

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2015330841 B2**

(54) Title
Nanopore-based polymer analysis with mutually-quenching fluorescent labels

(51) International Patent Classification(s)
C12Q 1/68 (2006.01)

(21) Application No: **2015330841**

(22) Date of Filing: **2015.10.08**

(87) WIPO No: **WO16/057829**

(30) Priority Data

(31) Number
62/062,256

(32) Date
2014.10.10

(33) Country
US

(43) Publication Date: **2016.04.14**

(44) Accepted Journal Date: **2019.08.15**

(71) Applicant(s)
Quantapore, Inc.

(72) Inventor(s)
Huber, Martin

(74) Agent / Attorney
Pizzseys Patent and Trade Mark Attorneys Pty Ltd, PO Box 291, WODEN, ACT, 2606, AU

(56) Related Art
GIBSON U E M ET AL, "A NOVEL METHOD FOR REAL TIME QUANTITATIVE RT-PCR", GENOME RESEARCH, COLD SPRING HARBOR LABORATORY PRESS, US, (1996-10-01), vol. 6, no. 10, ISSN 1088-9051, pages 995 - 1001

(51) International Patent Classification:
C12Q 1/68 (2006.01)(21) International Application Number:
PCT/US2015/054756(22) International Filing Date:
8 October 2015 (08.10.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/062,256 10 October 2014 (10.10.2014) US

(71) Applicant: QUANTAPORE, INC. [US/US]; 1455 Adams Court, Suite 1081, Menlo Park, CA 94025 (US).

(72) Inventor; and

(71) Applicant (for US only): HUBER, Martin [AT/US]; Quantapore, Inc., 1455 Adams Court, Suite 1081, Menlo Park, CA 94025 (US).

(74) Agent: MACEVICZ, Stephen, C.; Quantapore, Inc., 1455 Adams Court, Suite 1081, Menlo Park, CA 94025 (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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

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

— with sequence listing part of description (Rule 5.2(a))

(54) Title: NANOPORE-BASED POLYMER ANALYSIS WITH MUTUALLY-QUENCHING FLUORESCENT LABELS

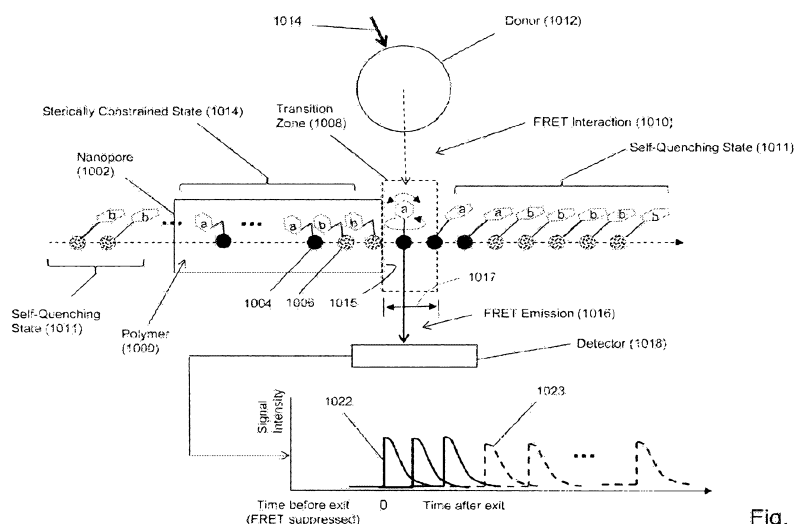


Fig. 1A

(57) Abstract: The invention is directed to a method for determining a monomer sequence of a polymer that is translocated through a nanopore. Monomers of the polymer are labeled with fluorescent labels such that in free solution fluorescent labels of adjacent monomers substantially quench each other and wherein the nanopore constrains fluorescent labels within its bore into a constrained state wherein no detectable fluorescent signal can be generated. By exciting the fluorescent label of each monomer as it exits the nanopore and transitions from a constrained state to a quenched state with an adjacent fluorescent label, a fluorescent signal can be generated by the exiting fluorescent label that allows its monomer to be identified, thereby permitting a monomer sequence to be determined from a sequence of fluorescent signals as the polymer translocates through the nanopore.

