

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2010301128 B2**

(54) Title
Ultrafast sequencing of biological polymers using a labeled nanopore

(51) International Patent Classification(s)
G01N 33/00 (2006.01) **B82B 3/00** (2006.01)

(21) Application No: **2010301128** (22) Date of Filing: **2010.05.13**

(87) WIPO No: **WO11/040996**

(30) Priority Data

(31) Number	(32) Date	(33) Country
61/277,939	2009.09.30	US

(43) Publication Date: **2011.04.07**

(44) Accepted Journal Date: **2014.09.18**

(71) Applicant(s)
Quantapore, Inc.

(72) Inventor(s)
Huber, Martin

(74) Agent / Attorney
**Watermark Patent and Trade Marks Attorneys, Level 2 302 Burwood Road,
HAWTHORN, VIC, 3122**

(56) Related Art
US 6528258
US 2005/0095599

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 April 2011 (07.04.2011)

(10) International Publication Number
WO 2011/040996 A1

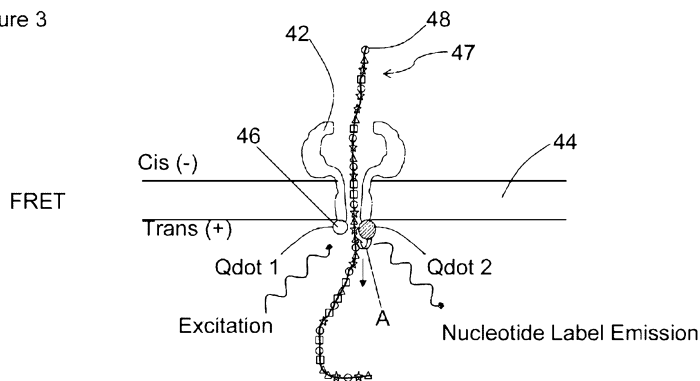
- (51) International Patent Classification:
G01N 33/00 (2006.01) *B82B 3/00* (2006.01)
- (21) International Application Number:
PCT/US2010/034809
- (22) International Filing Date:
13 May 2010 (13.05.2010)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/277,939 30 September 2009 (30.09.2009) US
- (71) Applicant (for all designated States except US): **QUANTAPORE, INC.** [US/US]; 1455 Adams Drive, Suite 1081, Menlo Park, California 94025 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **HUBER, Martin** [US/US]; 1066 Pine Street, Menlo Park, California 94025 (US).
- (74) Agents: **GIOVANNETTI, Steven M.** et al.; Levine Bagade Han LLP, 2400 Geng Road, Suite 120, Palo Alto, California 94303 (US).

- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report (Art. 21(3))

(54) Title: ULTRAFAST SEQUENCING OF BIOLOGICAL POLYMERS USING A LABELED NANOPORE

Figure 3



(57) Abstract: Methods and systems for sequencing a biological molecule or polymer, e.g., a nucleic acid, are provided. One or more donor labels, which are attached to a pore or nanopore, may be illuminated or otherwise excited. A polymer having a monomer labeled with one or more acceptor labels, may be translocated through the pore. Either before, after or while the labeled monomer of the polymer passes through, exits or enters the pore, energy may be transferred from the excited donor label to the acceptor label of the monomer. As a result of the energy transfer, the acceptor label emits energy, and the emitted energy is detected in order to identify the labeled monomer of the translocated polymer and to thereby sequence the polymer.

WO 2011/040996 A1

ULTRAFAST SEQUENCING OF BIOLOGICAL POLYMERS USING A LABELED
NANOPORE

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of priority to U.S. Prov. Pat. App. 61/277,939 filed September 30, 2009, which is incorporated herein by reference in its entirety.

BACKGROUND

[0002] DNA is a long bio-polymer made from repeating units called nucleotides. DNA polymers can be enormous molecules containing millions of nucleotides e.g. the human genome contains a total of 3 billion nucleotides. In living organisms, DNA does not usually exist as a single molecule, but instead as a tightly-associated pair of molecules. These two long strands intertwine like vines, in the shape of a double helix. The nucleotide repeats contain both a phosphate backbone which holds the chain together, and a base, which interacts with the other DNA strand in the helix. This interaction between the bases of the two DNA strands is called hydrogen bonds and they hold the double helix together. There are four different types of bases: Adenine (A), Cytosine (C), Guanine (G) and Thymine (T). Each type of base in one strand forms a hydrogen bond with just one type of base in the complementary strand, with A bonding only to T, and C bonding only to G.

[0003] The sequence of the four bases determines the genetic information contained in DNA. Revealing the sequence of the four building blocks of polynucleic acid is called sequencing. Polynucleic acid comprises bases of nucleosides chemically bound in a linear fashion. "DNA" (De-oxyribonucleic acid) and "RNA" (Ribonucleic acid) are examples of such polynucleic acid molecules.. The particular order or "sequence" of these bases in a given gene determines the structure of the protein encoded by the gene. Furthermore, the sequence of bases surrounding the gene typically contains information about how often the particular protein should be made, in which cell types etc.

[0004] The complete nucleotide sequence of all DNA polymers in a particular individual is known as that individual's "genome". In 2003 the human genome project was finished and a draft version of the human DNA sequence was presented. It took 13 years, 3

billion US \$ and the joint power of multiple sequencing centers to achieve this scientific milestone which was compared in significance to the arrival of men on the moon. The method used for this giant project is called Sanger sequencing (Sanger, F. et al., Proc. Natl. Acad. Sci. USA (1977) 74, 5463-5467 and Smith et al., US Pat. No. 5,821,058). Although major technical improvements were made during this time, the classical sequencing method has some key-disadvantages:

[0005] • Laborious sample preparation, including subcloning of DNA fragments in bacteria

[0006] • Expensive automation

[0007] • Cost prohibitive molecular biology reagents

[0008] • Limited throughput which results in years to finish sequencing whole genomes

[0009] Multiple diseases have a strong genetic component (Strittmatter, W.J. et al., Annual Review of Neuroscience 19 (1996): 53–77; Ogura, Y. et al., Nature 411, (2001): 603–606; Begovich, A.B. et al., American Journal of Human Genetics 75, (2004): 330–337). With the completion of the Human Genome Project and an ever deepening comprehension of the molecular basis of disease, medicine in the 21st century is poised for a revolution called “molecular diagnostics”. Most commercial and academic approaches in molecular diagnostics assess single nucleotide variations (SNPs) or mutations to identify DNA aberrations. These technologies, although powerful, will analyze only a small portion of the entire genome. The inability to accurately and rapidly sequence large quantities of DNA remains an important bottleneck for research and drug development (Shaffer, C., Nat Biotech 25 (2007): 149). Clearly, there is a need for the development of improved sequencing technologies that are faster, easier to use, and less expensive.

BRIEF SUMMARY

[0010] Variations described herein relate to methods, systems and/or devices for detecting the sequence composition of biological polymers. For example, methods and devices are described herein which are capable of ultrafast polymer sequencing utilizing a labeled pore or nanopore and a biological polymer with labeled monomer building blocks.

[0011] Methods and systems for sequencing a biological molecule or polymer, e.g., a nucleic acid, are provided. One or more donor labels, which are positioned on, attached or connected to a pore or nanopore, may be illuminated or otherwise excited. A polymer labeled with one or more acceptor labels, may be translocated through the nanopore. For

example, a polymer having one or more monomers labeled with one or more acceptor labels, may be translocated through the nanopore. Either before, after or while the labeled monomer of the polymer or molecule passes through, exits or enters the nanopore and when an acceptor label comes into proximity with a donor label, energy may be transferred from the excited donor label to the acceptor label of the monomer or polymer. As a result of the energy transfer, the acceptor label emits energy, and the emitted energy is detected or measured in order to identify the monomer, e.g., the nucleotides of a translocated nucleic acid molecule, which is associated with the detected acceptor label energy emission. The nucleic acid or other polymer may be deduced or sequenced based on the detected or measured energy emission from the acceptor labels and the identification of the monomers or monomer sub units.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0012] Figure 1A illustrates a variation of a synthetic nanopore having a pore label attached thereto.

[0013] Figure 1B illustrates a variation of a protein nanopore having a pore label attached thereto.

[0014] Figure 2A illustrates one variation of a FRET (Förster Resonance Energy Transfer) interaction between a pore label on a synthetic nanopore and a nucleic acid label on a nucleic acid which is being translocated through the synthetic nanopore.

[0015] Figure 2B illustrates translocation of the labeled nucleic acid through a synthetic nanopore at a point in time where no FRET is taking place.

[0016] Figure 2C illustrates one variation of a FRET interaction between a pore label on a protein nanopore and a nucleic acid label on a nucleic acid which is being translocated through the protein nanopore.

[0017] Figure 2D illustrates translocation of a labeled nucleic acid through a protein nanopore at a point in time where no FRET is taking place.

[0018] Figure 3 illustrates one variation of a multicolor FRET interaction between the donor labels (Quantum dots) of a protein nanopore and the acceptor labels of a nucleic acid. Each shape on the nucleic acid represents a specific acceptor label, where each label has a distinct emission spectra associated with a specific nucleotide such that each label emits light at a specific wavelength associated with a specific nucleotide.

[0019] Figure 4A illustrates partial contigs from nucleic acid sequencing utilizing a singly labeled nucleic acid.

[0020] Figure 4B illustrates how partial contig alignment may generate a first draft nucleic acid sequence.

[0021] Figure 5A illustrates one variation of a quenching interaction between a pore label on a synthetic nanopore and a nucleic acid label on a nucleic acid which is being translocated through the synthetic nanopore.

[0022] Figure 5B illustrates translocation of the labeled nucleic acid through a synthetic nanopore at a point in time where no quenching is taking place.

[0023] Figure 5C illustrates one variation of a quenching interaction between a pore label on a protein nanopore and a nucleic acid label on a nucleic acid which is being translocated through the protein nanopore.

[0024] Figure 5D illustrates translocation of a labeled nucleic acid through a protein nanopore at a point in time where no quenching is taking place.

[0025] Figure 6 shows an example of an absorption/emission spectra from a FRET pair containing a donor quantum dot and an acceptor fluorophore.

DETAILED DESCRIPTION

[0026] A method and/or system for sequencing a biological polymer or molecule (e.g., a nucleic acid) may include exciting one or more donor labels attached to a pore or nanopore. A biological polymer may be translocated through the pore or nanopore, where a monomer of the biological polymer is labeled with one or more acceptor labels. Energy may be transferred from the excited donor label to the acceptor label of the monomer as, after or before the labeled monomer passes through, exits or enters the pore or nanopore. Energy emitted by the acceptor label as a result of the energy transfer may be detected, where the energy emitted by the acceptor label may correspond to or be associated with a single or particular monomer (e.g., a nucleotide) of a biological polymer. The sequence of the biological polymer may then be deduced or sequenced based on the detection of the emitted energy from the monomer acceptor label which allows for the identification of the labeled monomer. A pore, nanopore, channel or passage, e.g., an ion permeable pore, nanopore, channel or passage may be utilized in the systems and methods described herein.

