection of compound **T6** (3'-allyl (DB)), which contains the 3'-allyl blocking group and the double bond pendant arm is approximately 1.5 times slower compared to rate of deprotection of compound **A1** (3'-AOM). With compound **T7** (3'-allyl (PA)), however, incomplete deprotection was observed, even after 30 minutes of incubation. Several side-products were observed in the UHPLC traces generated from **T7**, indicating the triple bond pendant arm on thymidine is not inert to the palladium reagent used.

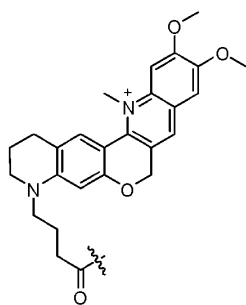




Example 4. Sequencing experiments with ffT(PA)-3' All-AOL-NR550s0 and ffT(DB)-3' Allyl-AOL-NR550s0

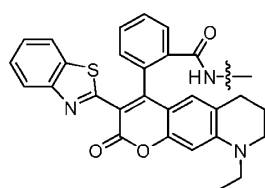
[0192] The sequencing compatibility of the 3'-allyl fully functional T, ffT(PA)-3'-All-AOL-NR550s0 and ffT(DB)-3' All-AOL-NR550s0, was tested on an Illumina iSeq100™ using blue/green 2-excitation/1-emission chemistry and a corresponding set of dyes (blue ffC, dual blue/green ffA and dark ffG). In addition to the ffT, the following nucleotides were included in the incorporation mix: Dark G, ffA(DB)-AOM-AOL-chromenoquinoline dye A, ffC(DB)-AOM-AOL-coumarin dye B. The ffA, ffC and ffG used in the experiments each contained a 3'-AOM as the 3' blocking group. A Pd(0)[THP]_x solution was used as cleavage reagent, and 10 mM sodium thiosulfate and lipoic acid were used as scavengers in the post-cleavage washes. A PhiX library was used as the template in individual single read 1x150 cycles experiments. The primary metrics of the sequencing experiments are shown in **FIGs. 3A** and **3B**, as well as in Table 1, in comparison with a control experiment performed with a 3'-AOM ffT (ffT(DB)-3' AOM-AOL-NR550s0).

[0193] Chromenoquinoline dye A is disclosed in U.S. Publication No. 2022/0195517 A1 (incorporated by reference). Chromenoquinoline dye A has the structure moiety:



when conjugated with a nucleotide.

[0194] Coumarin dye B is disclosed in U.S. Publication No. 2018/0094140 A1 (incorporated by reference). Coumarin dye B has the structure moiety:



when conjugated with a nucleotide.

[0195] The fluorescence intensity (in both green and blue channels) suffers a steeper decline when using ffT(PA)-3'All-AOL-NR550s0, while it is comparable between ffT(DB)-3'All-AOL-NR550s0 and ffT(DB)-3'AOM-AOL-NR550s0 experiments. Similar trends were observed in the %Q30. %PhiX error rate and phasing by cycle suffered a steeper increase with the compound ffT(PA)-3'All-AOL-NR550s0. The data suggests that the double bond pendant arm on T is essential for achieving high quality sequencing metrics with a palladium-based cleave mix.

Table 1.

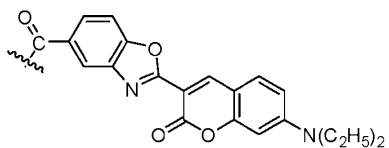
ffT	Phasing	Prephasing	%PhiX Error (150cycles)	%Q30	%Q30 (last 10 cycles)
ffT(PA)-3'All-AOL-NR550S0	0.21	0.11	3.5	78.9	57.3
ffT(DB)-3'All-AOL-NR550S0	0.11	0.11	0.9	88.8	85.8
ffT(DB)-3'AOM-AOL-NR550S0	0.03	0.08	0.7	91.8	89.8

Example 5. Sequencing experiments with ffT(DB)-3'All-(AOL)₂-Biotin

[0196] The sequencing compatibility of a 3'-allyl fully functional T, ffT(DB)-3'All-(AOL)₂-LC-biotin in a post-incorporation labeling workflow was tested on a modified Illumina MiSeq® using blue/green 2-excitation/2-emission chemistry and a corresponding set of dyes (blue ffA, blue and green ffC and dark ffG). In addition to the ffT, the following nucleotides were included in the incorporation mix for the post-incorporation labeling workflow: Dark G, ffA(DB)-AOM-AOL-BL-coumarin dye C (blue dye), ffA(DB)-AOM-AOL-AF670POPO (a known green dye), ffC(DB)-AOM-AOL-coumarin dye C (blue dye), ffC(DB)-AOM-AOL-NR550s0. The ffA,

ffC and ffG used in the experiments each contained a 3'-AOM as the 3' blocking group. An additional step was included after the incorporation step to label the DNA-incorporated ffT(DB)-3'-All-(AOL)2-LC-biotin with multiple green fluorophore labeled streptavidin conjugate (NR550C4)₆-streptavidin. A Pd(0)[THP]_x solution was used as cleavage reagent, and 10 mM sodium thiosulfate was included in the SBS wash buffer. A PhiX library was used as the template in individual single read 1x150 cycles experiments. The primary metrics of the sequencing experiments are shown in **FIGs. 4A** and **4B**, as well as Table 2, in comparison with a control experiment performed with a 3'-AOM ffT (ffT(DB)-3'AOM-AOL-AF550POPOs0). For the control SBS experiments, the following ffT was used: ffT(DB)-AOM-AOL-AF550POPOs0 (green dye) and the remaining ffNs were the same as the post-incorporation labeling workflow.