NANOPORE-BASED POLYMER ANALYSIS WITH MUTUALLY-QUENCHING FLUORESCENT LABELS

BACKGROUND

[0001] DNA sequencing technologies developed over the last decade have revolutionized the biological sciences, e.g. Lerner et al, *The Auk*, 127: 4-15 (2010); Metzker, *Nature Review Genetics*, 11: 31-46 (2010); Holt et al, *Genome Research*, 18: 839-846 (2008); and have the potential to revolutionize many aspects of medical practice, e.g. Voelkerding et al, *Clinical Chemistry*, 55: 641-658 (2009); Anderson et al, *Genes*, 1: 38-69 (2010); Freeman et al, *Genome Research*, 19: 1817-1824 (2009); Tucker et al, *Am. J. Human Genet.*, 85: 142-154 (2009). However, to realize such potential there are still a host of challenges that must be addressed, including reduction of per-run sequencing cost, simplification of sample preparation, reduction of run time, increasing read lengths, improving data analysis, and the like, e.g. Baker, *Nature Methods*, 7: 495-498 (2010); Kircher et al, *Bioessays*, 32: 524-536 (2010); Turner et al, *Annual Review of Genomics and Human Genetics*, 10: 263-284 (2009). Single molecule sequencing using nanopores may address some of these challenges, e.g., Maitra et al, *Electrophoresis*, 33: 3418-3428 (2012); Venkatesan et al, *Nature Nanotechnology*, 6: 615-624 (2011); however, this approach has its own set of technical difficulties, such as, reliable nanopore fabrication, control of DNA translocation rates, nucleotide discrimination, detection of electrical signals from large arrays of nanopore sensors, and the like, e.g. Branton et al, *Nature Biotechnology*, 26(10): 1146-1153 (2008); Venkatesan et al (cited above).

[0002] Optical detection of nucleotides has been proposed as a potential solution to some of the technical difficulties in the field of nanopore sequencing, e.g. Huber, U.S. patent 8,771,491; Russell, U.S. patent 6,528,258; Pittaro, U.S. patent publication 2005/0095599; Joyce, U.S. patent publication 2006/0019259; Chan, US. patent 6,355,420; McNally et al, *Nano Lett.*, 10(6): 2237-2244 (2010); and the like. However, optically-based nanopore sequencing has not been realized for a variety of reasons, including the lack of suitable fabrication techniques and understanding of how elements of such systems interact.

[0003] In view of the above, it would be advantageous to nanopore sensor technology in general and its particular applications, such as optically based nanopore sequencing, if

there were available materials and configurations of optical elements that permitted successful optical sensing and analysis of analytes, such as sequences of nucleic acids.

SUMMARY OF THE INVENTION

[0004] The present invention is directed to methods, kits and systems for optical detection and analysis of polymers, such as polynucleotides, in microfluidic and/or nanofluidic devices; in particular, the invention includes methods and systems using nanopores for determining nucleotide sequences of nucleic acids.

[0005] In one aspect, the invention include a method for determining a monomer sequence of a polymer comprising the following steps: (a) translocating a polymer through a nanopore, wherein monomers of the polymer are labeled with fluorescent labels such that in free solution fluorescent labels of adjacent monomers substantially quench each other's fluorescence emissions (that is, such labels are in a "quenched state" or "quenched configuration") and wherein the nanopore constrains fluorescent labels within its bore into a constrained state such that no detectable fluorescent signal, or substantially no detectable fluorescent signal, is generated; (b) exciting the fluorescent label of each monomer upon exiting the nanopore and prior to formation of a quenched configuration with an adjacent fluorescent label; (c) measuring a fluorescent signal generated by the exiting fluorescent label to identify the monomer to which the fluorescent label is attached; and (d) determining a monomer sequence of the polymer from a sequence of fluorescent signals.

