

MULTIVALENT BINDING COMPOSITION FOR NUCLEIC ACID ANALYSIS

CROSS-REFERENCE

[0001] This application is a continuation-in-part of U.S. Patent Application No. 16/579,794, filed on September 23, 2019, and claims the benefit of U.S. Provisional Application No. 62/897,172 filed on September 6, 2019, and of U.S. Provisional Application No. 62/852,876 filed on May 24, 2019, each of which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present disclosure relates generally to multivalent binding compositions and their use in analyzing nucleic acid molecules. In particular, the inventive concept relates to a multivalent binding composition having multiple copies of a nucleotide attached to a particle or polymer core which effectively increases the local concentration of the nucleotide and enhances the binding signals. The multivalent binding composition can be applied, for example, in the field of sequencing and biosensor microarrays.

BACKGROUND

[0003] Nucleic acid sequencing can be used to obtain information in a wide variety of biomedical contexts, including diagnostics, prognostics, biotechnology, and forensic biology. Various sequencing methods have been developed including Maxam-Gilbert sequencing and chain-termination methods, or de novo sequencing methods including shotgun sequencing and bridge PCR, or next-generation methods including polony sequencing, 454 pyrosequencing, Illumina sequencing, SOLiD sequencing, Ion Torrent semiconductor sequencing, HeliScope single molecule sequencing, SMRT® sequencing, and others. Despite advances in DNA sequencing, many challenges to cost effective, high throughput sequencing remain unaddressed. The present disclosure provides novel solutions and approaches to addressing many of the shortcomings of existing technologies.

SUMMARY

[0003a] In one aspect, the present disclosure refers to a method of determining an identity of a nucleotide in a target nucleic acid sequence, comprising:

- (a) providing a composition comprising:
 - (i) two or more copies of said target nucleic acid sequence;
 - (ii) two or more primer nucleic acid molecules that are complementary to one or more regions of said target nucleic acid sequence; and
 - (iii) two or more polymerase molecules;

(b) contacting said composition with a polymer-nucleotide conjugate under conditions sufficient to allow a multivalent binding complex to be formed between said polymer-nucleotide conjugate and said two or more copies of said target nucleic acid sequence in said composition of (a), wherein the polymer-nucleotide conjugate comprises two or more nucleotide moieties; and

(c) detecting said multivalent binding complex, thereby determining the identity of said nucleotide in the target nucleic acid sequence.

[0004] Disclosed herein are methods of determining an identity of a nucleotide in a target nucleic acid sequence comprising: a. providing a composition comprising: i. two or more copies of said target nucleic acid sequence; ii. two or more primer nucleic acid molecules that are complementary to one or more regions of said target nucleic acid sequence; and iii. two or more polymerase molecules; b. contacting said composition with a polymer nucleotide conjugate under conditions sufficient to allow a multivalent binding complex to be formed between said polymer-nucleotide conjugate and said two or more copies of said target nucleic acid sequence in said composition of (a), wherein the polymer-nucleotide conjugate comprises two or more copies of a nucleotide moiety and

optionally one or more detectable labels; and c. detecting said multivalent binding complex, thereby determining the identity of said nucleotide in the target nucleic acid sequence. In some examples, the target nucleic acid sequence is DNA. In some examples, the detection of the multivalent binding complex is performed in the absence of unbound or solution-borne polymer nucleotide conjugates. In some examples, the target nucleic acid sequence has been replicated or amplified or has been produced by replication or amplification. In some examples, the one or more detectable labels are fluorescent labels. In some examples, detecting the multivalent complex comprises a fluorescence measurement. In some examples, the contacting comprises use of one type of polymer-nucleotide conjugate. In some examples, the contacting comprises use of two or more types of polymer-nucleotide conjugates. In some examples, each type of the two or more types of polymer-nucleotide conjugate comprises a different type of nucleotide moiety. In some examples, the contacting comprises use of three types of polymer-nucleotide conjugate and wherein each type of the three types of polymer-nucleotide conjugate comprises a different type of nucleotide moiety. In some examples, the polymer-nucleotide conjugate comprises a blocked nucleotide moiety. In some examples, the blocked nucleotide is a 3'-O-azidomethyl nucleotide, a 3'-O-methyl nucleotide, or a 3'-O-alkyl hydroxylamine nucleotide. In some examples, said contacting occurs in the presence of an ion that stabilizes said multivalent binding complex. In some examples, the contacting is done in the presence of strontium ions, magnesium ions, calcium ions, or any combination thereof. In some examples, the polymerase molecules are catalytically inactive. In some examples, the polymerase molecules have been rendered catalytically inactive by mutation or chemical modification. In some examples, the polymerase molecules have been rendered catalytically inactive by the absence of a necessary ion or cofactor. In some examples, the polymerase molecules are catalytically active. In some examples, the polymer-nucleotide conjugate does not comprise a blocked nucleotide moiety. In some examples, the multivalent binding complex has a persistence time of greater than 2 seconds. In some examples, the method can be carried out at a temperature within a range of 25°C to 62°C. In some examples, the polymer-nucleotide conjugate further comprises one or more fluorescent labels and the two or more copies of the target nucleic acid sequence are deposited on, attached to, or hybridized to a surface, wherein a fluorescence image of the multivalent binding complex on the surface has a contrast to noise ratio in the detecting step of greater than 20. In some examples, the composition of (a) is deposited on a surface using a buffer that incorporates a polar aprotic solvent. In some examples, the contacting is performed under a condition that stabilizes said multivalent binding complex when said nucleotide moiety is complementary to a next base of said target nucleic acid sequence and destabilizes said multivalent binding complex when said nucleotide moiety is not complementary to

said next base of said target nucleic acid sequence. In some examples, said polymer-nucleotide conjugate comprises a polymer having a plurality of branches and said two or more nucleotide moieties are attached to said branches. In some examples, said polymer has a star, comb, cross-linked, bottle brush, or dendrimer configuration. In some examples, said polymer-nucleotide conjugate comprises one or more binding groups selected from the group consisting of an avidin, a biotin, an affinity tag, and combinations thereof. In some examples, the method further comprises a dissociation step that destabilizes said multivalent binding complex formed between the composition of (a) and the polymer-nucleotide conjugate, said dissociation step enabling removal of said polymer-nucleotide conjugate. In some examples, the method further comprises an extension step to incorporate a nucleotide that is complementary to a next base of the target nucleic acid sequence into said two or more primer nucleic acid molecules. In some examples, the extension step occurs concurrently with or after said dissociation step.