[0027] Nanopore energy transfer sequencing (NETS) can be used to sequence nucleic acid. NETS can enable the sequencing of whole genomes within days for a fraction of

today's cost which will revolutionize the understanding, diagnosis, monitoring and treatment of disease. The system or method can utilize a pore or nanopore (synthetic or protein-based) of which one side, either the cis (-) or trans (+) side of the pore is labeled with one or multiple or a combination of different energy absorbers or donor labels, such as fluorophores, fluorescent proteins, quantum dots, metal nanoparticles, nanodiamonds, etc. Multiple labels and methods of labeling a nanopore are described in U.S. Pat. No. 6,528,258, the entirety of which is incorporated herein by reference.

[0028] A nucleic acid can be threaded through a nanopore by applying an electric field through the nanopore (Kasianowicz, J.J. et al., Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci USA* 93 (1996): 13770-13773). A nucleic acid to be translocated through the nanopore may undergo a labeling reaction where naturally occurring nucleotides are exchanged with a labeled, energy emitting or absorbing counterpart or modified counterparts that can be subsequently modified with an energy emitting or absorbing label, i.e., an acceptor label. The labeled nucleic acid may then be translocated through the nanopore and upon entering, exiting or while passing through the nanopore a labeled nucleotide comes in close proximity to the nanopore or donor label. For example, within 1-10nm or 1-2nm of the nanopore donor label. The donor labels may be continuously illuminated with radiation of appropriate wavelength to excite the donor labels. Via a dipole-dipole energy exchange mechanism called FRET (Stryer, L. *Annu Rev Biochem.* 47 (1978): 819-846), the excited donor labels transfer energy to a bypassing nucleic acid or acceptor label. The excited acceptor label may then emit radiation, e.g., at a lower energy than the radiation that was used to excite the donor label. This energy transfer mechanism allows the excitation radiation to be "focused" to interact with the acceptor labels with sufficient resolution to generate a signal at the single nucleotide scale.

[0029] A nanopore may include any opening positioned in a substrate that allows the passage of a molecule through the substrate. For example, the nanopore may allow passage of a molecule that would otherwise not be able to pass through that substrate. Examples of nanopores include proteinaceous or protein based pores or synthetic pores. A nanopore may have an inner diameter of 1-10 nm or 1-5 nm or 1-3 nm.

[0030] Examples of protein pores include but are not limited to, alpha-hemolysin, voltage-dependent mitochondrial porin (VDAC), OmpF, OmpC, MspA and LamB (malto porin) (Rhee, M. et al., *Trends in Biotechnology*, 25(4) (2007): 174-181). Any protein pore that allows the translocation of single nucleic acid molecules may be

employed. A pore protein may be labeled at a specific site on the exterior of the pore, or at a specific site on the exterior of one or more monomer units making up the pore forming protein.

[0031] A synthetic pore may be created in various forms of solid substrates, examples of which include but are not limited to silicones (e.g. Si₃N₄, SiO₂), metals, metal oxides (e.g. Al₂O₃) plastics, glass, semiconductor material, and combinations thereof. A synthetic nanopore may be more stable than a biological protein pore positioned in a lipid bilayer membrane.

[0032] Synthetic nanopores may be created using a variety of methods. For example, synthetic nanopores may be created by ion beam sculpting (Li, J. et al., *Nature* 412 (2001): 166-169) where massive ions with energies of several thousand electron volts (eV) cause an erosion process when fired at a surface which eventually will lead to the formation of a nanopore. A synthetic nanopore may be created via latent track etching. For example, a single conical synthetic nanopore may be created in a polymer substrate by chemically etching the latent track of a single, energetic heavy ion. Each ion produces an etchable track in a polymer foil, forming a one-pore membrane (Heins, E.A. et al., *Nano Letters* 5 (2005): 1824-1829). A synthetic nanopore may also be created by a method called Electron beam-induced fine tuning. Nanopores in various materials have been fabricated by advanced nanofabrication techniques, such as FIB drilling and electron (E) beam lithography, followed by E-beam assisted fine tuning techniques. With the appropriate electron beam intensity applied, a previously prepared nanopore will start to shrink. The change in pore diameter may be monitored in real-time using a TEM (transmission electron microscope), providing a feedback mechanism to switch off the electron beam at any desired dimension of the nanopore (Lo, C.J. et al., *Nanotechnology* 17 (2006): 3264-67).

[0033] A synthetic nanopore may also be created by using a carbon nanotube embedded in a suitable substrate such as but not limited to polymerized epoxy. Carbon nanotubes can have uniform and well-defined chemical and structural properties. Various sized carbon nanotubes can be obtained, ranging from one to hundreds of nanometers. The surface charge of a carbon nanotube is known to be about zero, and as a result, electrophoretic transport of a nucleic acid through the nanopore becomes simple and predictable (Ito, T. et al., *Chem. Commun.* 12 (2003): 1482-83).

[0034] A pore may have two sides. One side is referred to as the "cis" side and faces the (-) negative electrode or a negatively charged buffer/ion compartment or solution. The

other side is referred to as the "trans" side and faces the (+) electrode or a positively charged buffer/ion compartment or solution. A biological polymer, such as a labeled nucleic acid molecule or polymer can be pulled or driven through the pore by an electric field applied through the nanopore, e.g., entering on the cis side of the nanopore and exiting on the trans side of the nanopore.

[0035] A nanopore or pore may be labeled with one or more donor labels. For example, the cis side or surface and/or trans side or surface of the nanopore may be labeled with one or more donor labels. The label may be attached to the base of a pore or nanopore or to another portion or monomer making up the nanopore or pore. A label may be attached to a portion of the membrane or substrate through which a nanopore spans or to a linker or other molecule attached to the membrane, substrate or nanopore. The nanopore or pore label may be positioned or attached on the nanopore, substrate or membrane such that the pore label can come into proximity with an acceptor label of a biological polymer, e.g., a nucleic acid, which is translocated through the pore. The donor labels may have the same or different emission or absorption spectra.

[0036] The labeling of a pore structure may be achieved via covalent or non-covalent interactions. Examples of such interactions include but are not limited to interactions based on hydrogen bonds, hydrophobic interactions, electrostatic interactions, ionic interactions, magnetic interactions, Van der Waals forces or combinations thereof.

[0037] A donor label may be placed as close as possible to the aperture of a nanopore without causing an occlusion that impairs translocation of a nucleic acid through the nanopore (see e.g., Figure 1). A pore label may have a variety of suitable properties and/or characteristics. For example, a pore label may have energy absorption properties meeting particular requirements. A pore label may have a large radiation energy absorption cross-section, ranging, for example, from about 0 to 1000 nm or from about 200 to 500 nm. A pore label may absorb radiation within a specific energy range that is higher than the energy absorption of the nucleic acid label. The absorption energy of the pore label may be tuned with respect to the absorption energy of a nucleic acid label in order to control the distance at which energy transfer may occur between the two labels. A pore label may be stable and functional for at least 10^6 or 10^9 excitation and energy transfer cycles.

[0038] Figure 1A shows a variation of a pore/substrate assembly 1. The pore/substrate assembly 1 includes a synthetic pore or nanopore 2 which has a pore label 6 attached thereto. The assembly may also include a substrate 4, e.g., a solid substrate, and the synthetic nanopore 2 is positioned in the substrate 4. The synthetic nanopore 2 is modified

at the trans (+) side with one or more pore labels 6. The pore label 6 is attached to the base of the synthetic nanopore 2 in a manner such that the label 6 does not lead to inclusion or impair the translocation of a nucleic acid through the synthetic nanopore 2.

[0039] Figure 1B shows a variation of a pore/lipid bilayer assembly 10. The pore/lipid bilayer assembly 10 includes a protein nanopore 12 which has a pore label 16 attached thereto. The assembly may also include a lipid bilayer 14 and the protein nanopore 12 is positioned in the lipid bilayer 14. The protein nanopore 12 is modified at the trans (+) side with one or more pore labels 16. The pore label 16 is attached to the base of the protein nanopore 12 in a manner such that the label 16 does not lead to inclusion or impair the translocation of a nucleic acid through the protein nanopore 12.

[0040] A protein nanopore can be embedded in a phospholipid bilayer or derivatizations thereof. Phospholipids are comprised of, but not limited to diphytanoyl-phosphatidylcholine, soybean azolectin, 1,2-Diphytanoyl-sn-glycero-3-phosphocholine. A lipid bilayer can also be prepared by a mixture of different phospholipids. Possible solvents for phospholipids are hexadecane, pentane, chloroform or any other suitable organic solvent. A lipid bilayer may be prepared in variety of ways known to those having ordinary skill in the art.

[0041] A lipid bilayer (e.g., including the above mentioned phospholipids) having a pore protein may be prepared according to the following method: A 10-25 μ m thick Teflonfilm (Septum) with a 1-100 μ m aperture separates two buffer compartments made out of Teflon. The septum/aperture is primed with 10% hexadecane in pentane on each side and after evaporation of the solvent the buffer compartments are filled with 1 molar KCl. 1,2-Diphytanoyl-sn-glycero-3-phosphocholine (Avanti, 10mg/mL in pentane) is added to each buffer compartment and the pentane is allowed to evaporate leaving behind lipid monolayers. Lowering and raising the liquid level in the chamber below and above the aperture caused lipid bilayers to be formed as needed. The formation of the bilayer is measured by applying voltage via Ag/AgCl electrodes.

[0042] Once the bilayer has formed the ionic current is completely eliminated. With the bilayer in place a dilute solution of pore protein is added to the cis-chamber. Pore proteins are chosen from a group of proteins such as, but not limited to, alpha-hemolysin, voltage-dependent mitochondrial porin (VDAC), Anthrax porin, OmpF, OmpC and LamB (maltoporin). The pore will self assemble and integrate into the lipid bilayer. Integration of the pore protein can be measured by a small but constant current. Typically, one inserted hemolysin pore can carry an ionic current of approximately 120 pA (picoAmperes), with an

applied voltage of +120 mV (milliVolts). The membrane can be protected by a second layer of a polymeric structure, comprising but not limited to Agarose, Polyacrylamide, etc. (Kang, X.-F. et al., *J Am Chem Soc.* 129 (2007): 4701-4705).

[0043] Various polymers or molecules may be attached to a lipid bilayer having a pore protein therein to provide additional stability and support to the pore/membrane assembly and the pore should the membrane be damaged.

[0044] A pore label may include one or more Quantum dots. A Quantum dot has been demonstrated to have many or all of the above described properties and characteristics found in suitable pore labels (Bawendi M.G. in US 6,251,303). Quantum Dots are nanometer scale semiconductor crystals that exhibit strong quantum confinement due to the crystals radius being smaller than the Bohr exciton radius. Due to the effects of quantum confinement, the bandgap of the quantum dots increases with decreasing crystal size thus allowing the optical properties to be tuned by controlling the crystal size (Bawendi M.G. et al., in US 7,235,361 and Bawendi M.G. et al., in US 6,855,551).