[0006] In another aspect, the invention includes a method of determining a nucleotide sequence of a at least one polynucleotide comprising the steps of: (a) translocating a at least one single stranded polynucleotide through a nanopore, wherein nucleotides of the single stranded polynucleotide are labeled with fluorescent labels such that in free solution fluorescent labels of adjacent nucleotides are in a quenched state quenching fluorescence emissions of the parts of the polynucleotide outside the nanopore (that is, parts of the polynucleotide that have not yet entered or that have already exited the nanopore), and wherein the nanopore forces the fluorescent labels within the nanopore into a constrained state wherein substantially no detectable signal is generated; (b) exciting the fluorescent label of each nucleotide upon exiting the nanopore and prior to forming a quenched state with an fluorescent label of an adjacent nucleotide; (c) measuring a fluorescent signal generated by the exiting fluorescent label to identify the nucleotide to which the fluorescent label is attached; and (d) determining a nucleotide sequence of the polynucleotide from a

sequence of fluorescent signals. In some embodiments of this aspect, nucleotides of the polynucleotides are labeled with second members of a FRET pair, each second member producing a FRET signal indicative of the nucleotide to which it is attached, so that nucleotides of the polynucleotide pass in sequence by a first member of the FRET pair positioned adjacent to the exit of the nanopore so that each second member upon exiting the nanopore passes within a FRET distance of the first member of the FRET pair.

[0007] In some embodiments, nanopores are fabricated in a solid phase membrane such that first members of a FRET pair are attached to the solid phase membrane adjacent to substantially each nanopore. In other embodiments, nanopores comprise protein nanopores disposed in apertures fabricated in a solid phase membrane wherein first members of a FRET pair are attached to the protein nanopore.

[0008] The present invention is exemplified in a number of implementations and applications, some of which are summarized below and throughout the specification.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] Fig. 1A illustrates schematically an exemplary embodiment of the invention.

[0010] Fig. 1B illustrates expected signals for different times to self-quenching after a fluorescent label or acceptor exits a nanopore.

[0011] Fig. 1C illustrates expected signals for specified time to self-quenching and compares to recorded signal from a labeled target polynucleotide translocating a nanopore.

[0012] Figs. 2A-2C illustrate one embodiment of a hybrid biosensor.

[0013] Fig. 2D illustrate an embodiment of the device of the invention with positioning of a member of a FRET pair using oligonucleotide hybridization.

[0014] Fig. 2E illustrates one embodiment of a hybrid nanopore where the surface of the solid state membrane (201) coated with a hydrophobic layer (202) to which a lipid layer is adhered (203). The lipids form a gigaohm seal with the inserted pore protein.

[0015] Figs. 3A-3H show reaction diagrams of various orthogonal linking chemistries for attaching fluorescent labels to bases of polynucleotides.

DETAILED DESCRIPTION OF THE INVENTION

[0016] While the invention is amenable to various modifications and alternative forms, specifics thereof have been shown by way of example in the drawings and will be described in detail. It should be understood, however, that the intention is not to limit the invention to the particular embodiments described. On the contrary, the intention is to cover all modifications, equivalents, and alternatives falling within the spirit and scope of the invention. For example, particular nanopore types and numbers, particular labels, FRET pairs, detection schemes, fabrication approaches of the invention are shown for purposes of illustration. It should be appreciated, however, that the disclosure is not intended to be limiting in this respect, as other types of nanopores, arrays of nanopores, and other fabrication technologies may be utilized to implement various aspects of the systems discussed herein. Guidance for aspects of the invention is found in many available references and treatises well known to those with ordinary skill in the art, including, for example, Cao, Nanostructures & Nanomaterials (Imperial College Press, 2004); Levinson, Principles of Lithography, Second Edition (SPIE Press, 2005); Doering and Nishi, Editors, Handbook of Semiconductor Manufacturing Technology, Second Edition (CRC Press, 2007); Sawyer et al, Electrochemistry for Chemists, 2nd edition (Wiley Interscience, 1995); Bard and Faulkner, Electrochemical Methods: Fundamentals and Applications, 2nd edition (Wiley, 2000); Lakowicz, Principles of Fluorescence Spectroscopy, 3rd edition (Springer, 2006); Hermanson, Bioconjugate Techniques, Second Edition (Academic Press, 2008); and the like, which relevant parts are hereby incorporated by reference.