[0197] Coumarin dye C has strong fluorescence and great stability. This dye is disclosed in U.S. Publication No. 2020/0277670 A1 (incorporated by reference), having the



structure moiety CN(C)Cc1ccc2c(c1)oc(=O)c3cc4c2oc(=O)c4c3 when conjugated with the ffA.

[0198] The fluorescence intensity (in both green and blue channels) and the signal decay are comparable between ffT(DB)-3'-All-(AOL)2-LC-Biotin labelled post-incorporation with a green dye-streptavidin conjugate and the standard ffT(DB)-3'AOM-AOL-AF550POPOs0. Similar trends were observed in the %Q30, %PhiX error rate and phasing metrics by cycle. The data suggests that 3'-allyl blocking group in combination the double bond pendant arm on T is compatible with a sequencing-by-synthesis workflow including a post-incorporation labelling step with a dye-labelled conjugate.

Table 2.

ffT	Phasing	Prephasing	%PhiX Error (150cycles)	%Q30	%Q30 (last 10 cycles)
ffT(DB)-3'AOM-AOL-AF550POPOs0	0.050	0.060	0.71	97.58	94.57
ffT(DB)-3'-All-(AOL)2-LC-Biotin	0.054	0.052	0.53	97.45	93.38

WHAT IS CLAIMED IS:

1. A method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

(i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;

(ii) the solid support comprises a plurality of different target polynucleotides; and

(iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein each nucleotide comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;

(c) imaging the solid support to determine the identity of incorporated nucleotides;

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove 3' allyl blocking groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support;

(e) contacting said solid support with an aqueous wash solution; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences.

2. The method of claim 1, wherein the aqueous solution comprising DNA polymerase in step (b) further comprises a palladium scavenger.

3. The method of claim 2, wherein the palladium scavenger in step (b) is a Pd(0) scavenger.

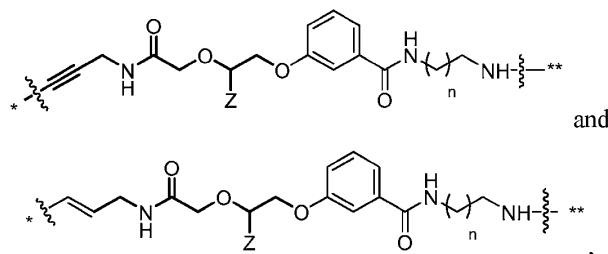
4. The method of any one of claims 1 to 3, wherein the aqueous wash solution in step (e) further comprises a palladium scavenger.

5. The method of claim 4, wherein the palladium scavenger in step (e) is a Pd(II) scavenger.

6. The method of any one of claims 1 to 5, wherein the aqueous deblocking solution further comprises ascorbate.

7. The method of any one of claims 1 to 6, wherein at least one type of nucleotide comprises a base attached to a detectable label via a cleavable linker.

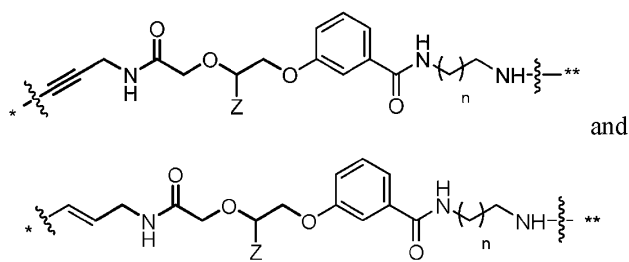
8. The method of claim 7, wherein the detectable label is a fluorescent dye, and the cleavable linker is selected from the group consisting of:



wherein Z is $-O-CH_2-CH=CH_2$; n is an integer of 1, 2, 3, 4 or 5; * indicates the attachment point of the cleavable linker to the base; and ** indicates the attachment point of the cleavable linker to the detectable label.

9. The method of any one of claims 1 to 6, wherein at least three types of nucleotides comprise a base attached to a detectable label via a cleavable linker.

10. The method of claim 9, wherein the detectable label of each of the at least three types of nucleotides is distinguishable from the other detectable labels, and the cleavable linker is selected from the group consisting of:



wherein Z is $-O-CH_2-CH=CH_2$; n is an integer of 1, 2, 3, 4 or 5; * indicates the attachment point of the cleavable linker to the base; and ** indicates the attachment point of the cleavable linker to the detectable label.

11. A method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

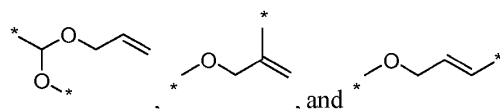
(i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;

(ii) the solid support comprises a plurality of different target polynucleotides; and

(iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C, and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein:

(i) each of at least three types of nucleotides independently comprises a base that is attached to a detectable label via a cleavable linker, and the cleavable linker comprises a moiety selected from the group consisting of:



, and * indicates where the moiety is connected to the remainder of the nucleotide;

(ii) each nucleotide comprises a 2' deoxyribose moiety with a 3' allyl group attached to the 3' oxygen atom; and

(iii) the DNA polymerase is an altered archaeal DNA polymerase;

(c) contacting the solid support with a solution comprising one or more radical scavengers and imaging the solid support to determine the identity of incorporated nucleotides;

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl blocking groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers;

(e) contacting said solid support with an aqueous wash solution comprising a palladium scavenger; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences.