[0045] One example of a Quantum dot which may be utilized as a pore label is a CdTe quantum dot which can be synthesized aqueously. A CdTe quantum dot may be functionalized with a nucleophilic group such as primary amines, thiols or functional groups such as carboxylic acids. A CdTe quantum dot may include a mercaptopropionic acid capping ligand, which has a carboxylic acid functional group that may be utilized to covalently link a quantum dot to a primary amine on the exterior of a protein pore. The cross-linking reaction may be accomplished using standard cross-linking reagents (homo-bifunctional as well as hetero-bifunctional) which are known to those having ordinary skill in the art of bioconjugation. Care may be taken to ensure that the modifications do not impair or substantially impair the translocation of a nucleic acid through the nanopore. This may be achieved by varying the length of the employed crosslinker molecule used to attach the donor label to the nanopore.

[0046] The primary amine of the Lysin residue 131 of the natural alpha hemolysin protein (Song, L. et al., *Science* 274, (1996): 1859-1866) may be used to covalently bind carboxy modified CdTe Quantum dots via 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride/ N-hydroxysulfosuccinimide (EDC/NHS) coupling chemistry.

[0047] A variety of methods, mechanisms and/or routes for attaching one or more pore labels to a pore protein may be utilized. A pore protein may be genetically engineered in a manner that introduces amino acids with known properties or various functional groups to

the natural protein sequence. Such a modification of a naturally occurring protein sequence may be advantageous for the bioconjugation of Quantum dots to the pore protein. For example, the introduction of a Cystein residue would introduce a thiol group that would allow for the direct binding of a Quantum dot, such as a CdTe quantum dot, to a pore protein. Also, the introduction of a Lysin residue would introduce a primary amine for binding a Quantum dot. The introduction of glutamic acid or aspartic acid would introduce a carboxylic acid moiety for binding a Quantum dot. These groups are amenable for bioconjugation with a Quantum dot using either homo- or hetero-bifunctional crosslinker molecules. Such modifications to pore proteins aimed at the introduction of functional groups for bioconjugation are known to those having ordinary skill in the art. Care should be taken to ensure that the modifications do not impair or substantially impair the translocation of a nucleic acid through the nanopore.

[0048] The nanopore label can be attached to a protein nanopore before or after insertion of said nanopore into a lipid bilayer. Where a label is attached before insertion into a lipid bilayer, care may be taken to label the base of the nanopore and avoid random labeling of the pore protein. This can be achieved by genetic engineering of the pore protein to allow site specific attachment of the pore label (see section 0047). An advantage of this approach is the bulk production of labeled nanopores. Alternatively, a labeling reaction of a pre-inserted nanopore may ensure site-specific attachment of the label to the base (trans-side) of the nanopore without genetically engineering the pore protein.

[0049] A biological polymer, e.g., a nucleic acid molecule or polymer, may be labeled with one or more acceptor labels. For a nucleic acid molecule, each of the four nucleotides or building blocks of a nucleic acid molecule may be labeled with an acceptor label thereby creating a labeled (e.g., fluorescent) counterpart to each naturally occurring nucleotide. The acceptor label may be in the form of an energy accepting molecule which can be attached to one or more nucleotides on a portion or on the entire strand of a converted nucleic acid.

[0050] A variety of methods may be utilized to label the monomers or nucleotides of a nucleic acid molecule or polymer. A labeled nucleotide may be incorporated into a nucleic acid during synthesis of a new nucleic acid using the original sample as a template ("labeling by synthesis"). For example, the labeling of nucleic acid may be achieved via PCR, whole genome amplification, rolling circle amplification, primer extension or the like or via various combinations and extensions of the above methods known to persons having ordinary skill in the art.

[0051] Labeling of a nucleic acid may be achieved by replicating the nucleic acid in the presence of a modified nucleotide analog having a label, which leads to the incorporation of that label into the newly generated nucleic acid. The labeling process can also be achieved by incorporating a nucleotide analog with a functional group that can be used to covalently attach an energy accepting moiety in a secondary labeling step. Such replication can be accomplished by whole genome amplification (Zhang, L. et al., *Proc. Natl. Acad. Sci. USA* 89 (1992): 5847) or strand displacement amplification such as rolling circle amplification, nick translation, transcription, reverse transcription, primer extension and polymerase chain reaction (PCR), degenerate oligonucleotide primer PCR (DOP-PCR) (Telenius, H. et al., *Genomics* 13 (1992): 718-725) or combinations of the above methods.

[0052] A label may comprise a reactive group such as a nucleophile (amines, thiols etc.). Such nucleophiles, which are not present in natural nucleic acids, can then be used to attach fluorescent labels via amine or thiol reactive chemistry such as NHS esters, maleimides, epoxy rings, isocyanates etc. Such nucleophile reactive fluorescent dyes (i.e. NHS-dyes) are readily commercially available from different sources. An advantage of labeling a nucleic acid with small nucleophiles lies in the high efficiency of incorporation of such labeled nucleotides when a "labeling by synthesis" approach is used. Bulky fluorescently labeled nucleic acid building blocks may be poorly incorporated by polymerases due to sterical hindrance of the labels during the polymerization process into newly synthesized DNA.

[0053] DNA can be directly chemically modified without polymerase mediated incorporation of labeled nucleotides. One example of a modification includes cis-platinum containing dyes that modify Guanine bases at their N7 position (Hoewel, T. et al., *Bio Techniques* 27 (1999): 1064-1067). Another example includes the modifying of pyrimidines with hydroxylamine at the C6 position which leads to 6-hydroxylamino derivatives. The resulting amine groups can be further modified with amine reactive dyes (e.g. NHS-Cy5).

[0054] A nucleic acid molecule may be directly modified with N-Bromosuccinimide which upon reacting with the nucleic acid will result in 5-Bromocystein, 8-Bromo adenine and 8-Bromoguanine. The modified nucleotides can be further reacted with di-amine nucleophiles. The remaining nucleophile can then be reacted with an amine reactive dye (e.g. NHS-dye) (Hermanson G. in *Bioconjugate Techniques*, Academic Press 1996, ISBN 978-0-12-342336-8).

[0055] A combination of 1, 2, 3 or 4 nucleotides in a nucleic acid strand may be exchanged with their labeled counterpart. The various combinations of labeled nucleotides can be sequenced in parallel, e.g., labeling a source nucleic acid or DNA with combinations of 2 labeled nucleotides in addition to the four single labeled samples, which will result in a total of 10 differently labeled sample nucleic acid molecules or DNAs (G, A, T, C, GA, GT, GC, AT, AC, TC). The resulting sequence pattern may allow for a more accurate sequence alignment due to overlapping nucleotide positions in the redundant sequence read-out.

[0056] A method for sequencing a polymer, such as a nucleic acid molecule includes providing a nanopore or pore protein (or a synthetic pore) inserted in a membrane or membrane like structure or other substrate. The base or other portion of the pore may be modified with one or more pore labels. The base may refer to the Trans side of the pore. Optionally, the Cis and/or Trans side of the pore may be modified with one or more pore labels. Nucleic acid polymers to be analyzed or sequenced may be used as a template for producing a labeled version of the nucleic acid polymer, in which one of the four nucleotides or up to all four nucleotides in the resulting polymer is/are replaced with the nucleotide's labeled analogue(s). An electric field is applied to the nanopore which forces the labeled nucleic acid polymer through the nanopore, while an external monochromatic or other light source may be used to illuminate the nanopore, thereby exciting the pore label. As, after or before labeled nucleotides of the nucleic acid pass through, exit or enter the nanopore, energy is transferred from the pore label to a nucleotide label, which results in emission of lower energy radiation. The nucleotide label radiation is then detected by a confocal microscope setup or other optical detection system or light microscopy system capable of single molecule detection known to people having ordinary skill in the art. Examples of such detection systems include but are not limited to confocal microscopy, epifluorescent microscopy and total internal reflection fluorescent (TIRF) microscopy. Other polymers (e.g., proteins and polymers other than nucleic acids) having labeled monomers may also be sequenced according to the methods described herein.

[0057] Energy may be transferred from a pore or nanopore donor label (e.g., a Quantum Dot) to an acceptor label on a polymer (e.g., a nucleic acid) when an acceptor label of an acceptor labeled monomer (e.g., nucleotide) of the polymer interacts with the donor label as, after or before the labeled monomer exits, enters or passes through a nanopore. For example, the donor label may be positioned on or attached to the nanopore on the cis or trans side or surface of the nanopore such that the interaction or energy

transfer between the donor label and acceptor label does not take place until the labeled monomer exits the nanopore and comes into the vicinity or proximity of the donor label outside of the nanopore channel or opening. As a result, interaction between the labels, energy transfer from the donor label to the acceptor label, emission of energy from the acceptor label and/or measurement or detection of an emission of energy from the acceptor label may take place outside of the passage, channel or opening running through the nanopore, e.g., within a cis or trans chamber on the cis or trans sides of a nanopore. The measurement or detection of the energy emitted from the acceptor label of a monomer may be utilized to identify the monomer.

[0058] The nanopore label may be positioned outside of the passage, channel or opening of the nanopore such that the label may be visible or exposed to facilitate excitation or illumination of the label. The interaction and energy transfer between a donor label and acceptor label and the emission of energy from the acceptor label as a result of the energy transfer may take place outside of the passage, channel or opening of the nanopore. This may facilitate ease and accuracy of the detection or measurement of energy or light emission from the acceptor label, e.g., via an optical detection or measurement device. The donor and acceptor label interaction may take place within a channel of a nanopore and a donor label could be positioned within the channel of a nanopore.

[0059] A donor label may be attached in various manners and/or at various sites on a nanopore. For example, a donor label may be directly or indirectly attached or connected to a portion or unit of the nanopore. Alternatively, a donor label may be positioned adjacent to a nanopore.

[0060] Each acceptor labeled monomer (e.g., nucleotide) of a polymer (e.g., nucleic acid) can interact sequentially with a donor label positioned on or next to or attached directly or indirectly to a nanopore or channel through which the polymer is translocated. The interaction between the donor and acceptor labels may take place outside of the nanopore channel or opening, e.g., after the acceptor labeled monomer exits the nanopore or before the monomer enters the nanopore. The interaction may take place within or partially within the nanopore channel or opening, e.g., while the acceptor labeled monomer passes through, enters or exits the nanopore.

[0061] When one of the four nucleotides of a nucleic acid is labeled, the time dependent signal arising from the single nucleotide label emission is converted into a sequence corresponding to the positions of the labeled nucleotide in the nucleic acid

sequence. The process is then repeated for each of the four nucleotides in separate samples and the four partial sequences are then aligned to assemble an entire nucleic acid sequence.

[0062] When multi-color labeled nucleic acid (DNA) sequences are analyzed, the energy transfer from one or more donor labels to each of the four distinct acceptor labels that may exist on a nucleic acid molecule may result in light emission at four distinct wavelengths or colors (each associated with one of the four nucleotides) which allows for a direct sequence read-out.