[0005] Disclosed herein are methods of determining an identity of a nucleotide in a target nucleic acid sequence comprising: a. providing a composition comprising: i. two or more copies of said target nucleic acid sequence; ii. two or more primer nucleic acid molecules that are complementary to one or more regions of said target nucleic acid sequence; and iii. two or more polymerase molecules; b. contacting said composition with a polymer nucleotide conjugate under conditions sufficient to allow a multivalent complex to be formed between said polymer-nucleotide conjugate and said two or more copies of said target nucleic acid sequence in said composition of (a), wherein the polymer-nucleotide conjugate comprises two or more copies of a reversibly terminated nucleotide moiety and optionally one or more cleavable detectable labels; and c. detecting said multivalent complex, thereby determining the identity of said nucleotide in the target nucleic acid sequence. In some examples, the target nucleic acid sequence is DNA. In some examples, the method further comprises contacting the composition of (a) with reversibly terminated nucleotides or polymer-nucleotide conjugates comprising two or more copies of a reversibly terminated nucleotide following the detection of said multivalent binding complex. In some examples, the target nucleic acid sequence has been replicated or amplified or has been produced by replication or amplification. In some examples, the one or more detectable labels are fluorescent labels. In some examples, detecting the multivalent complex comprises a fluorescence measurement. In some examples, the contacting comprises use of one type of polymer-nucleotide conjugate. In some examples, the contacting comprises use of two or more types of polymer-nucleotide conjugates. In some examples, each type of the two or more types of polymer-nucleotide conjugate comprises a different type of nucleotide moiety. In some examples, the contacting comprises use of three types of polymer-nucleotide conjugate and wherein each type of the three

types of polymer-nucleotide conjugate comprises a different type of nucleotide moiety. In some examples, the polymer-nucleotide conjugate comprises a blocked nucleotide moiety. In some examples, the blocked nucleotide is a 3'-O-azidomethyl, 3'-O-methyl, or 3'-O-alkyl hydroxylamine. In some examples, said contacting occurs in the presence of an ion that stabilizes said multivalent binding complex. In some examples, the polymerase molecules are catalytically inactive. In some examples, the polymerase molecules have been rendered catalytically inactive by mutation or chemical modification. In some examples, the polymerase molecules are catalytically active. In some examples, the polymer-nucleotide conjugate does not comprise a blocked nucleotide moiety. In some examples, the method can be carried out at a temperature within a range of 25°C to 80°C. In some examples, the polymer-nucleotide conjugate further comprises one or more fluorescent labels and the two or more copies of the target nucleic acid sequence are deposited on, attached to, or hybridized to a surface, wherein a fluorescence image of the multivalent binding complex on the surface has a contrast to noise ratio in the detecting step of greater than 20.

[0006] Also disclosed herein are systems comprising: a) one or more computer processors individually or collectively programmed to implement a method comprising: i) contacting a substrate comprising multiple copies of a target nucleic acid sequence tethered to a surface of the substrate with a reagent comprising a polymerase and one or more primer nucleic acid sequences that are complementary to one or more regions of said target nucleic acid sequence to form a primed target nucleic acid sequence; ii) contacting the substrate surface with a reagent comprising a polymer nucleotide conjugate under conditions sufficient to allow a multivalent binding complex to be formed between said polymer-nucleotide conjugate and two or more copies of said primed target nucleic acid sequence, wherein the polymer-nucleotide conjugate comprises two or more copies of a known nucleotide moiety and a detectable label; iii) acquiring and processing an image of the substrate surface to detect said multivalent binding complex, thereby determining the identity of a nucleotide in the target nucleic acid sequence. In some examples, the system further comprises a fluidics module configured to deliver a series of reagents to the substrate surface in a specified sequence and for specified time intervals. In some examples, the system further comprises an imaging module configured to acquire images of the substrate surface. In some examples, (ii) and (iii) are repeated two or more times thereby determining the identity of a series of two or more nucleotides in the target nucleic acid sequence. In some examples, the series of steps further comprises a dissociation step that destabilizes said multivalent binding complex, said dissociation step enabling removal of said polymer-nucleotide conjugate. In some examples, the series of steps further comprises an extension step to incorporate a nucleotide that is complementary to a next base of the target nucleic

acid sequence into said two or more primer nucleic acid molecules. In some examples, the extension step occurs concurrently with or after said dissociation step. In some examples, the detectable label comprises a fluorophore and the images comprise fluorescence images. In some examples, the fluorescence images of the multivalent binding complex on the substrate surface has a contrast-to-noise ratio of greater than 20 when the fluorophore is cyanine dye 3 (Cy3) and the image is acquired using an inverted fluorescence microscope equipped with a 20X objective, NA = 0.75, dichroic mirror optimized for 532 nm light, a bandpass filter optimized for Cyanine dye-3 emission, and a camera, under non-signal saturating conditions while the surface is immersed in 25 mM ACES, pH 7.4 buffer. In some examples, the series of steps is completed in less than 60 minutes. In some examples, the series of steps is completed in less than 30 minutes. In some examples, the series of steps is completed in less than 10 minutes. In some examples, an accuracy of base-calling is characterized by a Q-score of greater than 25 for at least 80% of the nucleotide identities determined. In some examples, an accuracy of base-calling is characterized by a Q-score of greater than 30 for at least 80% of the nucleotide identities determined. In some examples, an accuracy of base-calling is characterized by a Q-score of greater than 40 for at least 80% of the nucleotide identities determined.