[0017] In one aspect, the invention relates to the use of nanopores, fluorescent quenching, and fluorescent signaling to sequentially identify monomers of polymer analytes. Such analysis of polymer analytes may be carried out on single polymer analytes or on pluralities of polymer analytes in parallel at the same time, for example, by using an array of nanopores. In some embodiments, monomers are labeled with fluorescent labels that are capable of at least three states while attached to a target polymer: (i) A substantially quenched state wherein fluorescence of an attached fluorescent label is quenched by a fluorescent label on an immediately adjacent monomer; for example, a fluorescent label attached to a polymer in accordance with the invention is substantially quenched when the labeled polymer is free in conventional aqueous solution for studying and manipulating the polymer. (ii) A sterically constrained state wherein a labeled polymer is translocating through a nanopore such that the free-solution movements or

alignments of an attached fluorescent label is disrupted or limited so that there is little or no detectable fluorescent signal generated from the fluorescent label. (iii) A transition state wherein a fluorescent label attached to a polymer transitions from the sterically constrained state to the quenched state as the fluorescent label exits the nanopore (during a “transition interval”) while the polymer translocates through the nanopore.

[0018] In part, the invention is an application of the discovery that during the transition interval a fluorescent label (on an otherwise substantially fully labeled and self-quenched polymer) is capable of generating a detectable fluorescent signal. Without the intention of being limited by any theory underlying this discovery, it is believed that the fluorescent signal generated during the transition interval is due to the presence of a freely rotatable dipole in the fluorescent label emerging from the nanopore, which renders the fluorescent label temporarily capable of generating a fluorescent signal, for example, after direct excitation or via FRET. In both the sterically constrained state as well as the quenched state, the dipoles are limited in their rotational freedom thereby reducing or limiting the number of emitted photons. In some embodiments, the polymer is a polynucleotide, usually a single stranded polynucleotide, such as, DNA or RNA, but especially single stranded DNA. In some embodiments, the invention includes a method for determining a nucleotide sequence of a polynucleotide by recording signals generated by attached fluorescent labels as they exit a nanopore one at a time as a polynucleotide translocates through the nanopore. Upon exit, each attached fluorescent label transitions during a transition interval from a constrained state in the nanopore to a quenched state on the polynucleotide in free solution. In other words, in some embodiments, a step of the method of the invention comprises exciting each fluorescent label as it is transitioning from a constrained state in the nanopore to a quenched state on the polymer in free solution. As mentioned above, during this transition interval or period the fluorescent label is capable of emitting a detectable fluorescent signal indicative of the nucleotide it is attached to.

[0019] In some embodiments, the invention is an application of the discovery that fluorescent labels and nanopores may be selected so that during translocation of a polymer through a nanopore fluorescent labels attached to monomers are forced into a constrained state in which they are incapable (or substantially incapable) of producing a detectable fluorescent signal. In some embodiments, nanopores are selected that have a bore, or lumen, with a diameter in the range of from 1 to 4 nm; in other embodiments, nanopores are selected that have a bore or lumen with a diameter in the range of from 2 to 3 nm. In

some embodiments, such bore diameters are provided by a protein nanopore. In some embodiments, such nanopores are used to force fluorescent labels into a constrained state in accordance with the invention, so that whenever a fluorescent label exits a nanopore, it transitions from being substantially incapable of generating a fluorescent signal to being detectable and identifiable by a fluorescent signal it can be induced to emit. Thus, fluorescent labels attached to each of a sequence of monomers of a polymer may be detected in sequence as they suddenly generate a fluorescent signal in a region immediately adjacent to a nanopore exit (a “transition zone” or “transition volume”). In some embodiments, organic fluorescent dyes are used as fluorescent labels with nanopores of the above diameters. In some embodiments, at least one such organic fluorescent dye is selected from the set consisting of xanthene dyes, rhodamine dyes and cyanine dyes. Some embodiments for determining a monomer sequence of a polymer may be carried out with the following steps: (a) translocating a polymer through a nanopore, wherein monomers of the polymer are labeled with fluorescent labels wherein the nanopore constrains fluorescent labels within its bore into a constrained state such that substantially no detectable fluorescent signal is generated therein; (b) exciting the fluorescent label of each monomer upon exiting the nanopore; (c) measuring a fluorescent signal in a transition zone generated by the exiting fluorescent label to identify the monomer to which the fluorescent label is attached; and (d) determining a monomer sequence of the polymer from a sequence of fluorescent signals. In further embodiments, fluorescent labels are acceptors of a FRET pair and one or more donors of the FRET pair are attached to the nanopore within a FRET distance of the exit.

[0020] In some embodiments, “substantially quenched” as used above means a fluorescent label generates a fluorescent signal at least thirty percent reduced from a signal generated under the same conditions, but without adjacent mutually quenching labels. In some embodiments, “substantially quenched” as used above means a fluorescent label generates a fluorescent signal at least fifty percent reduced from a signal generated under the same conditions, but without adjacent mutually quenching labels.