[0063] During sequencing of a nucleic acid molecule, the energy transfer signal may be generated with sufficient intensity that a sensitive detection system can accumulate sufficient signal within the transit time of a single nucleotide through the nanopore to distinguish a labeled nucleotide from an unlabeled nucleotide. Therefore, the pore label may be stable, have a high absorption cross-section, a short excited state lifetime, and/or temporally homogeneous excitation and energy transfer properties. The nucleotide label may be capable of emitting and absorbing sufficient radiation to be detected during the transit time of the nucleotide through the pore. The product of the energy transfer cross-section, emission rate, and quantum yield of emission may yield sufficient radiation intensity for detection within the single nucleotide transit time. A nucleotide label may also be sufficiently stable to emit the required radiation intensity and without transience in radiation emission.

[0064] The excitation radiation source may be of high enough intensity that when focused to the diffraction limit on the nanopore, the radiation flux is sufficient to saturate the pore label. The detection system may filter out excitation radiation and pore label emission while capturing nucleic acid label emission during pore transit with sufficient signal-to-noise ratio (S/N) to distinguish a labeled nucleotide from an unlabeled nucleotide with high certainty. The collected nucleic acid label radiation may be counted over an integration time equivalent to the single nucleotide pore transit time.

[0065] A software signal analysis algorithm may then be utilized which converts the binned radiation intensity signal to a sequence corresponding to a particular nucleotide. Combination and alignment of four individual nucleotide sequences (where one of the four nucleotides in each sequence is labeled) allows construction of the complete nucleic acid sequence via a specifically designed computer algorithm.

[0066] A system for sequencing one or more biological polymers, e.g., nucleic acid molecules, may include a fixture or pore holder. The pore holder may include a nanopore membrane assembly wherein one or more nanopores span a lipid bilayer membrane. The

nanopore membrane assembly has a Cis (-) side and a Trans (+) side. One or more labels may be attached to the nanopores. Alternatively, a label may be attached to a portion of the membrane or substrate through which the nanopore spans or to a linker or other molecule attached to the membrane, substrate or nanopore. An aqueous buffer solution is provided which surrounds the nanopore membrane assembly. The pore holder may contain two electrodes. A negative electrode or terminal may be positioned on the Cis side of the nanopore membrane assembly and a positive electrode or terminal may be positioned on the Trans side of the nanopore membrane assembly.

[0067] A flow of fluid or solution is provided on the side of the nanopore where the translocated polymer or nucleic acid exits after translocation through the nanopore. The flow may be continuous or constant such that the fluid or solution does not remain static for an extended period of time. The fluid flow or motion helps move or transfer translocated polymers away from the nanopore channel such the translocated polymers do not linger or accumulated near the nanopore channel exit or opening and cause fluorescent background or noise which could disrupt or prevent an accurate reading, measurement or detection of the energy emitted by a polymer acceptor label. Translocated polymers may include labels that were not fully exhausted, i.e. haven't reached their fluorescent lifetime and are still able to emit light. Such labels could interfere with the energy transfer between donor labels and subsequent monomer labels or emit energy that may interfere with the emission from other labels and disrupt an accurate reading or detection of energy from a labeled monomer.

[0068] One or more polymers, e.g., nucleic acid polymers or molecules, to be analyzed may also be provided. A polymer or nucleic acid polymer or molecule may include one or more labels, e.g., one or more monomers or nucleotides of the polymer may be labeled. A nucleic acid molecule may be loaded into a port positioned on the Cis side of the nanopore membrane assembly. The membrane segregates the nucleic acids to be analyzed to the Cis side of the nanopore membrane assembly. An energy source for exciting the nanopore label is provided, e.g., an illumination source. An electric field may be applied to or by the electrodes to force the labeled nucleic acid to translocate through the nanopore into the Cis side and out of the Trans side of the nanopore, from the Cis to the Trans side of the membrane, e.g., in a single file (Kasianowicz, J.J. et al., *Proc. Natl. Acad. Sci USA* 93 (1996): 13770-13773). Optionally, an electrical field may be applied utilizing other mechanisms to force the labeled nucleic acid to translocate through the nanopore. When a nucleic acid molecule is translocated through the nanopore and a labeled nucleotide comes

into close proximity with the nanopore label, e.g., upon or after exiting the nanopore, energy is transferred from the excited nanopore label to a nucleotide label. A detector or detection system, e.g., optical detection system, for detecting or measuring energy emitted from the nucleotide label as a result of the transfer of energy from the nanopore label to the nucleotide label may also be provided.

[0069] The pore may be labeled with one or more donor labels in the form of quantum dots, metal nanoparticles, nano diamonds or fluorophores. The pore may be illuminated by monochromatic laser radiation. The monochromatic laser radiation may be focused to a diffraction limited spot exciting the quantum dot pore labels. As the labeled nucleic acid (e.g., labeled with an acceptor label in the form of a fluorophore) is translocated through the nanopore, the pore donor label (also “pore label” or “donor label”) and a nucleotide acceptor label come into close proximity with one another and participate in a FRET (Förster resonance energy transfer) energy exchange interaction between the pore donor label and nucleic acid acceptor label (Ha, T. et al., *Proc. Natl.Acad. Sci USA* 93 (1996): 6264-6268).

[0070] FRET is a non-radiative dipole-dipole energy transfer mechanism from a donor to acceptor fluorophore. The efficiency of FRET may be dependent upon the distance between donor and acceptor as well as the properties of the fluorophores (Stryer, L., *Annu Rev Biochem.* 47 (1978): 819–846).

[0071] A fluorophore may be any construct that is capable of absorbing light of a given energy and re-emitting that light at a different energy. Fluorophores include, e.g., organic molecules, rare-earth ions, metal nanoparticles, nanodiamonds and semiconductor quantum dots.

[0072] Figure 2A shows one variation of a FRET interaction between a pore donor label 26 on a synthetic nanopore 22 and a nucleic acid acceptor label 28 on a nucleic acid 27 (e.g., a single or double stranded nucleic acid), which is being translocated through the synthetic nanopore 22. The synthetic nanopore 22 is positioned in a substrate 24. FRET is a non-radiative dipole-dipole energy transfer mechanism from a donor label 26 to an acceptor label 28 (e.g., a fluorophore). The efficiency of the energy transfer is, among other variables, dependent on the physical distance between acceptor label 28 and the donor label.

[0073] The nucleic acid acceptor label 28 positioned on a nucleotide of the nucleic acid moves into close proximity with an excited nanopore donor label 26, e.g., as or after the label 28 or labeled nucleotide exits the nanopore 22, and gets excited via FRET (indicated

by the arrow A showing energy transfer from the pore label 26 to the nucleic acid label 28). As a result, the nucleic acid label 28 emits light of a specific wavelength, which can then be detected with the appropriate optical equipment or detection system in order to identify the labeled nucleotide corresponding to or associated with the detected wavelength of emitted light.

[0074] Figure 2B shows translocation of the labeled nucleic acid 27 at a point in time where no FRET is taking place (due to the acceptor and donor labels not being in close enough proximity to each other). This is indicated by the lack of any arrows showing energy transfer between a pore label 26 and a nucleic acid label 28.

[0075] Figure 2C shows one variation of a FRET interaction between a pore donor label 36 on a proteinaceous or protein nanopore 32 and a nucleic acid acceptor label 38 on a nucleic acid 37 (e.g., a single or double stranded nucleic acid), which is being translocated through the protein pore or nanopore 32. The pore protein 32 is positioned in a lipid bilayer 34. The nucleic acid acceptor label 38 positioned on a nucleotide of the nucleic acid moves into close proximity with an excited nanopore donor label 36, e.g., as or after the label 38 or labeled nucleotide exits the nanopore 32, and gets excited via FRET (indicated by the arrow A showing energy transfer from the pore label 36 to the nucleic acid label 38). As a result, the nucleic acid label 38 emits light of a specific wavelength, which can be detected with the appropriate optical equipment or detection system in order to identify the labeled nucleotide corresponding to or associated with the detected wavelength of emitted light.

[0076] Figure 2D shows translocation of the labeled nucleic acid 37 at a point in time where no FRET is taking place (due to the labels not being in close enough proximity to each other). This is indicated by the lack of arrows showing energy transfer between a pore donor label 36 and a nucleic acid label 38.

[0077] Three equations are also shown below: Equation (1) gives the Förster radius which is defined as the distance that energy transfer efficiency from donor to acceptor is 50%. The Förster distance depends on the refractive index (n_D), quantum yield of the donor (Q_D), spatial orientation (K) and the spectral overlap of the acceptor and donor spectrum (I). N_A is the Avogadro number with $N_A = 6.022 \times 10^{23} \text{ mol}^{-1}$ (see equation below). Equation (2) describes the overlap integral for the donor and acceptor emission and absorption spectra respectively; Equation (3) shows the FRET energy transfer efficiency as a function of distance between the acceptor and donor pair. The equations demonstrate that spectral overlap controls the Förster radius, which determines the energy transfer efficiency for a

given distance between the FRET pair. Therefore by tuning the emission wavelength of the donor, the distance at which energy transfer occurs can be controlled.

[0078] (1)

$$R_0 = \left(\frac{9000(\ln 10) \kappa_A^2 Q_D}{N_A 128 \pi^3 n_D^4} \right)^{1/6}$$

[0079] (2)

$$I = \int I(\lambda) d\lambda = \int PL_{\text{donor}}(\lambda) \times \lambda^4 \times \epsilon_A(\lambda) d\lambda$$

[0080] (3)

$$E = \frac{k_{DA}}{k_{DA} + \tau_D^{-1}} = \frac{R_0^6}{R_0^6 + r^6}$$

[0081] With respect to Quantum dots, due to the size dependent optical properties of quantum dots, the donor emission wavelength may be adjusted. This allows the spectral overlap between donor emission and acceptor absorption to be adjusted so that the Förster radius for the FRET pair may be controlled. The emission spectrum for Quantum dots is narrow, (e.g., 25nm Full width-half maximum -FWHM- is typical for individual quantum dots) and the emission wavelength is adjustable by size, enabling control over the donor label-acceptor label interaction distance by changing the size of the quantum dots. Another important attribute of quantum dots is their broad absorption spectrum, which allows them to be excited at energies that do not directly excite the acceptor label. The properties allow quantum dots of the properly chosen size to be used to efficiently transfer energy with sufficient resolution to excite individual labeled nucleotides as, after or before the labeled nucleotides travel through a donor labeled pore.