[0007] Disclosed herein are compositions comprising: a) a polymer core; and b) two or more nucleotide, nucleotide analog, nucleoside, or nucleoside analog moieties attached to the polymer core; wherein the length of the linker is dependent on the nucleotide, nucleotide analog, nucleoside, or nucleoside analog moiety that is attached to the polymer core. Also disclosed herein are compositions comprising: a) a mixture of polymer-nucleotide conjugates, wherein each polymer-nucleotide conjugate comprises: i) a polymer core; and ii) two or more nucleotide, nucleotide analog, nucleoside, or nucleoside analog moieties attached to the polymer core, wherein the length of the linker is dependent on the nucleotide, nucleotide analog, nucleoside, or nucleoside analog moiety that is attached to the polymer core; and wherein the mixture comprises polymer-nucleotide conjugates having at least two different types of attached nucleotide, nucleotide analog, nucleoside, or nucleoside analog moiety. In some examples, the polymer core comprises a polymer having a plurality of branches and the two or more nucleotide, nucleotide analog, nucleoside, or nucleoside analog moieties are attached to said branches. In some examples, polymer has a star, comb, cross-linked, bottle brush, or dendrimer configuration. In some examples, the polymer-nucleotide conjugate comprises one or more binding groups selected from the group consisting of an avidin, a biotin, an affinity tag, and combinations thereof. In some examples, the polymer core comprises a branched polyethylene glycol (PEG) molecule. In some examples, the polymer-nucleotide conjugate comprises a blocked nucleotide moiety. In some examples, the blocked

nucleotide is a 3'-O-azidomethyl nucleotide, a 3'-O-methyl nucleotide, or a 3'-O-alkyl hydroxylamine nucleotide. In some examples, the polymer-nucleotide conjugate further comprises one or more fluorescent labels.

[0008] In some examples the present disclosure provides methods of determining the identity of a nucleotide in a target nucleic acid comprising the steps, without regard to any particular order of operations, 1) providing a composition comprising: a target nucleic acid comprising two or more repeats of an identical sequence; two or more primer nucleic acids complementary to one or more regions of said target nucleic acid; and two or more polymerase molecules; 2) contacting said composition with a multivalent binding or incorporation composition comprising a polymer-nucleotide conjugate under conditions sufficient to allow a binding or incorporated complex to be formed between said polymer-nucleotide conjugate and the composition of step (a), wherein the polymer-nucleotide conjugate comprises two or more copies of a nucleotide and optionally one or more detectable labels; and 3) detecting said binding or incorporated complex, thereby establishing the identity of said nucleotide in the target nucleic acid polymer. In some further examples, the present disclosure provides said method, wherein the target nucleic acid is DNA, and/or wherein the target nucleic acid has been replicated, such as by any commonly practiced method of DNA replication or amplification, such as rolling circle amplification, bridge amplification, helicase dependent amplification, isothermal bridge amplification, rolling circle multiple displacement amplification (RCA/MDA) and/or recombinase based methods of replication or amplification. . In some further examples, the present disclosure provides said method, wherein the detectable label is a fluorescent label and/or wherein detecting the complex comprises a fluorescence measurement. In some further examples, the present disclosure provides said method wherein the multivalent binding composition comprises one type of polymer-nucleotide conjugate, wherein the multivalent binding composition comprises two or more types of polymer-nucleotide conjugates, and/or wherein each type of the two or more types of polymer-nucleotide conjugates comprises a different type of nucleotide. In some examples, the present disclosure provides said method wherein the binding complex or incorporated complex further comprises a blocked nucleotide, especially wherein the blocked nucleotide is a 3'-O-azidomethyl nucleotide, a 3'-O-alkyl hydroxylamino nucleotide, or a 3'-O-methyl nucleotide. In some further examples, the present disclosure provides said method wherein the contacting is done in the presence of strontium ions, barium, magnesium ions, and/or calcium ions. In some examples, the present disclosure provides said method wherein the polymerase molecule is catalytically inactive, such as where the polymerase molecule been rendered catalytically inactive by mutation, by chemical modification, or by the absence of a necessary ion or cofactor. In some examples, the present disclosure also provides said method wherein the

polymerase molecule is catalytically active, and/or wherein the binding complex does not comprise a blocked nucleotide. In some examples, the present disclosure provides said method wherein the binding complex has a persistence time of greater than 2 seconds and/or wherein the method is or may be carried out at a temperature of at or above 15 °C, at or above 20 °C, at or above 25 °C, at or above 35 °C, at or above 37 °C, at or above 42 °C at or above 55 °C at or above 60 °C, or at or above 72 °C, or within a range defined by any of the foregoing. In some examples, the present disclosure provides said method wherein the binding complex is deposited on, attached to, or hybridized to, a surface showing a contrast to noise ratio in the detecting step of greater than 20. In some examples, the present disclosure provides said method wherein the composition is deposited under buffer conditions incorporating a polar aprotic solvent. In some examples, the present disclosure provides said method wherein the contacting is performed under a condition that stabilizes said binding complex when said nucleotide is complementary to a next base of said target nucleic acid and destabilizes said binding complex when said nucleotide is not complementary to said next base of said target nucleic acid. In some examples, the present disclosure provides said method wherein said polymer-nucleotide conjugate comprises a polymer having a plurality of branches and said plurality of copies of said first nucleotide are attached to said branches, especially wherein said first polymer has a star, comb, cross-linked, bottle brush, or dendrimer configuration. In some examples, the present disclosure provides said method wherein said polymer-nucleotide conjugate comprises one or more binding groups selected from the group consisting of avidin, biotin, affinity tag, and combinations thereof. In some examples, the present disclosure provides said method further comprising a dissociation step that destabilizes said binding complex formed between the composition of (a) and the polymer-nucleotide conjugate to remove said polymer-nucleotide conjugate. In some examples, the present disclosure provides said method further comprising an extension step to incorporate into said primer nucleic acid a nucleotide that is complementary to said next base of the target nucleic acid, and optionally wherein the extension step occurs currently as or after said dissociation step.