[0021] In some embodiments, a nucleotide sequence of a target polynucleotide is determined by carrying out four separate reactions in which copies of the target polynucleotide have each of its four different kinds of nucleotide (A, C, G and T) labeled with a single fluorescent label. In a variant of such embodiments, a nucleotide sequence of a target polynucleotide is determined by carrying out four separate reactions in which

copies of the target polynucleotide have each of its four different kinds of nucleotide (A, C, G and T) labeled with one fluorescent label while at the same time the other nucleotides on the same target polynucleotide are labeled with a second fluorescent label. For example, if a first fluorescent label is attached to A's of the target polynucleotide in a first reaction, then a second fluorescent label is attached to C's, G's and T's (i.e. to the "not-A" nucleotides) of the target polynucleotides in the first reaction. Likewise, in continuance of the example, in a second reaction, the first label is attached to C's of the target polynucleotide and the second fluorescent label is attached to A's, G's and T's (i.e. to the "not-C" nucleotides) of the target polynucleotide. And so on, for nucleotides G and T.

[0022] The same labeling scheme may be expressed in terms of conventional terminology for subsets of nucleotide types; thus, in the above example, in a first reaction, a first fluorescent label is attached to A's and a second fluorescent label is attached to B's; in a second reaction, a first fluorescent label is attached to C's and a second fluorescent label is attached to D's; in a third reaction, a first fluorescent label is attached to G's and a second fluorescent label is attached to H's; and in a fourth reaction, a first fluorescent label is attached to T's and a second fluorescent label is attached to V's.

[0023] In some embodiments, a polymer, such as a polynucleotide or peptide, may be labeled with a single fluorescent label attached to a single kind of monomer, for example, every T (or substantially every T) of a polynucleotide is labeled with a fluorescent label, e.g. a cyanine dye. In such embodiments, a collection, or sequence, of fluorescent signals from the polymer may form a signature or fingerprint for the particular polymer. In some such embodiments, such fingerprints may or may not provide enough information for a sequence of monomers to be determined.

[0024] In some embodiments, a feature of the invention is the labeling of substantially all monomers of a polymer analyte with fluorescent dyes or labels that are members of a mutually quenching set. The use of the term "substantially all" in reference to labeling polymer analytes is to acknowledge that chemical and enzymatic labeling techniques are typically less than 100 percent efficient. In some embodiments, "substantially all" means at least 80 percent of all monomer have fluorescent labels attached. In other embodiments, "substantially all" means at least 90 percent of all monomer have fluorescent labels attached. In other embodiments, "substantially all" means at least 95 percent of all monomer have fluorescent labels attached. Mutually quenching sets of fluorescent dyes

have the following properties: (i) each member quenches fluorescence of every member (for example, by FRET or by static or contact mechanisms), and (ii) each member generates a distinct fluorescent signal when excited and when in a non-quenched state. That is, if a mutually quenching set consists of two dyes, D1 and D2, then (i) D1 is self-quenched (e.g. by contact quenching with another D1 molecule) and it is quenched by D2 (e.g. by contact quenching) and (ii) D2 is self-quenched (e.g. by contact quenching with another D2 molecule) and it is quenched by D1 (e.g. by contact quenching). Guidance for selecting fluorescent dyes or labels for mutually quenching sets may be found in the following references, which are incorporated herein by reference: Johansson, *Methods in Molecular Biology*, 335: 17-29 (2006); Marras et al, *Nucleic Acids Research*, 30: e122 (2002); and the like. In some embodiments, members of a mutually quenching set comprise organic fluorescent dyes that components or moieties capable of stacking interactions, such as aromatic ring structures. Exemplary mutually quenching sets of fluorescent dyes, or labels, may be selected from rhodamine dyes, fluorescein dyes and cyanine dyes. In one embodiment, a mutually quenching set may comprise the rhodamine dye, TAMRA, and the fluorescein dye, FAM. In another embodiment, mutually quenching sets of fluorescent dyes may be formed by selecting two or more dyes from the group consisting of Oregon Green 488, Fluorescein-EX, fluorescein isothiocyanate, Rhodamine Red-X, Lissamine rhodamine B, Calcein, Fluorescein, Rhodamine, one or more BODIPY dyes, Texas Red, Oregon Green 514, and one or more Alexa Fluors. Representative BODIPY dyes include BODIPY FL, BODIPY R6G, BODIPY TMR, BODIPY 581/591, BODIPY TR, BODIPY 630/650 and BODIPY 650/665. Representative Alexa Fluors include Alexa Fluor 350, 405, 430, 488, 500, 514, 532, 546, 555, 568, 594, 610, 633, 635, 647, 660, 680, 700, 750 and 790.