[0082] Following a FRET energy transfer, the pore donor label may return to the electronic ground state and the nucleotide acceptor label can re-emit radiation at a lower energy. Where fluorophore labeled nucleotides are utilized, energy transferred from the fluorophore acceptor label results in emitted photons of the acceptor label. The emitted photons of the acceptor label may exhibit lower energy than the pore label emission. The detection system for fluorescent nucleotide labels may be designed to collect the maximum number of photons at the acceptor label emission wavelength while filtering out emission

from a donor label (e.g., quantum dot donors) and laser excitation. The detection system counts photons from the labeled monomers as a function of time. Photon counts are binned into time intervals corresponding to the translocation time of, for instance, a monomer comprising a single nucleotide in a nucleic acid polymer crossing the nanopore. Spikes in photon counts correspond to labeled nucleotides translocating across the pore. To sequence the nucleic acid, sequence information for a given nucleotide is determined by the pattern of spikes in photon counts as a function of time. An increase in photon counts is interpreted as a labeled nucleotide.

[0083] Translocation of nucleic acid polymers through the nanopore may be monitored by current measurements arising from the flow of ions through the nanopore.

Translocating nucleic acids partially block the ionic flux through the pore resulting in a measurable drop in current. Thus, detection of a current drop represents detection of a nucleic acid entering the pore, and recovery of the current to the original value represents detection of a nucleic acid exiting the pore.

[0084] As mentioned supra, a multicolor FRET interaction is utilized to sequence a molecule such a nucleic acid. Figure 3 shows one variation of a multicolor FRET interaction between one or more donor labels 46 (e.g., Quantum dots) of a protein nanopore 42 (optionally, a synthetic nanopore may be utilized) and one or more acceptor labels 48 of a nucleic acid molecule 47 (e.g., a single or double stranded nucleic acid). Each shape on the nucleic acid 47 represents a specific type of acceptor label labeling a nucleotide, where each label has a distinct emission spectra associated with or corresponding to a specific nucleotide such that each label emits light at a specific wavelength or color associated with a specific nucleotide.

[0085] In Figure 3, each of the four shapes (triangle, rectangle, star, circle) represents a specific acceptor label 48, each label having a distinct emission spectra (e.g., 4 different emission spectra). Each of the acceptor labels 48 can form a FRET pair with a corresponding donor label or quantum dot 46 attached to the base of the nanopore. Qdot1 and Qdot2 represent two different Quantum dots as donor labels 46 that form specific FRET pairs with a nucleic acid acceptor label 48. The Quantum dot donor labels 46 are in an excited state and depending on the particular acceptor label 48 that comes in proximity to the Quantum dots during, after or before a labeled nucleotide translocation through the nanopore 42, an energy transfer (arrow A) from the donor label 46 to the nucleotide acceptor label 48 takes place, resulting in a nucleotide label 48 energy emission. As a result, each nucleotide may emit light at a specific wavelength or color (due to the distinct

emission spectrum of the nucleotide's label), which can be detected (e.g., by optical detection) and used to identify or deduce the nucleotide sequence of the nucleic acid 47 and the nucleic acid 47 sequence.

[0086] Different pore labels exhibiting different spectral absorption maxima may be attached to a single pore. The nucleic acid may be modified with corresponding acceptor dye labeled nucleotides where each donor label forms FRET pairs with one acceptor labeled nucleotide (i.e. multi-color FRET). Each of the four nucleotides may contain a specific acceptor label which gets excited by one or more of the pore donor labels. The base of the pore may be illuminated with different color light sources to accommodate the excitation of the different donor labels. Alternatively, e.g., where Quantum dots are used as donor labels, the broad absorption spectra characteristic of Quantum dots may allow for a single wavelength light source to sufficiently illuminate/excite the different donor labels which exhibit different spectral absorption maxima.

[0087] A single pore donor label (e.g., a single Quantum dot) may be suitable for exciting one nucleic acid acceptor label. For example, four different pore donor labels may be provided where each donor label can excite one of four different nucleic acid acceptor labels resulting in the emission of four distinct wavelengths. A single pore donor label (e.g., a single Quantum dot) may be suitable for exciting two or more nucleic acid acceptor labels that have similar absorption spectra overlapping with the donor label emission spectrum and show different emission spectra (i.e. different Stoke's shifts), where each acceptor label emits light at a different wavelength after excitation by the single donor label. Two different pore donor labels (e.g., two Quantum dots having different emission or absorption spectra) may be suitable for exciting four nucleic acid acceptor labels having different emission or excitation spectra, which each emit light at different wavelengths. One donor label or Quantum dot may be capable of exciting two of the nucleic acid acceptor labels resulting in their emission of light at different wavelengths, and the other Quantum dot may be capable of exciting the other two nucleic acid acceptor labels resulting in their emission of light at different wavelengths. The above arrangements provide clean and distinct wavelength emissions from each nucleic acid acceptor label for accurate detection.

[0088] A nanopore may include one or more monomers or attachment points, e.g., about 7 attachment points, one on each of the seven monomers making up a particular protein nanopore, such as alpha-hemolysin. One or more different donor labels, e.g., Quantum dots, may attach one to each of the attachment points, e.g., a nanopore may have

up to seven different Quantum dots attached thereto. A single donor label or Quantum dot may be used to excite all four different nucleic acid acceptor labels resulting in a common wavelength emission suitable for detecting a molecule or detecting the presence of a molecule, e.g., in a biosensor application.

[0089] For accumulation of the raw signal data where a multi-color FRET interaction is utilized, the emission wavelength of the four different acceptor labels may be filtered and recorded as a function of time and emission wavelength, which results in a direct read-out of sequence information.

[0090] As mentioned supra, a nucleic acid sample may be divided into four parts to sequence the nucleic acid. The four nucleic acid or DNA samples may be used as a template to synthesize a labeled complementary nucleic acid polymer. Each of the four nucleic acid samples may be converted in a way such that one of the four nucleotide types (Guanine, Adenine, Cytosine or Thymine) are replaced with the nucleotide's labeled counterpart or otherwise labeled by attaching a label to a respective nucleotide. The same label may be used for each nucleotide or optionally, different labels may be used. The remaining nucleotides are the naturally occurring nucleic acid building blocks. Optionally, two, three or each nucleotide of a nucleic acid may be replaced with a nucleotide carrying a distinct acceptor label.

[0091] To perform the sequence read-out where a single nucleotide label is utilized with the target nucleic acid split into four samples, each having one nucleotide labeled with the same, or optionally, a different acceptor label, a specially designed algorithm may be utilized which (i) corrects, (ii) defines, and (iii) aligns the four partial sequences into one master sequence. Each partial sequence displays the relative position of one of the four nucleotides in the context of the whole genome sequence, thus, four sequencing reactions may be required to determine the position of each nucleotide.

[0092] The algorithm may correct for missing bases due to inefficient labeling of the nucleic acid. One type of nucleotide in a DNA molecule can be completely substituted with the nucleotide's fluorescent counterpart. Various inefficiencies in labeling may result in less than 100% coverage from this substitution. Fluorescently labeled nucleotides usually come at a purity of around 99%, i.e., approximately 1% of the nucleotides do not carry a label. Consequently, even at a 100% incorporation of modified nucleotides, 1% of the nucleotides may be unlabeled and may not be detectable by nanopore transfer sequencing.

[0093] One solution to this problem is a redundant coverage of the nucleic acid to be sequenced. Each sequence may be read multiple times, e.g., at least 50 times per sequencing reaction (i.e. a 50 fold redundancy). Thus, the algorithm will compare the 50 sequences which will allow a statistically sound determination of each nucleotide call.

[0094] The algorithm may define the relative position of the sequenced nucleotides in the template nucleic acid. For example, the time of the current blockage during the translocation process may be used to determine the relative position of the detected nucleotides. The relative position and the time of the occurrence of two signals may be monitored and used to determine the position of the nucleotides relative to each other. Optionally, a combination of the above methods may be used to determine the position of the nucleotides in the sequence.

[0095] The nucleic acid or DNA to be analyzed may be separated into four samples. Each sample will be used to exchange one form of nucleotide (A, G, T, or C) with the nucleotide's fluorescent counterpart. Four separate nanopore sequencing reactions may reveal the relative positions of the four nucleotides in the DNA sample through optical detection. A computer algorithm will then align the four sub-sequences into one master sequence. The same acceptor label capable of emitting light at a specific wavelength or color may be utilized in all four samples. Optionally, different labels having different wavelength emissions may be utilized.

[0096] For example, Figure 4A shows partial contigs from nucleic acid sequencing utilizing a singly labeled nucleic acid. Four separate nanopore sequencing reactions take place. Each of the four separate nanopore sequencing reactions, which are created by the same type of nucleotide acceptor label, generates a sub-sequence that displays the relative position of one of the four nucleotides. A redundant coverage of each sequence may ensure statistical sound base calls and read-outs. A computer algorithm may be utilized to deduce the four partial contig sequences which are the common denominators of the multiple covered sub-sequences (i.e. G-contig, A-contig, T-contig, and C-contig).

[0097] Figure 4B shows how partial contig alignment may generate a first draft nucleic acid sequence. For example, the second bioinformatic step involves alignment of the four contigs. Software searches for matching sequence stretches of the four contigs that complement each other. This step results in a finished draft sequence.

[0098] Optionally, both optical and electrical read-outs/ detection may be utilized to sequence a nucleic acid. Electrical read-outs may be utilized to measure the number of non-labeled nucleotides in a sequence to help assess the relative position of a detected

labeled nucleotide on a nucleic acid sequence. The length of the nucleic acid can be calculated by measuring the change in current through the nanopore and the duration of that current change. The methods and systems described herein may utilize solely optical read-outs or optical detection of energy emission or light emission by a labeled monomer to identify and sequence the monomer and to sequence a polymer including the monomer. Optionally, a combination of optical and electrical readouts or detection may be used.

[0099] A nucleotide acceptor label may be in the form of a quencher which may quench the transferred energy. In the case of a quenching nucleotide label, radiation emission from the pore donor label will decrease when a labeled nucleotide is in proximity to the donor label. The detection system for quenching pore labels is designed to maximize the radiation collected from the pore labels, while filtering out laser excitation radiation. For a quenching label, a decrease in photon counts of the pore label, such as a quantum dot, is interpreted as a labeled nucleotide.

[0100] Figure 5A shows one variation of a quenching interaction between a pore donor label 56 on a synthetic pore or nanopore 52 and a nucleic acid quenching label 58 on a nucleic acid 57 (e.g., a single or double stranded nucleic acid), which is being translocated through the synthetic nanopore 52. The synthetic nanopore 52 is positioned in a substrate 54, e.g., a solid substrate.

[0101] During a continuous or substantially continuous illumination of the pore label 56, the pore label 56 emits light at a certain wavelength which is detected with an appropriate optical or other detection system. The quenching label 58 positioned on a nucleotide of nucleic acid 57 comes in close proximity to the pore label 56, e.g., as or after the label 58 or labeled nucleotide exits the nanopore 52, and thereby quenches the pore label 56. The quenching label 58 acts by absorption of energy from the illuminated pore label 56 (which is indicated by arrow B) causing the photon emission from the pore label 56 to change. For example, the quenching may be detected by detecting a change, such as a decrease or diminishing, in the photons emitted by the nanopore label. A degree of photon emission change may be associated with or correspond to a single nucleotide of the nucleic acid molecule and as such, the nucleic acid molecule sequence may be deduced based on detecting the change in photon emission by the donor label caused by the quenching label.