12. The method of claim 11, wherein the palladium scavenger in the aqueous wash solution of step (e) is a Pd(II) scavenger.

13. The method of claim 11 or 12, wherein the aqueous solution comprising DNA polymerase in step (b) further comprises a palladium scavenger.

14. The method of claim 13, wherein the palladium scavenger in step (b) is a Pd(0) scavenger.

15. A method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

- (i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;
- (ii) the solid support comprises a plurality of different target polynucleotides; and
- (iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous solution comprising DNA polymerase and nucleotides A, G, C, and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein:

- (i) at least one of the nucleotides comprise a base that is attached to a detectable label via a cleavable linker; and

- (ii) the nucleotides each comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;

(c) contacting the solid support with a solution comprising one or more radical scavengers and imaging the solid support to determine the identity of incorporated nucleotides;

(d) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers;

(e) contacting said solid support with an aqueous wash solution; and

(f) repeating steps (b)-(e) to determine target polynucleotide sequences.

16. The method of claim 15, wherein the aqueous solution comprising DNA polymerase in step (b) further comprises a palladium scavenger.

17. The method of claim 16, wherein the palladium scavenger in step (b) is a Pd(0) scavenger.

18. The method of claim 16 or 17, wherein the aqueous wash solution in step (e) further comprises a palladium scavenger.

19. The method of claim 18, wherein the palladium scavenger in step (e) is a Pd(II) scavenger.

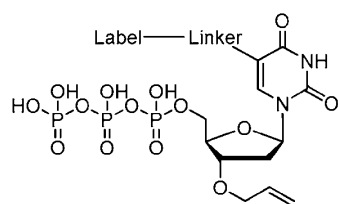
20. The method of any one of claims 1 to 19, wherein the tris(hydroxyalkyl)phosphine is tris(hydroxypropyl)phosphine (THPP).

21. The method of any one of claims 1 to 20, wherein the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise concatemers comprising said multiple copies of target polynucleotides.

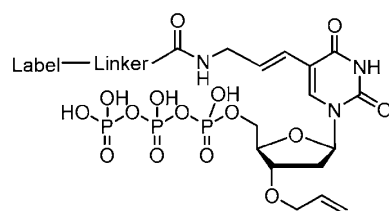
22. The method of any one of claims 1 to 20, wherein the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise clusters of immobilized nucleic acid molecules comprising said multiple copies of target polynucleotides.

23. The method of any one of claims 1 to 22, wherein the bases for the A and G nucleotides are deazapurines.

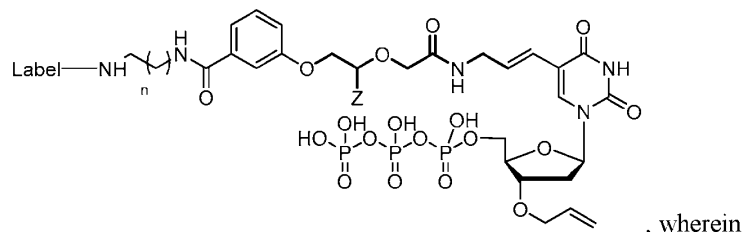
24. The method of any one of claims 1 to 23, wherein the T nucleotide has the structure:



25. The method of claim 24, wherein the T nucleotide has the structure:

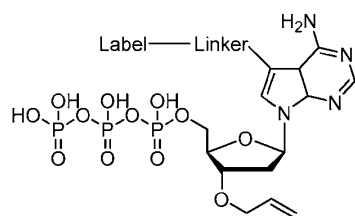


26. The method of claim 24 or 25, wherein the T nucleotide has the structure:

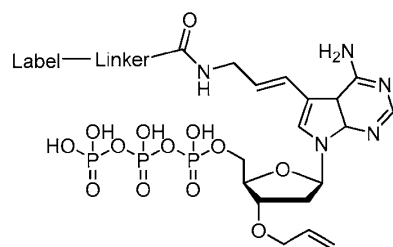


Z is -O-CH₂-CH=CH₂, and n is an integer of 1, 2, 3, 4 or 5.

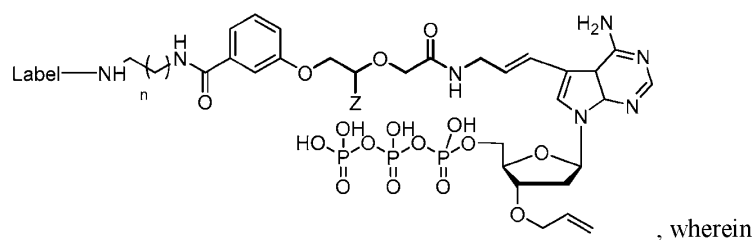
27. The method of any one of claims 1 to 26, wherein the A nucleotide has the structure:



28. The method of claim 27, wherein the A nucleotide has the structure:

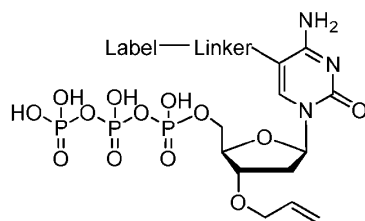


29. The method of claim 27 or 28, wherein the A nucleotide has the structure:

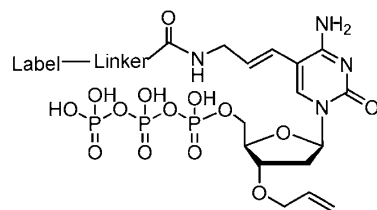


Z is $-O-CH_2-CH=CH_2$, and n is an integer of 1, 2, 3, 4 or 5.