[0009] In some examples, the present disclosure provides a composition comprising a branched polymer having two or more branches and two or more copies of a nucleotide, wherein said nucleotide is attached to a first plurality of said branches or arms, and optionally, wherein one or more interaction moieties are attached to a second plurality of said branches or arms. In some examples, said composition may further comprise one or more labels on the polymer. In some examples, the present disclosure provides said composition wherein the nucleoside has a surface density of at least 4 nucleotides per polymer. In some examples, the present disclosure provides said composition comprising or incorporating a nucleotide or nucleotide analog that is modified so

as to prevent its incorporation into an extending nucleic acid chain during a polymerase reaction. In some examples, said composition may comprise or incorporate a nucleotide or nucleotide analog that is reversibly modified so as to prevent its incorporation into an extending nucleic acid chain during a polymerase reaction. In some examples, the present disclosure provides said composition wherein one or more labels comprise a fluorescent label, a FRET donor, and/or a FRET acceptor. In some examples, said composition may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more branches or arms, or 2, 4, 8, 16, 32, 64, or more, branches or arms. In some examples, the branches or arms may radiate from a central moiety. In some examples, said composition may comprise one or more interaction moieties, which interaction moieties may comprise avidin or streptavidin; a biotin moiety; an affinity tag; an enzyme, antibody, minibody, receptor, or other protein; a non-protein tag; a metal affinity tag, or any combination thereof. In some examples, the present disclosure provides said composition wherein the polymer comprises polyethylene glycol, polypropylene glycol, polyvinyl acetate, polylactic acid, or polyglycolic acid. In some examples, the present disclosure provides said composition wherein the nucleotide or nucleotide analog is attached to the branch or arm through a linker; and especially wherein the linker comprises PEG, and wherein the PEG linker moiety has an average molecular weight of about 1K Da, about 2K Da, about 3K Da, about 4K Da, about 5K Da, about 10K Da, about 15K Da, about 20K Da, about 50K Da, about 100K Da, about 150K Da, or about 200K Da, or greater than about 200K Da. In some examples, the present disclosure provides said composition wherein the linker comprises PEG, and wherein the PEG linker moiety has an average molecular weight of between about 5K Da and about 20K Da. In some examples, the present disclosure provides said composition wherein at least one nucleotide or nucleotide analog comprises a deoxyribonucleotide, a ribonucleotide, a deoxyribonucleoside, or a ribonucleoside; and/or wherein the nucleotide or nucleotide analog is conjugated to the linker through the 5' end of the nucleotide or nucleotide analog. In some examples, the present disclosure provides said composition wherein one of the nucleotides or nucleotide analogs comprises deoxyadenosine, deoxyguanosine, thymidine, deoxyuridine, deoxycytidine, adenosine, guanosine, 5-methyl-uridine, and/or cytidine; and wherein the length of the linker is between 1 nm and 1,000 nm. In some examples, the present disclosure provides said composition wherein at least one nucleotide or nucleotide analog is a nucleotide that has been modified to inhibit elongation during a polymerase reaction or a sequencing reaction, such as wherein the at least one nucleotide or nucleotide analog is a nucleotide that lacks a 3' hydroxyl group; a nucleotide that has been modified to contain a blocking group at the 3' position; and/or a nucleotide that has been modified with a 3'-O-azido group, a 3'-O-azidomethyl group, , a 3'-O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-O-malonyl group, or a 3'-O-benzyl group.

In some examples, the present disclosure provides said composition wherein at least one nucleotide or nucleotide analog is a nucleotide that has not been modified at the 3' position.

[0010] In some examples, the present disclosure provides a method of determining the sequence of a nucleic acid molecule comprising the steps, without regard to any particular order, of 1) providing a nucleic acid molecule comprising a template strand and a complementary strand that is at least partially complementary to the template strand; 2) contacting the nucleic acid molecule with the one or more nucleic acid binding composition according to any of the examples disclosed herein; 3) detecting binding of the nucleic acid binding composition to the nucleic acid molecule, and 4) determining an identity of a terminal nucleotide to be incorporated into said complementary strand of said nucleic acid molecule. In some examples, the present disclosure provides a method of determining the sequence of a nucleic acid molecule comprising the steps, without regard to any particular order, of 1) providing a nucleic acid molecule comprising a template strand and a complementary strand that is at least partially complementary to the template strand; 2) contacting the nucleic acid molecule with the one or more nucleic acid binding composition according to any of the examples disclosed herein; 3) detecting partial or complete incorporation of the nucleic acid binding composition to the nucleic acid molecule, and 4) determining an identity of a terminal nucleotide to be incorporated into said complementary strand of said nucleic acid molecule from the partial or complete incorporation of the examples described herein. In some examples, the present disclosure provides said method, further comprising incorporating said terminal nucleotide into said complementary strand, and repeating said contacting, detecting, and incorporating steps for one or more additional iterations, thereby determining the sequence of said template strand of said nucleic acid molecule. In some examples, the present disclosure provides said method, wherein said nucleic acid molecule is tethered to a solid support; and especially wherein the solid support comprises a glass or polymer substrate, at least one hydrophilic polymer coating layer, and a plurality of oligonucleotide molecules attached to at least one hydrophilic polymer coating layer. In some examples, the present disclosure provides said method, further comprising examples wherein at least one hydrophilic polymer coating layer comprises PEG; and/or wherein at least one hydrophilic polymer layer comprises a branched hydrophilic polymer having at least 8 branches. In some examples, the present disclosure provides said method, wherein the plurality of oligonucleotide molecules is present at a surface density of at least 500 molecules/mm², at least 1,000 molecules/mm², at least 5,000 molecules/mm², at least 10,000 molecules/mm², at least 20,000 molecules/mm², at least 50,000 molecules/mm², at least 100,000 molecules/mm², or at least 500,000 molecules/mm². In some examples, the present disclosure provides said method, wherein said nucleic acid molecule has been clonally-amplified on a solid support. In some examples, the