[0025] As above, in some embodiments, a monomer sequence of a target polymer is determined by carrying out separate reactions (one for each kind of monomer) in which copies of the target polymer have each different kind of monomer labeled with a mutually- or self-quenching fluorescent label. In other embodiments, a monomer sequence of a target polymer is determined by carrying out separate reactions (one for each kind of monomer) in which copies of the target polymer have each different kind of monomer labeled with a different mutually quenching fluorescent label selected from the same mutually quenching set. In embodiments in which a mutually quenching set contains only two dyes, then a selected monomer (say, monomer X) is labeled with a first mutually quenching dye and

every other kind of monomer (i.e., not-monomer X) is labeled with a second mutually quenching dye from the same set. Thus, steps of the embodiment generate a sequence of two different fluorescent signals, one indicating monomer X and another indicating not-monomer X.

[0026] In some embodiments, a single fluorescent label (for example, attached to a single kind of monomer in a polymer comprising multiple kinds of monomers) may be used that is self-quenching when attached to adjacent monomers (of the same kind) on a polymer, such as adjacent nucleotides of a polynucleotide. Exemplary self-quenching fluorescent labels include, but are not limited to, Oregon Green 488, fluorescein-EX, FITC, Rhodamine Red-X, Lissamine rhodamine B, calcein, fluorescein, rhodamine, BODIPYS, and Texas Red, e.g. which are disclosed in Molecular Probes Handbook, 11th Edition (2010).

[0027] In some embodiments, fluorescent labels are members of a FRET pair. A FRET pair generally is one or more FRET donors and one or more FRET acceptors where each donor is capable of a FRET reaction with each acceptor. In one aspect, this means that the donors of the FRET pair have an emission spectrum that substantially overlaps the absorption spectrum of the acceptors. In another aspect, the transition dipole of the donor and the acceptor have to be aligned in a way that allows efficient energy transfer. In some aspects, the invention in part is based on the discovery and appreciation of a fluorescence, particularly, FRET suppressing property of nanopores and the application of this property to enable detection of labeled polymers translocating through a nanopore. It is believed, although the invention is not intended to be limited thereby, that a nanopore may be selected with a bore dimensioned so that a FRET pair label cannot orient to engage in a FRET interaction while translocating through the nanopore. The dipoles of the labels of the polynucleotide in the bore of the nanopore are constrained in their rotational freedom based on the limited diameter of the nanopore. This reduction in dipole alignment with the alignment of the corresponding FRET pair attached to the nanopore limits the FRET efficiency dramatically. Labeled polynucleotides can engage in a FRET interaction after exiting the nanopore at which point the FRET acceptor or donor on the polymer (e.g. polynucleotide) regains rotational freedom which allows for a FRET event.

[0028] The invention may have a wide range of embodiments depending on the type of analytes being detected, the types of donors and acceptors employed, the physical

arrangement of the nanopore, donors and acceptors, whether analytes are labeled with donors or with acceptors, and the like. In some embodiments, analytes measured by the invention are acceptor-labeled polymers, especially acceptor-labeled polynucleotides. In one species of the latter embodiment, different nucleotides of a polynucleotide analyte are labeled with one or more different kinds of acceptors, so that a nucleotide sequence of the polynucleotide may be determined from measuring FRET signals generated as it translocates through a nanopore. In another embodiment, analytes measured by the invention are donor-labeled polymers, especially donor-labeled polynucleotides. The sequence of the polynucleotide may be determined from measuring FRET signals as it translocates through a nanopore. In yet another embodiment of the present invention, at least one of the four nucleotides of a polynucleotide analyte is labeled with a member of a FRET pair. The positions of the labeled nucleotides in the polynucleotide are determined by translocating the labeled polynucleotide through a labeled nanopore and measuring FRET events. By labeling the remaining nucleotides of the same polynucleotide sample and subsequently translocating said samples through a labeled nanopore, sub-sequences of the polynucleotide are generated. Such sub-sequences can be aligned resulting in a full sequence of the polynucleotide.