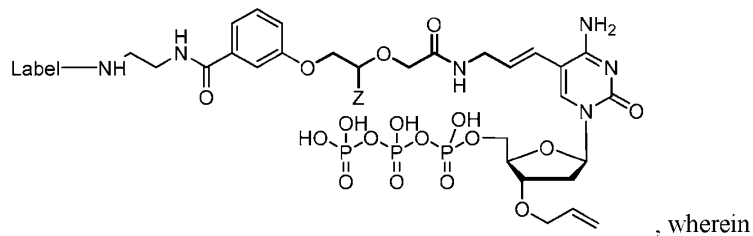
30. The method of any one of claims 1 to 29, wherein the C nucleotide has the structure:



31. The method of claim 30, wherein the C nucleotide has the structure:

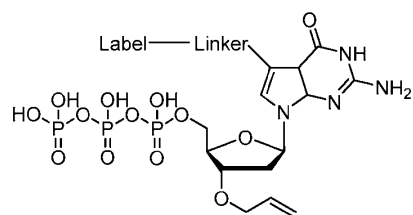


32. The method of claim 30 or 31, wherein the C nucleotide has the structure:

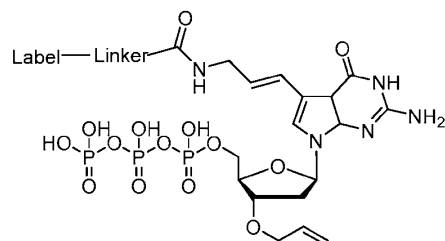


Z is $-O-CH_2-CH=CH_2$, and n is an integer of 1, 2, 3, 4 or 5.

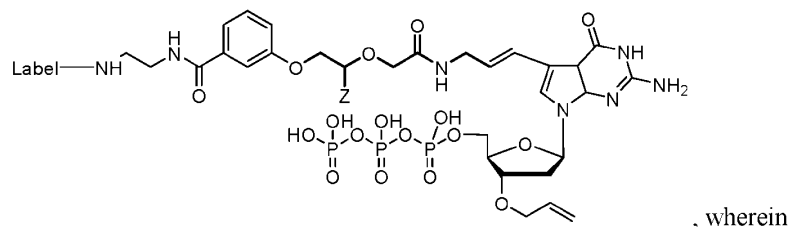
33. The method of any one of claims 1 to 32, wherein the G nucleotide has the structure:



34. The method of claim 33, wherein the G nucleotide has the structure:

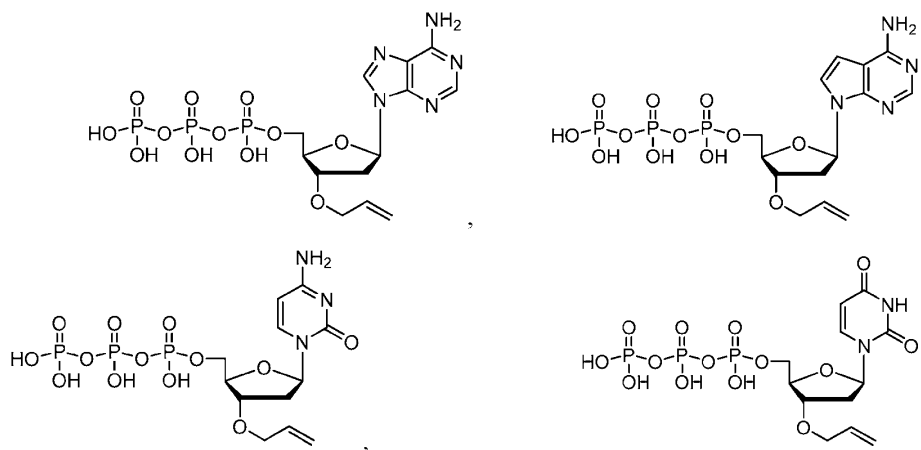


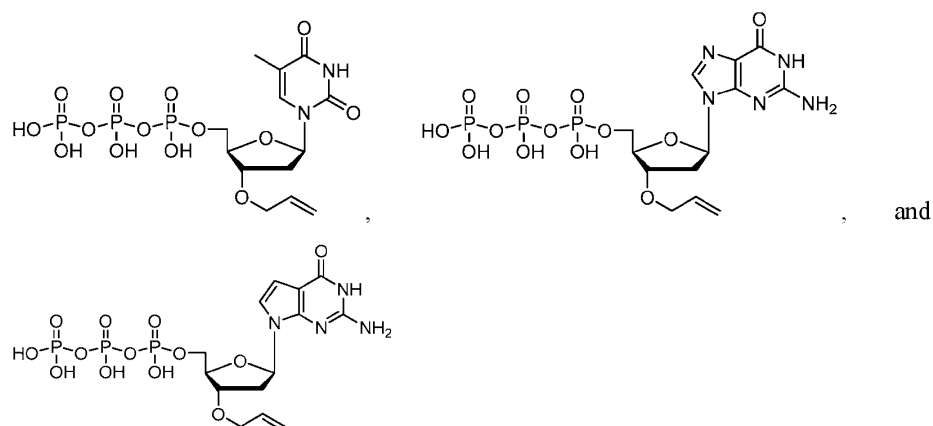
35. The method of claim 33 or 34, wherein the G nucleotide has the structure:



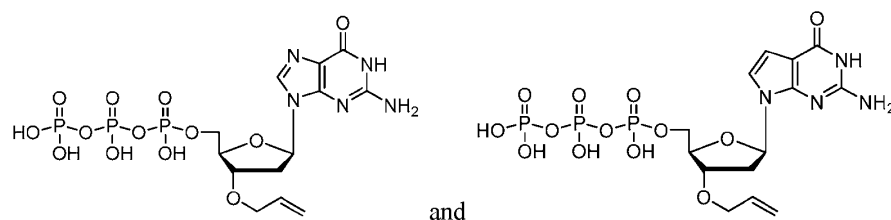
Z is $-O-CH_2-CH=CH_2$, and n is an integer of 1, 2, 3, 4 or 5.

36. The method of any one of claims 1 to 22, wherein at least one type of nucleotide has a structure selected from the group consisting of:





37. The method of claim 36, wherein at least one type of nucleotide has a structure selected from the group consisting of:



38. A method of determining the sequence of a plurality of different target polynucleotides in parallel, the method comprising:

(a) contacting a solid support with a solution comprising sequencing primers under hybridization conditions, wherein:

- (i) the solid support comprises at least 5,000,000 spatially distinguishable sites/cm² that comprise multiple copies of target polynucleotides;
- (ii) the solid support comprises a plurality of different target polynucleotides; and
- (iii) the sequencing primers are complementary to at least a portion of the different target polynucleotides;

(b) contacting the solid support with an aqueous incorporation mixture comprising DNA polymerase and one or more of four types of nucleotides A, G, C, and T or U under conditions suitable for DNA polymerase-mediated primer extension, wherein:

- (i) the nucleotides each comprises a 2' deoxyribose moiety with a 3' allyl blocking group  attached to the 3' oxygen atom;
- (ii) at least two types of nucleotides are unlabeled; and

(iii) the first type of unlabeled nucleotides comprises a first functional moiety;

(c) contacting the extended copy polynucleotides with an aqueous labeling mixture comprising a first labeling reagent, wherein the first labeling reagent comprises one or more first detectable labels and a first binding moiety that is capable of specific binding to the first functional moiety of the first type of unlabeled nucleotide;

(d) imaging the solid support and performing one or more fluorescent measurements to determine the identity of incorporated nucleotides;

(e) contacting the solid support with an aqueous deblocking solution comprising a palladium catalyst and tris(hydroxyalkyl)phosphine under conditions suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support;

(f) contacting said solid support with an aqueous wash solution; and

(g) repeating steps (b)-(f) to determine target polynucleotide sequences.

39. The method of claim 38, wherein the first functional moiety of the first type of unlabeled nucleotide is bound to the first labeling reagent by either covalent bonding or noncovalent interaction via a cleavable linker.

40. The method of claim 38 or 39, wherein each of the four types of nucleotides in the aqueous incorporation mixture is unlabeled, the second type of unlabeled nucleotides comprises a second functional moiety, wherein the aqueous labeling mixture comprises a second labeling reagent, and the second labeling reagent comprises one or more second detectable labels and a second binding moiety that is capable of specific binding to the second functional moiety of the second type of unlabeled nucleotides.

41. The method of claim 40, wherein the second functional moiety of the second type of unlabeled nucleotides is bound to the second labeling reagent by either covalent bonding or noncovalent interaction via a cleavable linker.

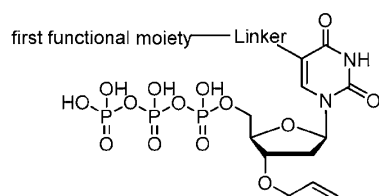
42. The method of claim 40 or 41, wherein the third type of unlabeled nucleotides comprises a third functional moiety, wherein the aqueous labeling mixture comprises a third labeling reagent, and the third labeling reagent comprises one or more third detectable labels and a third binding moiety that is capable of specific binding to the third functional moiety of the third type of unlabeled nucleotides.

43. The method of claim 40 or 41, wherein the third type of unlabeled nucleotide comprises a mixture of the third type of unlabeled nucleotides comprising the first functional moiety and the third type of unlabeled nucleotides comprising the second functional moiety, and

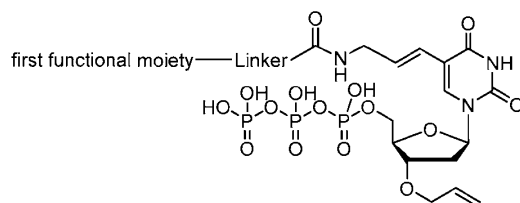
wherein both the first labeling reagent and the second labeling reagent are capable of specific binding to the third type of unlabeled nucleotides.

44. The method of claim 42 or 43, wherein the fourth type of unlabeled nucleotides is not capable of specific binding with any of the first, second, or third labeling reagent.