present disclosure provides said method, wherein the clonal amplification comprises the use of a polymerase chain reaction (PCR), multiple displacement amplification (MDA), transcription-mediated amplification (TMA), nucleic acid sequence-based amplification (NASBA), strand displacement amplification (SDA), real-time SDA, bridge amplification, isothermal bridge amplification, rolling circle amplification (RCA), circle-to-circle amplification, helicase-dependent amplification, recombinase-dependent amplification, single-stranded binding (SSB) protein-dependent amplification, or any combination thereof. In some examples, the present disclosure provides said method, wherein the one or more nucleic acid binding compositions are labeled with fluorophores and the detecting step comprises use of fluorescence imaging; and especially wherein the fluorescence imaging comprises dual wavelength excitation/four wavelength emission fluorescence imaging. In some examples, the present disclosure provides said method, wherein four different nucleic acid binding compositions, each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein the four different nucleic acid binding compositions are labeled with separate respective fluorophores, and wherein the detecting step comprises simultaneous excitation at a wavelength sufficient to excite all four fluorophores and imaging of fluorescence emission at wavelengths sufficient to detect each respective fluorophore. In some examples, the present disclosure provides said method, wherein four different nucleic acid binding compositions, each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein the four different nucleic acid binding compositions are labeled with cyanine dye 3 (Cy3), cyanine dye 3.5 (Cy3.5), cyanine dye 5 (Cy5), and cyanine dye 5.5. (Cy5.5) respectively, and wherein the detecting step comprises simultaneous excitation at any two of 532 nm, 568 nm and 633 nm, and imaging of fluorescence emission at about 570 nm, 592 nm, 670 nm, and 702 nm respectively; and/or wherein the fluorescence imaging comprises dual wavelength excitation / dual wavelength emission fluorescence imaging. In some examples, the present disclosure provides said method, wherein four different nucleic acid binding compositions, each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein one, two, three, or four different nucleic acid binding compositions are respectively labeled, each with a with distinct fluorophore or set of fluorophores, and wherein the detecting step comprises simultaneous excitation at a wavelength sufficient to excite one, two, three, or four fluorophores or sets of fluorophores, and imaging of fluorescence emission at wavelengths sufficient to detect each respective fluorophore. In some examples, the present disclosure provides said method, wherein three different nucleic acid binding or incorporation compositions, each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein one, two, or three different

nucleic acid binding or incorporation compositions are respectively labeled, each with a with distinct fluorophore or set of fluorophores, and wherein the detecting step comprises simultaneous excitation at a wavelength sufficient to excite one, two, or three, fluorophores or sets of fluorophores, and imaging of fluorescence emission at wavelengths sufficient to detect each respective fluorophore, and wherein detection of the fourth nucleotide is determined or determinable with reference to the location of “dark” or unlabeled spots or target nucleotides. In some examples, the present disclosure provides said method, wherein the multivalent binding or incorporation composition may comprise three types of polymer-nucleotide conjugates and wherein each type of the three types of polymer-nucleotide conjugates comprises a different type of nucleotide. In some examples, the present disclosure provides said method, wherein the detection of the binding or incorporation complex is performed in the absence of unbound or solution-borne polymer nucleotide conjugates.

[0011] In some examples, the present disclosure provides said method, wherein four different nucleic acid binding compositions, or three different nucleic acid binding or incorporation compositions, each comprising a different nucleotide or nucleotide analog, are used to determine the identity of the terminal nucleotide, wherein one of the four or three different nucleic acid binding or incorporation compositions is labeled with a first fluorophore, one is labeled with a second fluorophore, one is labeled with both the first and second fluorophore, and one is not labeled or is absent, and wherein the detecting step comprises simultaneous excitation at a first excitation wavelength and a second excitation wavelength and images are acquired at a first fluorescence emission wavelength and a second fluorescence emission wavelength. In some examples, the present disclosure provides said method, wherein the first fluorophore is Cy3, the second fluorophore is Cy5, the first excitation wavelength is 532 nm or 568 nm, the second excitation wavelength is 633 nm, the first fluorescence emission wavelength is about 570 nm, and the second fluorescence emission wavelength is about 670 nm. In some examples, the present disclosure provides said method, wherein the detection label can comprise one or more portions of a fluorescence resonance energy transfer (FRET) pair, such that multiple classifications can be performed under a single excitation and imaging step. In some examples, the present disclosure provides said method, wherein a sequencing reaction cycle comprising the contacting, detecting, and incorporating/extending steps is performed in less than 30 minutes in less than 20 minutes, or in less than 10 minutes. In some examples, the present disclosure provides said method, wherein an average Q-score for base calling accuracy over a sequencing run is greater than or equal to 30, and/or greater than or equal to 40. In some examples, the present disclosure provides said method, wherein at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the terminal nucleotides identified have a Q-score of greater than 30 and/or greater than or equal to 40. In some

examples, the present disclosure provides said method, herein at least 95% of the terminal nucleotides identified have a Q-score of greater than 30.

[0012] In some examples, the present disclosure provides a reagent comprising one or more nucleic acid binding compositions as disclosed herein and a buffer. For example, in some examples, the present disclosure provides a reagent, wherein said reagent comprises 1, 2, 3, 4, or more nucleic acid binding or incorporation compositions, wherein each nucleic acid binding or incorporation composition comprises a single type of nucleotide. In some examples, a reagent of the present disclosure comprises 1, 2, 3, 4, or more nucleic acid binding or incorporation compositions, wherein each nucleic acid binding or incorporation composition comprises a single type of nucleotide or nucleotide analog, and wherein said nucleotide or nucleotide analog may respectively correspond to one or more from the group consisting of adenosine triphosphate (ATP), adenosine diphosphate (ADP), adenosine monophosphate (AMP), deoxyadenosine triphosphate (dATP), deoxyadenosine diphosphate (dADP), and deoxyadenosine monophosphate (dAMP); one or more from the group consisting of thymidine triphosphate (TTP), thymidine diphosphate (TDP), thymidine monophosphate (TMP), deoxythymidine triphosphate (dTTP), deoxythymidine diphosphate (dTDP), deoxythymidine monophosphate (dTMP), uridine triphosphate (UTP), uridine diphosphate (UDP), uridine monophosphate (UMP), deoxyuridine triphosphate (dUTP), deoxyuridine diphosphate (dUDP), and deoxyuridine monophosphate (dUMP); one or more from the group consisting of cytidine triphosphate (CTP), cytidine diphosphate (CDP), cytidine monophosphate (CMP), deoxycytidine triphosphate (dCTP), deoxycytidine diphosphate (dCDP), and deoxycytidine monophosphate (dCMP); and one or more from the group consisting of guanosine triphosphate (GTP), guanosine diphosphate (GDP), guanosine monophosphate (GMP), deoxyguanosine triphosphate (dGTP), deoxyguanosine diphosphate (dGDP), and deoxyguanosine monophosphate (dGMP). In some other examples or some further examples, the present disclosure provides a reagent comprising or further comprising 1, 2, 3, 4, or more nucleic acid binding or incorporation compositions, wherein each nucleic acid binding or incorporation composition comprises a single type of nucleotide or nucleotide analog, and wherein said nucleotide or nucleotide analog may respectively correspond to one or more from the group consisting of ATP, ADP, AMP, dATP, dADP, dAMP, TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, dUMP, CTP, CDP, CMP, dCTP, dCDP, dCMP, GTP, GDP, GMP, dGTP, dGDP, and dGMP.