[0102] Figure 5B shows translocation of the labeled nucleic acid 57 at a point in time where no quenching is taking place (due to the labels not being in close enough proximity to each other). This is indicated by the lack of any arrows showing energy transfer between a pore label 56 and a nucleic acid label 58.

[0103] Figure 5C shows one variation of a quenching interaction between a pore donor label 66 on a proteinaceous or protein pore or nanopore 62 and a nucleic acid quenching label 68 on a nucleic acid 67 (e.g., a single or double stranded nucleic acid), which is being translocated through the protein nanopore 62. The protein nanopore 62 is positioned in a lipid bilayer 64.

[0104] During a continuous illumination of the pore label 66 the pore label 66 emits light at a certain wavelength which is detected with an appropriate optical or other detection system. The quenching label 68 positioned on a nucleotide of nucleic acid 67 comes in close proximity to the pore label 66, e.g., as or after the label 68 or labeled nucleotide exits the nanopore 62, and thereby quenches the pore label 66 (which is indicated by arrow B). This quenching is detected by a decrease or sharp decrease in measured photons emitted from the nanopore label.

[0105] Figure 5D shows translocation of the labeled nucleic acid 67 at a point in time where no quenching is taking place (due to the labels not being in close enough proximity to each other). This is indicated by the lack of any arrows showing energy transfer between a pore label 66 and a nucleic acid label 68.

[0106] The energy transfer reaction, energy emission or pore label quenching as described above may take place as or before the label or labeled nucleotide enters the nanopore, e.g., on the cis side of the nanopore.

[0107] The labeling system may be designed to emit energy continuously without intermittency or rapid photobleaching of the fluorophores. For example, the buffer compartment of a pore holder may contain an oxygen depletion system that will remove dissolved Oxygen from the system via enzymatical, chemical or electrochemical means thereby reducing photobleaching of the fluorophore labeled nucleic acid.

[0108] An oxygen depletion system is a buffer solution containing components that selectively react with dissolved oxygen. Removing oxygen from the sequencing buffer solution helps prevent photobleaching of the fluorophore labels. An example of a composition of an oxygen depletion buffer is as follows: 10 mM tris-Cl, pH 8.0, 50 mM NaCl, 10 mM MgCl₂, 1% (v/v) 2-mercaptoethanol, 4 mg/ml glucose, 0.1 mg/ml glucose oxidase, and 0.04 mg/ml catalase (Sabanayagam, C.R. et al., J. Chem. Phys. 123 (2005): 224708). The buffer is degassed by sonication before use to extend the buffer's useful lifetime by first removing oxygen mechanically. The buffer system then removes oxygen via the enzymatic oxidation of glucose by glucose oxidase.

[0109] The sequencing buffer may also contain components that prevent fluorescence intermittency, also referred to as "blinking," in one or both of the quantum dot labeled pores and fluorophore labeled nucleic acids. The phenomenon of blinking occurs when the excited fluorophore transitions to a non-radiative triplet state. Individual fluorophores may display fluorescence intermittency known as blinking in which the fluorophore transitions to and from the fluorophore's emitting and dark state. Blinking can interfere with certain aspects of the sequencing schemes. Blinking may be prevented or left alone. The triplet state is responsible for blinking in many organic fluorophores and that blinking can be suppressed with chemicals that quench the triplet state.

[0110] Molecules such as Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid) are effective in eliminating blinking for fluorophores or dyes such as Cy5 (Rasnik, I. et al., *Nat Methods* 11 (2006): 891-893). Certain Quantum dots may display blinking, however, CdTe quantum dots produced by aqueous synthesis in the presence of mercaptopropionic acid have recently been shown to emit continuously without blinking (He, H. et al., *Angew. Chem. Int. Ed.* 45 (2006): 7588 -7591). CdTe quantum dots are ideally suited as labels to be utilized in the methods described herein, since they are water soluble with high quantum yield and can be directly conjugated through the terminal carboxylic acid groups of the mercaptopropionic acid ligands.

[0111] The labels may be made resistant to photobleaching and blinking. With an efficient oxygen depletion system, Cy5 fluorophores can undergo $\sim 10^5$ cycles of excitation and emission before irreversible degradation. If the incident laser light is of high enough efficiency that excitation of the Cy5 fluorophore is saturated (re-excited immediately after emission) then the rate of photon emission is determined by the fluorescence lifetime of the Cy5 fluorophore. Since the Cy5 fluorophore has a lifetime on the order of ns, and an assumed FRET efficiency of 10%, up to 10,000 photons can be emitted as the Cy5 labeled nucleotide transverses the nanopore. Microscopes used for single molecule detection are typically around 3% efficient in light collection. This can provide ~ 300 photons detected for a given label, which provides sufficiently high signal to noise ratio for single base detection.

[0112] A polymer or nucleic acid may be translocated through a nanopore having a suitable diameter (the diameter may vary, e.g., the diameter may be about 2 to 6 nm) at an approx. speed of 1,000 to 100,000 or 1,000 to 10,000 nucleotides per second. Each base of the nucleic acid may be fluorescently labeled with a distinct fluorophore. The base of the nanopore may be labeled with a quantum dot. When the nucleotide label comes in close

proximity to the quantum dot, a non-radiative, quantum resonance energy transfer occurs which results in light emission of a specific wavelength from the nucleotide label.

[0113] Figure 6 shows an example of an absorption/emission spectra from a FRET pair containing a donor quantum dot and an acceptor fluorophore. The characteristic broad absorption peak (thin dashed line) of the quantum dot allows for a short excitation wavelength which doesn't interfere with the detection of the longer emission wavelength. The emission peak of the quantum dot (thin solid line) has a significant spectral overlap with the absorption peak of the acceptor fluorophore (thick dashed line). This overlap may result in an energy transfer from the quantum dot to the fluorophore which then emits photons of a specific wavelength (thick solid line). These fluorophore emitted photons are subsequently detected by an appropriate optical system. The efficiency of the energy transfer may be highly dependent on the distance between the donor and acceptor, with a 50% efficiency at the so called Foerster radius.

[0114] Sequencing may be performed by utilizing one or more pores or nanopores simultaneously. For example, a plurality of nanopores may be positioned in parallel or in any configuration in one or more lipid bilayers or substrates in order to expedite the sequencing process and sequence many nucleic acid molecules or other biological polymers at the same time.

[0115] A plurality of pores may be configured on a rotatable disc or substrate. When donor labels or quantum dots become substantially or completely used, burned out or exhausted (i.e., they reached their fluorescent lifetime), the disc or substrate may be rotated, thereby rotating a fresh pore with fresh donor labels or quantum dots into place to receive nucleic acids and continue sequencing. The electrical field which pulls the nucleic acid through the pore may be turned off during rotation of the disc and then turned back on once a new pore is in position for sequencing. Optionally, the electric field may be left on continuously.

[0116] Each of the individual variations described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other variations. Modifications may be made to adapt a particular situation, material, composition of matter, process, process act(s) or step(s) to the objective(s), spirit or scope of the present invention.

[0117] Methods recited herein may be carried out in any order of the recited events which is logically possible, as well as the recited order of events. Furthermore, where a range of values is provided, every intervening value between the upper and lower limit of

that range and any other stated or intervening value in that stated range is encompassed within the invention. Also, any optional feature of the inventive variations described may be set forth and claimed independently, or in combination with any one or more of the features described herein.

[0118] All existing subject matter mentioned herein (e.g., publications, patents, patent applications and hardware) is incorporated by reference herein in its entirety except insofar as the subject matter may conflict with that of the present invention (in which case what is present herein shall prevail). The referenced items are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such material by virtue of prior invention.

[0119] Reference to a singular item, includes the possibility that there are plural of the same items present. More specifically, as used herein and in the appended claims, the singular forms “a,” “an,” “said” and “the” include plural referents unless the context clearly dictates otherwise. It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as “solely,” “only” and the like in connection with the recitation of claim elements, or use of a “negative” limitation. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

[0120] This disclosure is not intended to be limited to the scope of the particular forms set forth, but is intended to cover alternatives, modifications, and equivalents of the variations described herein. Further, the scope of the disclosure fully encompasses other variations that may become obvious to those skilled in the art in view of this disclosure. The scope of the present invention is limited only by the appended claims.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method for sequencing a polymer comprising:
 - exciting a donor label attached to a nanopore wherein the donor label emits energy at a first wavelength;
 - translocating a polymer through the nanopore, wherein the polymer comprises a monomer labeled with an acceptor label;
 - transferring energy from the excited donor label to the acceptor label of the monomer after the labeled monomer passes through and exits the nanopore, wherein the energy transfer causes the acceptor label to emit energy at a second wavelength which is different from the first wavelength;
 - detecting the energy emitted by the acceptor label as a result of the energy transfer, wherein the energy emitted by the acceptor label is associated with a single monomer; and
 - deducing the polymer sequence based on the detection of the emitted energy from the acceptor label and identification of the monomer.
2. The method of claim 1, wherein the labeled polymer is threaded through the nanopore by applying an electric field through the nanopore.
3. The method of claim 1, wherein the donor label is a Quantum dot.
4. The method of claim 1, wherein the acceptor label is a fluorophore.
5. The method of claim 1, wherein a plurality of donor labels exhibiting different spectral emission maxima are attached to the nanopore and a plurality of acceptor labels having distinct absorption/emission spectra each associated with a specific monomer are attached to the polymer, and wherein each donor label and corresponding acceptor label undergo a FRET (Förster Resonance Energy Transfer) interaction such that each acceptor labeled monomer of the polymer emits a specific wavelength that is detected and used to sequence the polymer.

6. The method of claim 1, wherein detecting the energy emitted by the acceptor label comprises optical detection of a specific wavelength emitted by the acceptor label and associated with a specific nucleotide.
7. The method of claim 1, wherein the polymer is a nucleic acid and the monomer is a nucleotide.
8. The method of claim 7, wherein one of four nucleotide bases is labeled and the time dependent signal arising from a nucleotide label emission is converted into a nucleic acid sequence corresponding to positions of the labeled nucleotide in the nucleic acid sequence.
9. The method of claim 8, wherein the method is repeated for each of the four nucleotides in separate samples and four partial sequences are deduced and then aligned to assemble an entire nucleic acid sequence.
10. A system for sequencing a polymer comprising:
 - a nanopore membrane assembly having a donor labeled nanopore which spans a lipid bilayer membrane, the nanopore membrane assembly having a negatively charged side and a positively charged side;
 - an aqueous buffer solution surrounding the nanopore membrane assembly;
 - a negative electrode positioned on the negatively charged (cis) side of the nanopore membrane assembly;
 - a positive electrode positioned on the positively charged (trans) side of the nanopore membrane assembly;
 - a polymer comprising a monomer having an acceptor label, wherein the polymer is loaded into a port positioned on the negatively charged side of the nanopore membrane assembly;
 - an energy source for exciting the nanopore donor label, wherein the nanopore donor label is configured to emit energy at a first wavelength;
 - an electric field applied by the electrodes to force the labeled polymer to translocate through the nanopore from the negatively charged side to the positively charged side in a single file, wherein energy is transferred from the nanopore donor label

to the monomer acceptor label after the labeled monomer passes through and exits the nanopore and comes in proximity to the nanopore donor label, wherein the energy transfer causes the monomer acceptor label to emit energy at a second wavelength which is different from the first wavelength;

an optical detector for detecting the energy emitted from the monomer acceptor label as a result of the transfer of energy from the nanopore donor label to the monomer acceptor label, wherein the emitted energy is associated with a single monomer and facilitates the sequencing of the polymer.