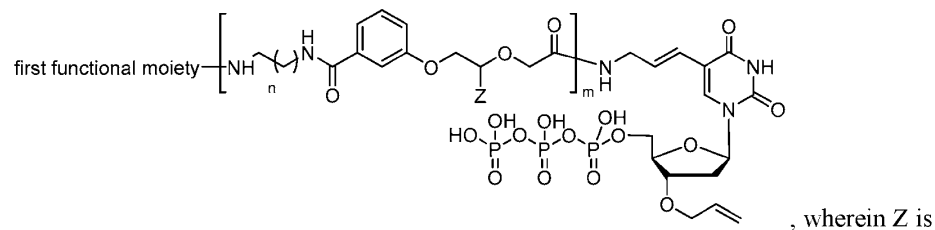
45. The method of any one of claims 38 to 44, wherein the T nucleotide has the structure:



46. The method of claim 45, wherein the T nucleotide has the structure:

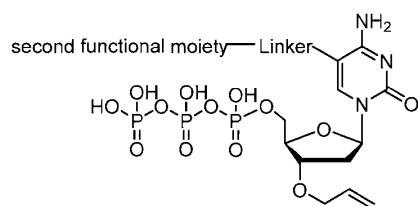


47. The method of claim 45 or 46, wherein the T nucleotide has the structure:

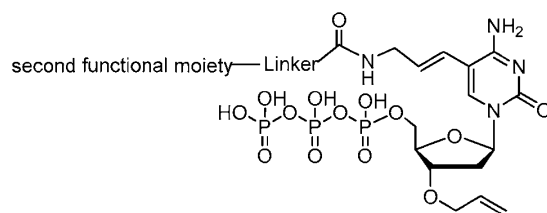


-O-CH₂-CH=CH₂, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

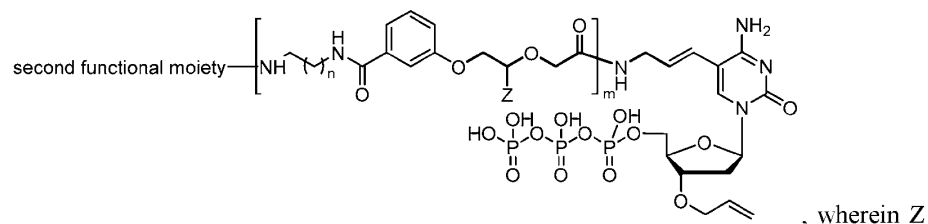
48. The method of any one of claims 38 to 47, wherein the C nucleotide has the structure:



49. The method of claim 48, wherein the C nucleotide has the structure:

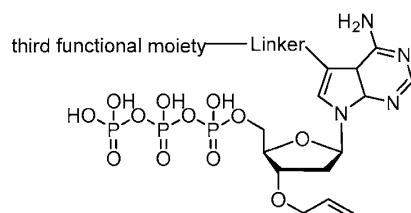


50. The method of claim 48 or 49, wherein the C nucleotide has the structure:

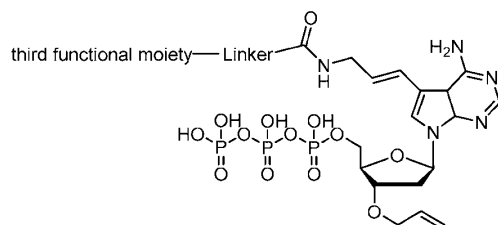


is -O-CH₂-CH=CH₂, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

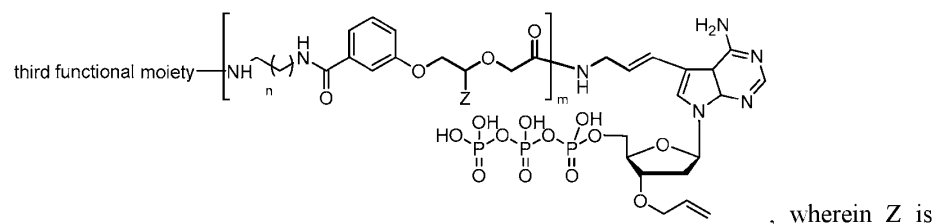
51. The method of any one of claims 38 to 50, wherein the A nucleotide has the structure:



52. The method of claim 51, wherein the A nucleotide has the structure:

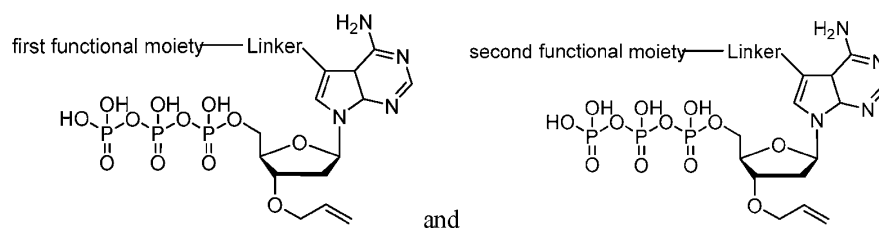


53. The method of claim 51 or 52, wherein the A nucleotide has the structure:

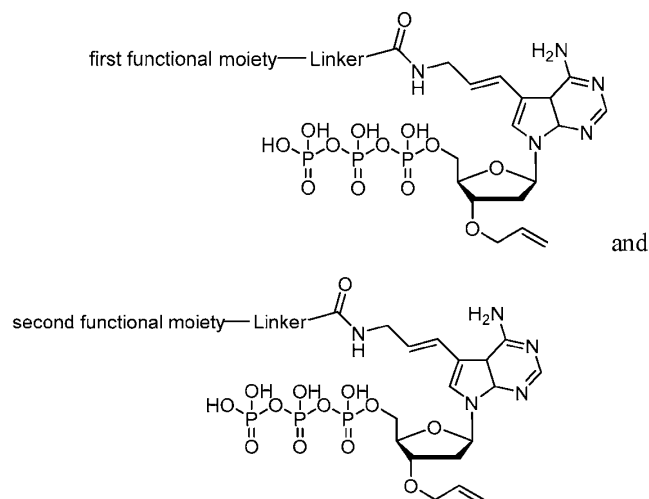


-O-CH₂-CH=CH₂, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

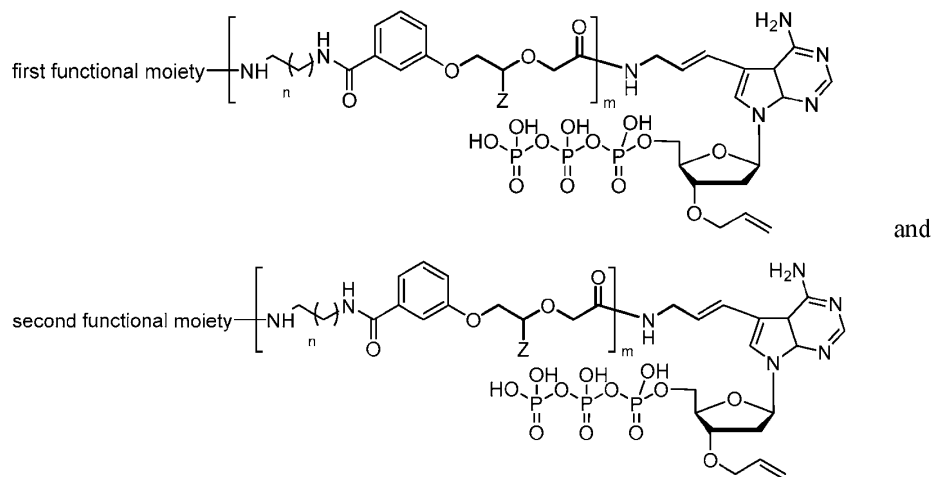
54. The method of any one of claims 38 to 50, wherein the A nucleotide has the structure:



55. The method of claim 54, wherein the A nucleotide has the structure:

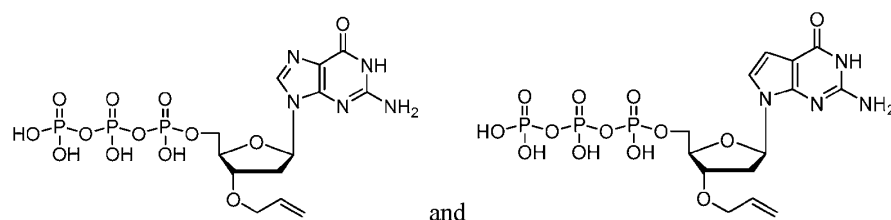


56. The method of claim 54 or 55, wherein the A nucleotide has the structure:



wherein Z is -O-CH₂-CH=CH₂, and each of m and n is independently an integer of 1, 2, 3, 4 or 5.

57. The method of any one of claims 38 to 56, wherein G nucleotide has a structure selected from the group consisting of:



58. The method of any one of claims 1 to 57, wherein said contacting the solid support with a deblocking solution is performed for 4-5 seconds.

59. The method of any one of claims 1 to 58, wherein said contacting the solid support with a deblocking solution is performed via continuous flow without pausing to incubate.

60. The method of any one of claims 1 to 59, wherein sequencing cycles are repeated at least about 20 times, 30 times, 50 times, 100 times, 150 times, 200 times, 250 times, 300 times, 350 times, 400 times, 450 times or 500 times.

61. The method of claim 60, wherein after about 50 repeated cycles the pre-phasing value is less than 0.18.

62. The method of claim 61, wherein after about 50 repeated sequencing cycles the phasing value is less than 0.18.

63. The method of claim 62, wherein after about 50 repeated sequencing cycles the pre-phasing value is less than 0.07.

64. The method of claim 60, wherein after about 100 repeated sequencing cycles the pre-phasing value is less than 0.10 and the phasing value is less than 0.10.

65. The method of claim 60, wherein after about 150 repeated sequencing cycles the pre-phasing value is less than 0.25 and the phasing value is less than 0.25.

66. The method of claim 60, wherein after about 150 repeated sequencing cycles the pre-phasing value is less than 0.10 and the phasing value is less than 0.10.

67. The method of any one of claims 1 to 66, wherein the deblocking solution further comprises one or more buffer reagents selected from the group consisting of a primary amine, a secondary amine, a tertiary amine, a carbonate salt, a phosphate salt, and a borate salt, and combinations thereof.

68. The method of claim 67, wherein the buffer reagents are selected from the group consisting of ethanolamine (EA), tris(hydroxymethyl)aminomethane (Tris), glycine, a carbonate salt, a phosphate salt, a borate salt, 2-dimethylaminoethanol (DMEA), 2-diethylaminoethanol (DEEA), N,N,N',N'-tetramethylethylenediamine (TEMED), N,N,N',N'-tetraethylethylenediamine (TEEDA), and (2-hydroxyethyl)piperidine, and combinations thereof.