[0013] Disclosed herein are kits comprising the nucleic acid binding or incorporation composition of any of the examples disclosed herein and/or a reagent of any of the examples disclosed herein, and/or one or more buffers; and instructions for the use thereof.

[0014] Disclosed herein are systems for performing the method of any example disclosed herein, comprising a nucleic acid binding or incorporation composition of any of the examples disclosed herein, and/or a reagent of any of the examples disclosed herein. In some examples, a system is configured to iteratively perform the sequential contacting of tethered, primed nucleic acid molecules with said nucleic acid binding or incorporation compositions and/or said reagents; and for the detection of binding or incorporation of the disclosed nucleic acid binding or incorporation compositions to the one or more primed nucleic acid molecules.

[0015] In some examples, the present disclosure provides a composition comprising a particle (*e.g.*, a nanoparticle or polymer core), said particle comprising a plurality of enzyme or protein binding or incorporation substrates, wherein the enzyme or protein binding or incorporation substrates bind with one or more enzymes or proteins to form one or more binding or incorporation complexes (*e.g.*, a multivalent binding or incorporation complex), and wherein said binding or incorporation may be monitored or identified by observation of the location, presence, or persistence of the one or more binding or incorporation complexes. In some examples, said particle may comprise a polymer, branched polymer, dendrimer, liposome, micelle, nanoparticle, or quantum dot. In some examples, said substrate may comprise a nucleotide, a nucleoside, a nucleotide analog, or a nucleoside analog. In some examples, the enzyme or protein binding or incorporation substrate may comprise an agent that can bind with a polymerase. In some examples, the enzyme or protein may comprise a polymerase. In some examples, said observation of the location, presence, or persistence of one or more binding or incorporation complexes may comprise fluorescence detection. In some examples, the present disclosure provides a composition comprising multiple distinct particles as disclosed herein, wherein each particle comprises a single type of nucleoside or nucleoside analog, and wherein each nucleoside or nucleoside analog is associated with a fluorescent label of a detectably different emission or excitation wavelength. In some examples, the present disclosure provides said composition further comprising one or more labels, *e.g.*, fluorescence labels, on the particle. In some examples, the present disclosure provides said composition wherein the composition comprises at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or more than 20 tethered nucleotides, nucleotide analogs, nucleosides, or nucleoside analogs tethered to the particle. In some examples, the present disclosure provides said composition wherein the nucleoside or nucleoside analog is present at a surface density of between 0.001 and 1,000,000 per μm^2 , between 0.01 and 1,000,000 per μm^2 , between 0.1 and 1,000,000 per μm^2 , between 1 and 1,000,000 per μm^2 , between 10 and 1,000,000 per μm^2 , between 100 and 1,000,000 per μm^2 , between 1,000 and 1,000,000 per μm^2 , between 1,000 and 100,000 per μm^2 , between 10,000 and 100,000 per μm^2 , or between 50,000 and 100,000 per μm^2 , or within a range defined by any two of the foregoing

values. In some examples, the present disclosure provides said composition wherein the nucleoside or nucleoside analog is present within a nucleotide or nucleotide analog. In some examples, the present disclosure provides said composition wherein the composition comprises or incorporates a nucleotide or nucleotide analog that is modified so as to prevent its incorporation into an extending nucleic acid chain during a polymerase reaction. In some examples, the present disclosure provides said composition wherein the composition comprises or incorporates a nucleotide or nucleotide analog that is reversibly modified so as to prevent its incorporation into an extending nucleic acid chain during a polymerase reaction. In some examples, the present disclosure provides said composition wherein one or more labels comprise a fluorescent label, a FRET donor, and/or a FRET acceptor. In some examples, the present disclosure provides said composition wherein the substrate (*e.g.*, nucleotide, nucleotide analog, nucleoside, or nucleoside analog) is attached to the particle through a linker. In some examples, the present disclosure provides said composition wherein at least one nucleotide or nucleotide analog is a nucleotide that has been modified to inhibit elongation during a polymerase reaction or a sequencing reaction, such as, for example, a nucleotide that lacks a 3' hydroxyl group; a nucleotide that has been modified to contain a blocking group at the 3' position; a nucleotide that has been modified with a 3'-O-azido group, a 3'-O-azidomethyl group, a 3'-O-alkyl hydroxylamino group, a 3'-phosphorothioate group, a 3'-O-malonyl group, or a 3'-O-benzyl group; and/or a nucleotide that has not been modified at the 3' position.