11. The system of claim 10, wherein the donor label comprises one or more quantum dots and the acceptor label comprises one or more fluorophores.
12. The system of claim 10, wherein a plurality of nanopores are incorporated into the membrane in parallel, such that a plurality of polymers may be sequenced simultaneously.
13. The system of claim 10, wherein a plurality of nanopores are incorporated into a rotatable disc, such that the nanopores may be rotatably interchanged.
14. The method of claim 10, wherein the polymer is a nucleic acid and the monomer is a nucleotide.

QUANTAPORE, INC

WATERMARK PATENT AND TRADE MARKS ATTORNEYS

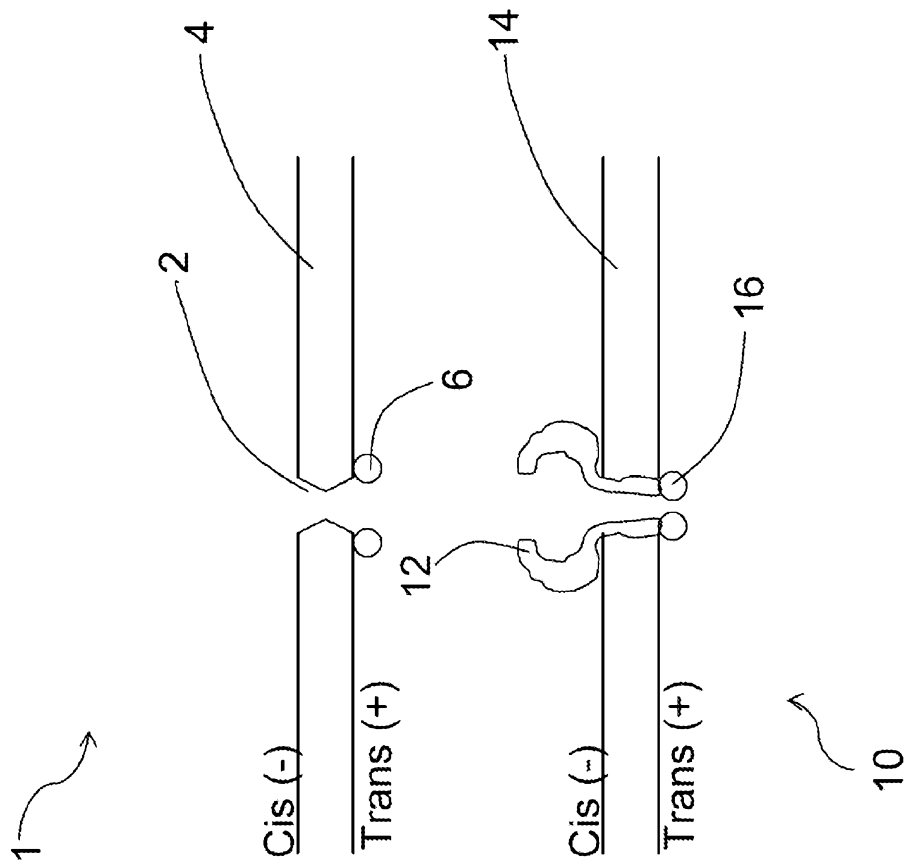


Figure 1A

Figure 1B

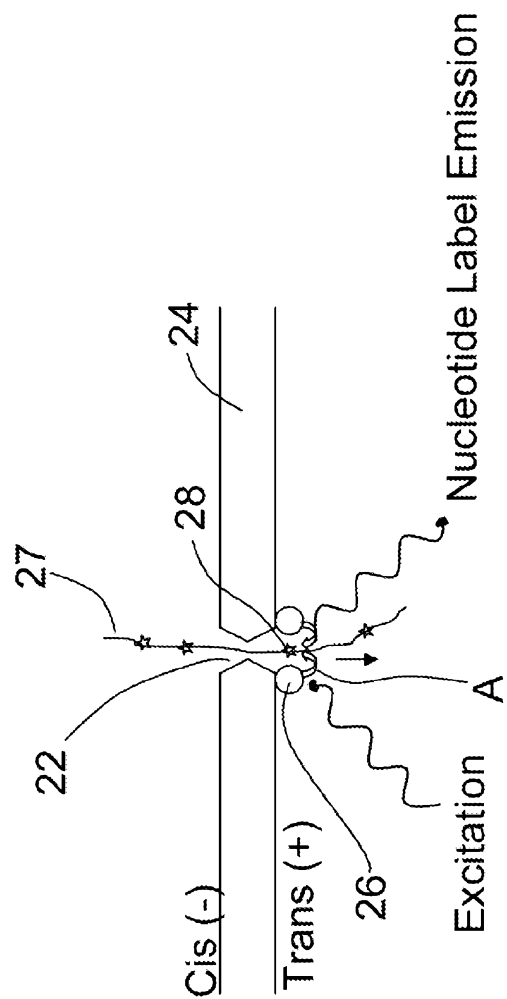


Figure 2A

FRET

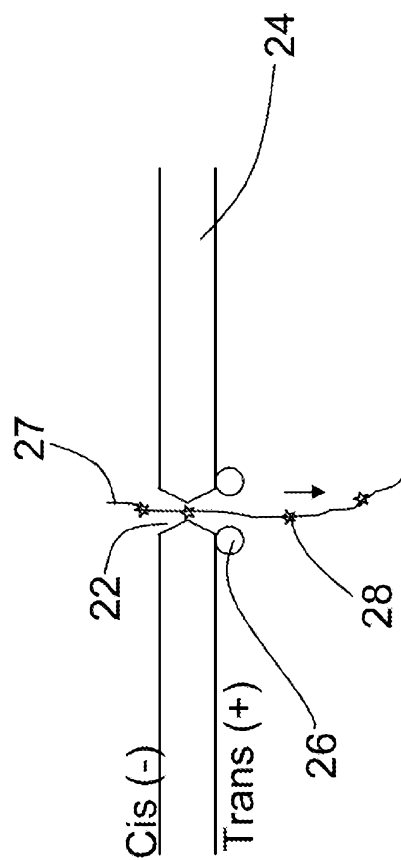
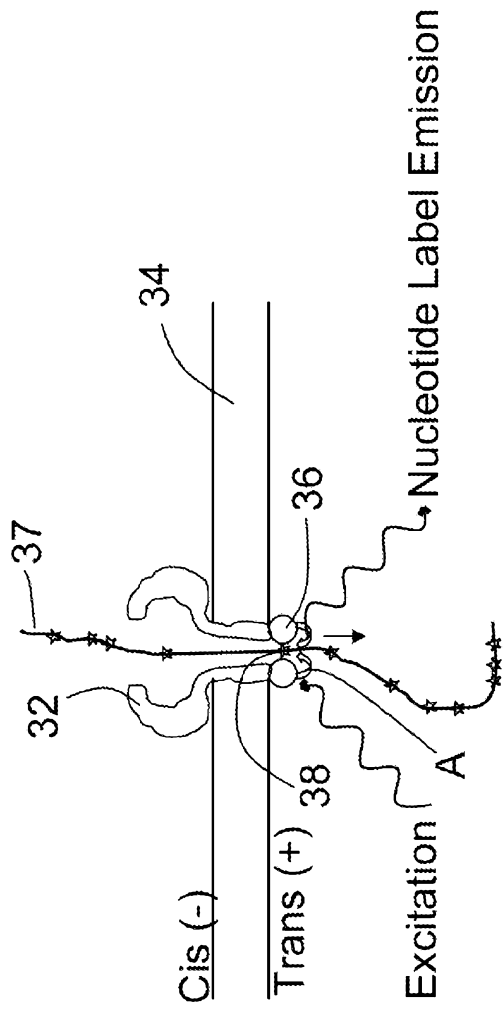