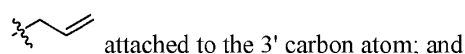
69. The method of any one of claims 1 to 68, wherein the DNA polymerase is an altered family B archaeal DNA polymerase comprising a 3-amino acid region that is functionally

equivalent or homologous to amino acids 408-410 in 9° N DNA polymerase, wherein the first amino acid of the 3-amino acid region is an amino acid selected from the group consisting of isoleucine (I), alanine (A), valine (V), and serine (S); the second amino acid of the 3-amino acid region is an amino acid selected from the group consisting of alanine (A) and glycine (G); and the third amino acid of the 3-amino acid region is an amino acid selected from the group consisting of alanine (A), isoleucine (I), valine (V), leucine (L), threonine (T), and proline (P).

70. A sequencing kit comprising:

(a) an incorporation mixture comprising DNA polymerase and nucleotides A, G, C, and T or U, wherein:

(i) the nucleotides comprise a 2' deoxyribose moiety with a 3' allyl group



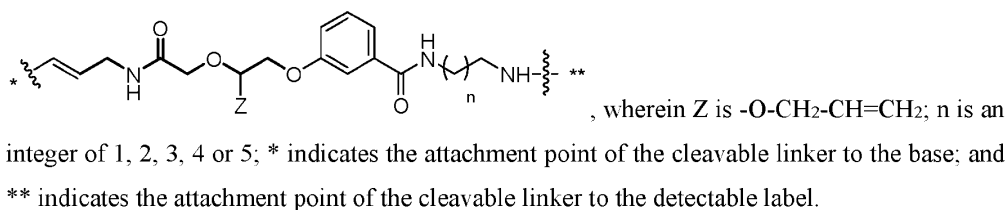
(ii) the DNA polymerase is an altered archaeal DNA polymerase;

(b) an aqueous deblocking solution comprising a palladium catalyst, tris(hydroxyalkyl)phosphine, and one or more buffer reagents that is suitable to chemically remove (i) 3' allyl groups from incorporated nucleotides to expose a 3'-OH group for further nucleotide incorporation on the solid support, and (ii) detectable labels attached via cleavable linkers; and

(c) an aqueous wash solution comprising a Pd(II) scavenger;

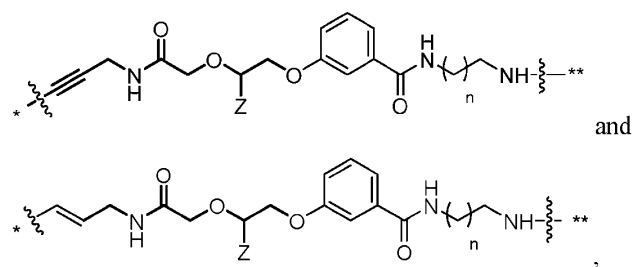
wherein said kit is configured for performing at least about 100 cycles of sequencing-by-synthesis.

71. The kit of claim 70, wherein at least one type of the nucleotides comprises a base that is attached to a detectable label via a cleavable linker, and the cleavable linker is



72. The kit of claim 71, wherein the nucleobase of T nucleotide is attached to the detectable label via the cleavable linker.

73. The kit of any one of claims 70 to 72, wherein each of at least three of the nucleotides independently comprises a base that is attached to a detectable label via a cleavable linker, and the cleavable linker is selected from the group consisting of:



wherein Z is $-O-CH_2-CH=CH_2$; n is an integer of 1, 2, 3, 4 or 5; * indicates the attachment point of the cleavable linker to the base; and ** indicates the attachment point of the cleavable linker to the detectable label.

74. The kit of claim 70, wherein two or more types of nucleotides A, G, C, and T or U are unlabeled, and wherein the first type of unlabeled nucleotides comprises a first functional moiety, and the kit further comprises a first labeling reagent, wherein the first labeling reagent comprises one or more first detectable labels and a first binding moiety that is capable of specific binding to the first functional moiety of the first type of unlabeled nucleotide.

75. The kit of claim 74, wherein each of the four types of nucleotides is unlabeled, and wherein the second type of unlabeled nucleotides comprises a second functional moiety, and the kit further comprises a second labeling reagent, wherein the second labeling reagent comprises one or more second detectable labels and a second binding moiety that is capable of specific binding to the second functional moiety of the second type of unlabeled nucleotide.

76. The kit of claim 75, wherein the third type of unlabeled nucleotides comprises a third functional moiety, and the kit further comprises a third labeling reagent, wherein the third labeling reagent comprises one or more third detectable labels and a third binding moiety that is capable of specific binding to the third functional moiety of the third type of unlabeled nucleotide.

77. The kit of claim 75, wherein the third type of unlabeled nucleotides comprises a mixture of the third type of unlabeled nucleotides comprising the first functional moiety and the third type of unlabeled nucleotides comprising the second functional moiety, and wherein both the first labeling reagent and the second labeling reagent are capable of specific binding to the third type of unlabeled nucleotides.

78. The kit of claim 76 or 77, wherein the fourth type of unlabeled nucleotides is not capable of specific binding with any of the first, second, or third labeling reagent.

79. The kit of any one of claims 70 to 78, wherein the incorporation mixture further comprises a Pd(0) scavenger.

80. The kit of any one of claims 70 to 79, wherein the tris(hydroxyalkyl)phosphine is tris(hydroxypropyl)phosphine (THPP).