[0016] In some examples, the present disclosure provides a method of determining the sequence of a nucleic acid molecule comprising the steps, without regard to order, of 1) providing a nucleic acid molecule comprising a template strand and a complementary strand that is at least partially complementary to the template strand; 2) contacting the nucleic acid molecule with the one or more nucleic acid binding or incorporation composition according to any of the examples disclosed herein; 3) detecting binding or incorporation of the nucleic acid binding or incorporation composition to the nucleic acid molecule, and 4) determining an identity of a terminal nucleotide to be incorporated into said complementary strand of said nucleic acid molecule. In some examples, said method may further comprise incorporating said terminal nucleotide into said complementary strand, and repeating said contacting, detecting, and incorporating steps for one or more additional iterations, thereby determining the sequence of said template strand of said nucleic acid molecule. In some examples, the present disclosure provides said method wherein said nucleic acid molecule has been clonally-amplified on a solid support. In some examples, the present disclosure provides said method wherein the clonal amplification comprises the use of a polymerase chain reaction (PCR), multiple displacement amplification (MDA), transcription-mediated amplification

(TMA), nucleic acid sequence-based amplification (NASBA), strand displacement amplification (SDA), real-time SDA, bridge amplification, isothermal bridge amplification, rolling circle amplification, circle-to-circle amplification, helicase-dependent amplification, recombinase-dependent amplification, single-stranded binding (SSB) protein-dependent amplification, or any combination thereof. In some examples, the present disclosure provides said method wherein a sequencing reaction cycle comprising the contacting, detecting, and incorporating steps is performed in less than 30 minutes, less than 20 minutes, or in less than 10 minutes. In some examples, the present disclosure provides said method wherein an average Q-score for base calling accuracy over a sequencing run is greater than or equal to 30, or greater than or equal to 40. In some examples, the present disclosure provides said method wherein at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the terminal nucleotides identified have a Q-score of greater than 30; or greater than 40. In some examples, the present disclosure provides said method wherein at least 95% of the terminal nucleotides identified have a Q-score of greater than 30.

[0017] In some examples, the present disclosure provides a reagent comprising one or more nucleic acid binding or incorporation compositions as disclosed herein, and a buffer. In some examples, the present disclosure provides said reagent, wherein said reagent comprises 1, 2, 3, 4, or more nucleic acid binding or incorporation compositions, wherein each nucleic acid binding or incorporation composition comprises a single type of nucleotide or nucleotide analog, and wherein said nucleotide or nucleotide analog comprises a nucleotide, nucleotide analog, nucleoside, or nucleoside analog. In some examples, the present disclosure provides said method wherein said reagent comprises 1, 2, 3, 4, or more nucleic acid binding or incorporation compositions, wherein each nucleic acid binding or incorporation composition comprises a single type of nucleotide or nucleotide analog, and wherein said nucleotide or nucleotide analog may respectively correspond to one or more from the group consisting of ATP, ADP, AMP, dATP, dADP, and dAMP; one or more from the group consisting of TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, and dUMP; one or more from the group consisting of CTP, CDP, CMP, dCTP, dCDP, and dCMP; and one or more from the group consisting of GTP, GDP, GMP, dGTP, dGDP, and dGMP. In some examples, the present disclosure provides said method wherein said reagent comprises 1, 2, 3, 4, or more nucleic acid binding or incorporation compositions, wherein each nucleic acid binding or incorporation composition comprises a single type of nucleotide or nucleotide analog, and wherein said nucleotide or nucleotide analog may respectively correspond to one or more from the group consisting of ATP, ADP, AMP, dATP, dADP, dAMP, TTP, TDP, TMP, dTTP, dTDP, dTMP, UTP, UDP, UMP, dUTP, dUDP, dUMP, CTP, CDP, CMP, dCTP, dCDP, dCMP, GTP, GDP, GMP, dGTP, dGDP, and dGMP.

[0018] In some examples, the present disclosure provides a kit comprising any of the compositions disclosed herein; and/or any of the reagents disclosed herein; one or more buffers; and instructions for the use thereof.

[0019] In some examples, the present disclosure provides a system for performing any of the methods disclosed herein; wherein said methods may comprise use of any of the compositions as disclosed herein; and/or any of the reagents as disclosed herein; one or more buffers, and one or more nucleic acid molecules optionally tethered or attached to a solid support, wherein said system is configured to iteratively perform for the sequential contacting of said nucleic acid molecules with said composition and/or said reagent; and for the detection of binding or incorporation of the nucleic acid binding or incorporation compositions to the one or more nucleic acid molecules.

[0020] In some examples, the present disclosure provides a composition as disclosed herein for use in increasing the contrast to noise ratio (CNR) of a labeled nucleic acid complex bound to or associated with a surface.

[0021] In some examples, the present disclosure provides a composition as disclosed herein for use in establishing or maintaining control over the persistence time of a signal from a labeled nucleic acid complex bound to or associated with a surface.

[0022] In some examples, the present disclosure provides a composition as disclosed herein for use in establishing or maintaining control over the persistence time of a fluorescence, luminescence, electrical, electrochemical, colorimetric, radioactive, magnetic, or electromagnetic signal from a labeled nucleic acid complex bound to or associated with a surface.

[0023] In some examples, the present disclosure provides a composition as disclosed herein for use in increasing the specificity, accuracy, or read length of a nucleic acid sequencing and/or genotyping application.

[0024] In some examples, the present disclosure provides a composition as disclosed herein for use in increasing the specificity, accuracy, or read length in a sequencing by binding or incorporation, sequencing by synthesis, single molecule sequencing, or ensemble sequencing method.

[0025] In some examples, the present disclosure provides a reagent as disclosed herein for use in increasing the contrast to noise ratio (CNR) of a labeled nucleic acid complex bound to or associated with a surface.

[0026] In some examples, the present disclosure provides a reagent as disclosed herein for use in establishing or maintaining control over the persistence time of a signal from a labeled nucleic acid complex bound to or associated with a surface.

[0027] In some examples, the present disclosure provides a reagent as disclosed herein for use in establishing or maintaining control over the persistence time of a fluorescence, luminescence,

electrical, electrochemical, colorimetric, radioactive, magnetic, or electromagnetic signal from a labeled nucleic acid complex bound to or associated with a surface.

[0028] In some examples, the present disclosure provides a reagent as disclosed herein for use in increasing the specificity, accuracy, or read length of a nucleic acid sequencing and/or genotyping application.

[0029] In some examples, the present disclosure provides a reagent as disclosed herein for use in increasing the specificity, accuracy, or read length in a sequencing by binding or incorporation, sequencing by synthesis, single molecule sequencing, or ensemble sequencing method.