Figure 2B

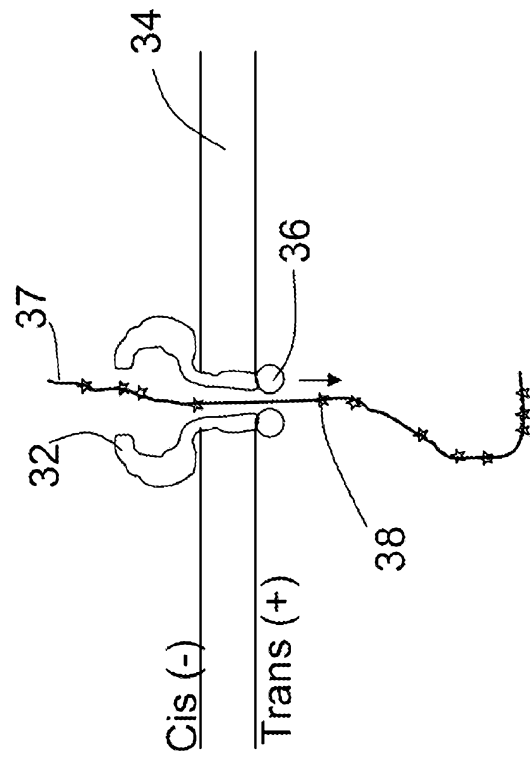
No FRET

Figure 2C



FRET

Figure 2D



No FRET

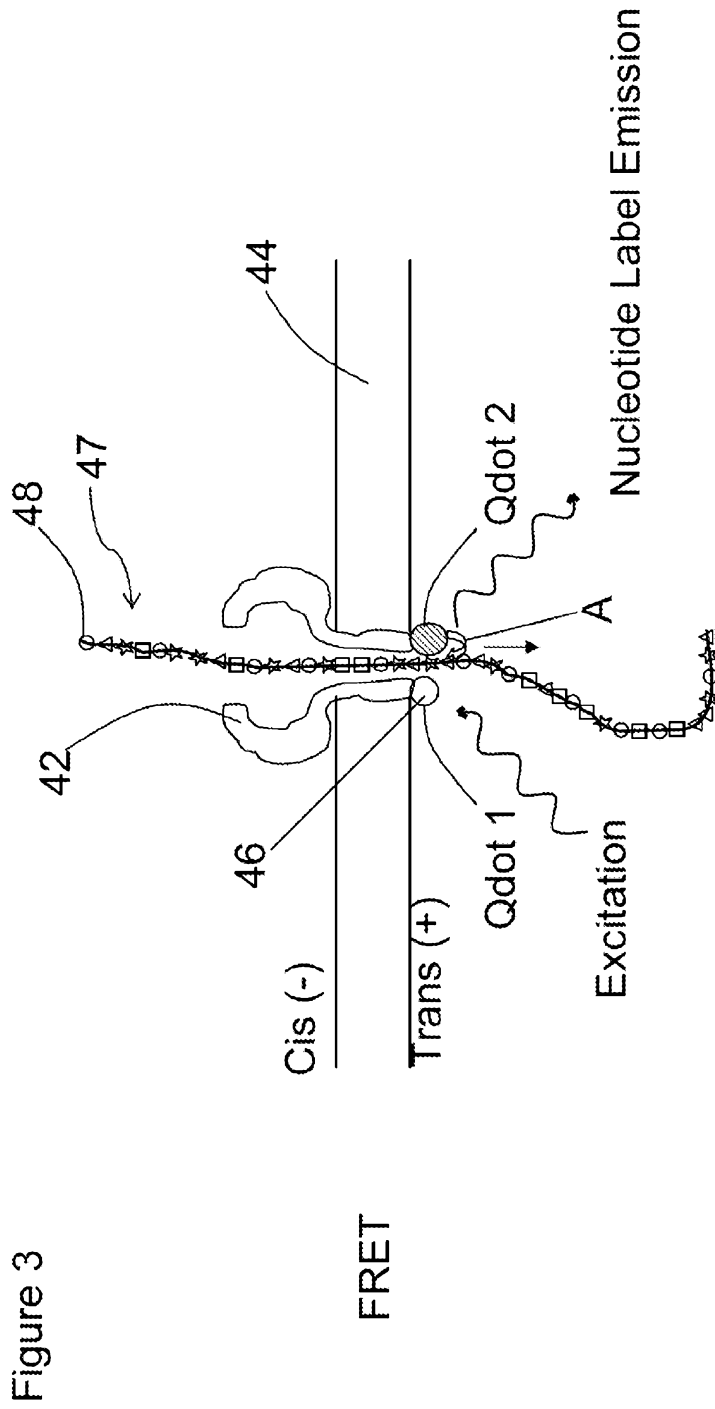


Figure 4A

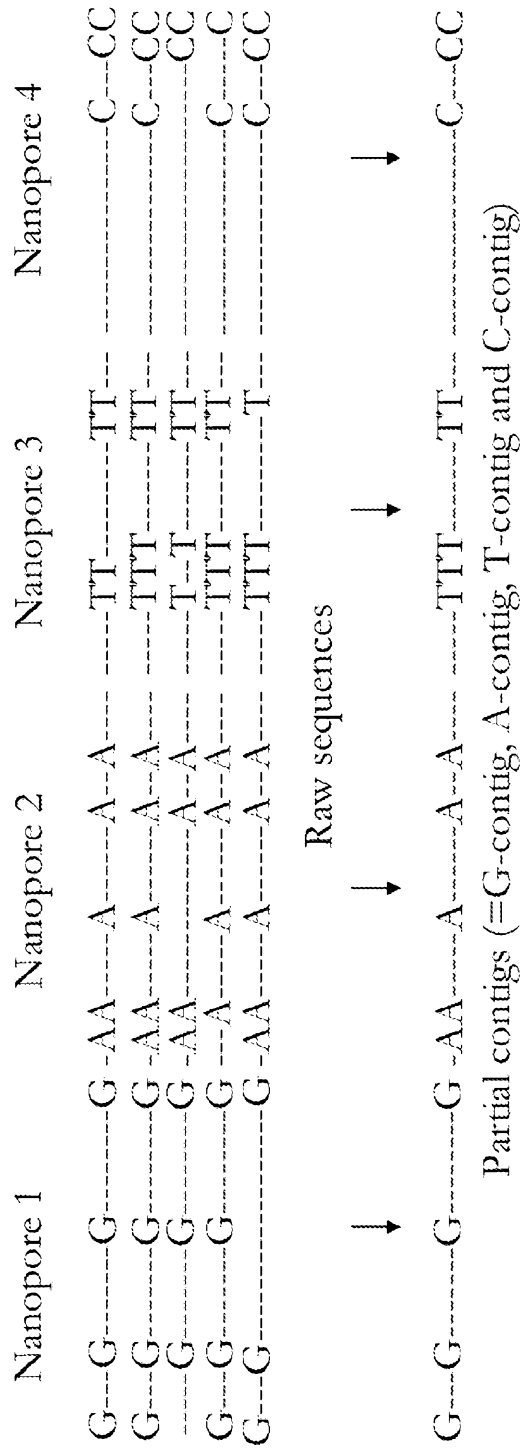
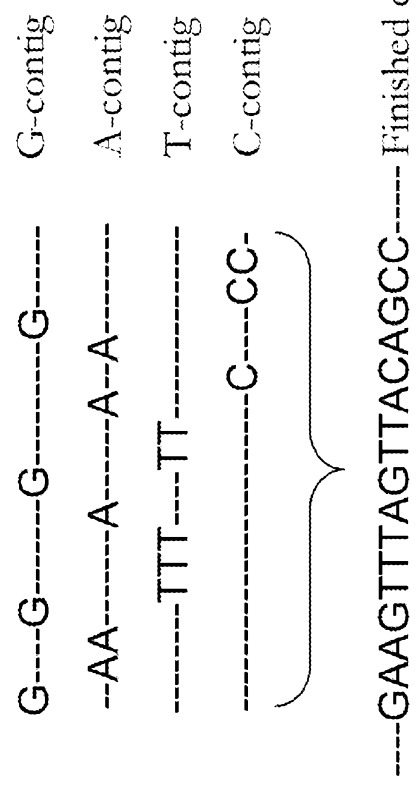


Figure 4B



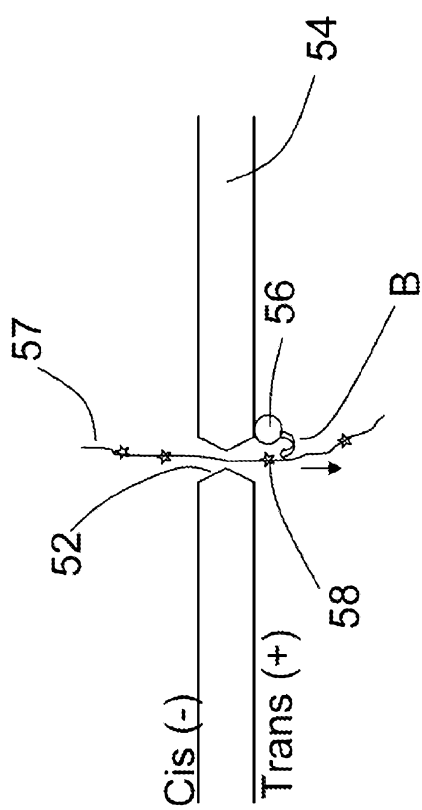


Figure 5A

Quenching

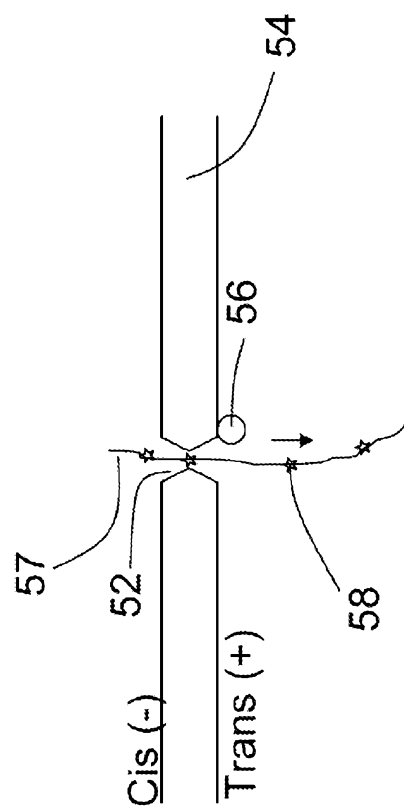


Figure 5B

No Quenching

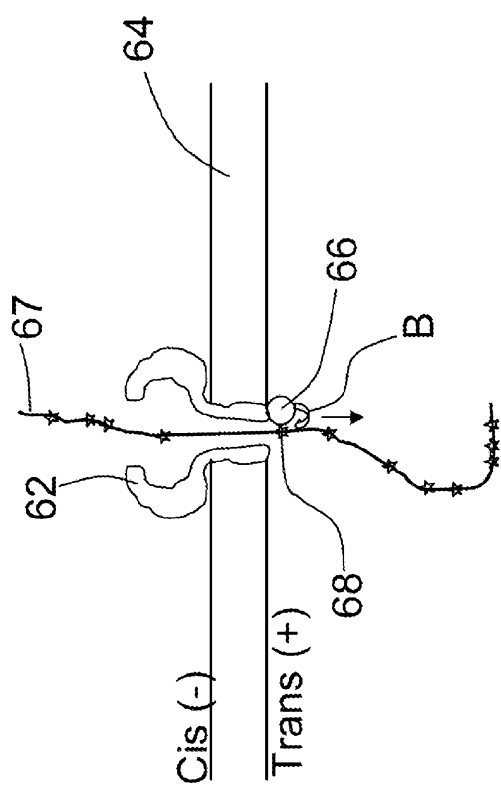


Figure 5C

Quenching

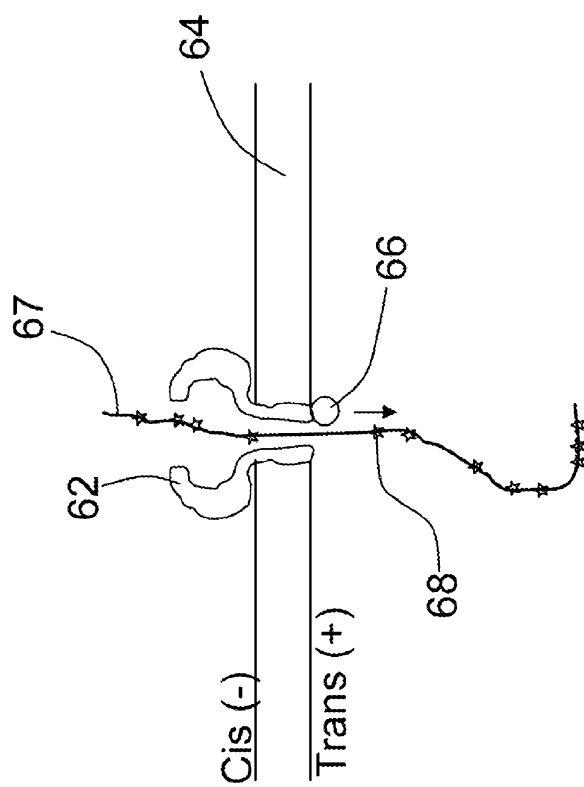


Figure 5D

No Quenching

Figure 6
Absorption/Emission spectra of a theoretical FRET pair