INCORPORATION BY REFERENCE

[0030] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference in its entirety. In the event of a conflict between a term herein and a term in an incorporated reference, the term herein controls.

BRIEF DESCRIPTION OF THE DRAWINGS

[0031] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0032] The novel features of the inventive concepts disclosed herein are set forth with particularity in the appended claims. A better understanding of the features and advantages of the disclosed compositions, methods, and systems will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the inventive concepts are utilized, and the accompanying drawings of which:

[0033] **FIGS. 1A-1H** illustrate the steps utilizing a non-limiting examples of multivalent binding composition for sequencing a target nucleic acid: **FIG. 1A** illustrates a non-limiting example of attaching target nucleic acid to a surface; **FIG. 1B** illustrates clonally the target nucleic acid to form clusters of amplified target nucleic acid molecules; **FIG. 1C** illustrates a non-limiting example of priming the target nucleic acid to produce a primed target nucleic acid; **FIG. 1D** illustrates a non-limiting example of contacting the primed target nucleic acid to the multivalent binding composition and polymerase to form a binding complex; **FIG. 1E** illustrates a non-limiting example of the images of the binding complex captured on the surface; **FIG. 1F** illustrates a non-limiting example of extending the primer strand by one nucleotide; **FIG. 1G** illustrates a non-limiting example of another cycle of contacting the primed target nucleic acid to the multivalent binding composition and

CLAIMS

WHAT IS CLAIMED IS:

1. A method of determining an identity of a nucleotide in a target nucleic acid sequence, comprising:
 - (a) providing a composition comprising:
 - (i) two or more copies of said target nucleic acid sequence;
 - (ii) two or more primer nucleic acid molecules that are complementary to one or more regions of said target nucleic acid sequence; and
 - (iii) two or more polymerase molecules;
 - (b) contacting said composition with a polymer-nucleotide conjugate under conditions sufficient to allow a multivalent binding complex to be formed between said polymer-nucleotide conjugate and said two or more copies of said target nucleic acid sequence in said composition of (a), wherein the polymer-nucleotide conjugate comprises two or more nucleotide moieties; and
 - (c) detecting said multivalent binding complex, thereby determining the identity of said nucleotide in the target nucleic acid sequence.
2. The method of claim 1, wherein the target nucleic acid sequence is DNA.
3. The method of claim 1, wherein the detection of the binding complex is performed in the absence of unbound or solution-borne polymer nucleotide conjugates.
4. The method of claim 1, wherein the target nucleic acid sequence has been replicated or amplified or has been produced by replication or amplification.
5. The method of claim 1, wherein detecting the multivalent binding complex comprises a fluorescence measurement.
6. The method of claim 1, wherein the contacting comprises use of one type of polymer-nucleotide conjugate.
7. The method of claim 1, wherein the contacting comprises use of two or more types of polymer-nucleotide conjugate.

8. The method of claim 7, wherein each type of the two or more types of polymer-nucleotide conjugate comprises a different type of nucleotide moiety.
9. The method of claim 8, wherein the contacting comprises use of three types of polymer-nucleotide conjugate and wherein each type of the three types of polymer-nucleotide conjugate comprises a different type of nucleotide moiety.
10. The method of claim 1, wherein the polymer-nucleotide conjugate comprises a blocked nucleotide moiety.
11. The method of claim 10, wherein the blocked nucleotide is a 3'-O-azidomethyl, 3'-O-methyl, or 3'-O-alkyl hydroxylamine nucleotide.
12. The method of claim 1, wherein said contacting occurs in the presence of an ion that stabilizes said multivalent binding complex.
13. The method of claim 1, wherein the contacting is done in the presence of strontium ions, magnesium ions, calcium ions, or any combination thereof.
14. The method of claim 1, wherein the polymerase molecules are catalytically inactive.
15. The method of claim 1, wherein the polymerase molecules have been rendered catalytically inactive by mutation or chemical modification.
16. The method of claim 1, wherein the polymerase molecules have been rendered catalytically inactive by the absence of a necessary ion or cofactor.
17. The method of claim 1, wherein the polymerase molecules are catalytically active.
18. The method of claim 1, wherein the polymer-nucleotide conjugate does not comprise a blocked nucleotide moiety.
19. The method of claim 1, wherein the multivalent binding complex has a persistence time of greater than 2 seconds.
20. The method of claim 1, wherein the method is carried out at a temperature within a range of 25°C to 42°C.
21. The method of claim 1, wherein the polymer-nucleotide conjugate further comprises one or more fluorescent labels and the two or more copies of the target nucleic acid sequence are deposited on, attached to, or hybridized to a surface, wherein a fluorescence image of the multivalent binding complex on the surface has a contrast to noise ratio in the detecting step of greater than 20.

22. The method of claim 1, wherein the composition of (a) is deposited on a surface using a buffer that incorporates a polar aprotic solvent.
23. The method of claim 1, wherein the contacting is performed under a condition that stabilizes said multivalent binding complex when said nucleotide moiety is complementary to a next base of said target nucleic acid sequence and destabilizes said multivalent binding complex when said nucleotide moiety is not complementary to said next base of said target nucleic acid sequence.
24. The method of claim 1, wherein said polymer-nucleotide conjugate comprises a polymer having a plurality of branches and said two or more nucleotide moieties are attached to said branches.
25. The method of claim 24, wherein said first polymer has a star, comb, cross-linked, bottle brush, or dendrimer configuration.
26. The method of claim 1, wherein said polymer-nucleotide conjugate comprises one or more binding groups selected from the group consisting of an avidin, a biotin, an affinity tag, and combinations thereof.
27. The method of claim 1, further comprising a dissociation step that destabilizes said multivalent binding complex formed between the composition of (a) and the polymer-nucleotide conjugate, said dissociation step enabling removal of said polymer-nucleotide conjugate.
28. The method of claim 27, further comprising an extension step to incorporate a nucleotide that is complementary to a next base of the target nucleic acid sequence into said two or more primer nucleic acid molecules.
29. The method of claim 28, wherein the extension step occurs concurrently with or after said dissociation step.
30. The method of claim 1, wherein said polymer-nucleotide conjugate comprises one or more detectable labels.