[0029] Some of the above aspects and embodiments of the invention are illustrated diagrammatically in Fig. 1A. Polymer analyte (1000), such as a polynucleotide, is driven, e.g. electrophoretically, through nanopore (1002), which constrains the conformation of polymer (1000) so that its monomeric units translocate through the nanopore in the same order as their primary sequence in the polymer. In the embodiment shown in Fig. 1A, fluorescent labels are assumed to be members of FRET pairs, but this is not intended to limit the present invention; fluorescent labels may also include fluorescent labels that are directly excited, for example with a laser emitting at an appropriate wavelength, to generate a fluorescent signal.

[0030] As mentioned above, whenever an acceptor-labeled monomeric unit is within the bore of nanopore (1002), FRET interactions between such acceptors and the donors of its FRET pair are suppressed because acceptors are in a constrained state (1014). Such suppression typically means that no detectable FRET signal is produced even if such acceptors are within a FRET distance of a donor, for example, due to unfavorable orientation of the acceptor and donor dipoles, or due to contact quenching, or like mechanism. On the other hand, when an acceptor-labeled monomeric unit emerges from

the bore of, or exits, the nanopore into transition zone (1008), FRET interaction (1010) occurs and FRET emission (1016) is produced and detected by detector (1018) until the acceptor enters a self-quenching state (1011) with an adjacent acceptor and as the distance between the acceptor and donor increases with the movement of polymer (1000) out of FRET interaction distance. Signal (1022) is produced by a single acceptor as it moves through transition zone (1008). Transition zone (1008), which is a spatial region immediately adjacent to exit (1015) of nanopore (1002), is defined by several factors, including the speed of the translocation of polymer (1000) through nanopore (1002), the vibrational and rotational mobility of the fluorescent labels, the physiochemical nature of the fluorescent labels, and the like. In some embodiments, transition zone (1008) may be defined by a perpendicular distance (1017) between the exit (1015) of nanopore (1002) and the point at which an exiting fluorescent label takes on a quenched configuration with an adjacent fluorescent label. In some embodiments, transition zone (1008) may be defined by its corresponding transition interval, or the time it takes a fluorescent label to travel distance (1017). In some embodiments, transition distance (1017) is in the range of from 20 to 50 angstroms; in other embodiments, transition distance is in the range of from 20 to 40 angstroms. In some embodiments, corresponding transition intervals are in the range of from 0.2 to 2.0 msec; in still other embodiments, transition intervals are in the range of from 0.2 to 1.0 msec. In Fig. 1A, only one type of monomeric unit, illustrated as solid circles (1004) carries a first fluorescent label (designated as “a”); the rest of the monomeric units, illustrated as speckled circles (1006), carry a second fluorescent label (designated as “b”). In this embodiment, first fluorescent labels quench adjacent first fluorescent labels and adjacent second fluorescent labels; likewise, second fluorescent labels quench adjacent first fluorescent labels and adjacent second fluorescent labels; moreover, the first and second fluorescent labels generate FRET signals that are distinguishable from one another, for example, recorded signal (1022) for label “a” and recorded signal (1023) for label “b” in Fig. 1A, so that each fluorescent label (and hence, monomer) may be identified by a signal detected by detector (1018).

[0031] As illustrated in Fig. 1B, the degree to which successive signals (1022) or (1023) are resolved by detector (1018) depend at least in part on the translocation speed of polymer (1000). Curve A and curve B of Fig. 1B illustrate results from simulations of fluorescent signal generation based on the Forrester equation under different auto-quenching conditions. As illustrated, under both conditions readily discernable signals are

generated. In Fig. 1C, a further simulation showing signal peaks (1033) is compared to actual data (1035) generated as a fluorescently labeled single stranded DNA analyte translocated through a nanopore. The single stranded DNA used to generate the data (1035) was 200nt long and each cytosine was exchanged with a fluorescently labeled counterpart. The labeled DNA was translocated through a continuously excited hybrid nanopore at an applied potential of 300mV and FRET events were captured using a cmos camera operated at 2kHz acquisition rate. At the 3' end of the labeled DNA a short homopolymer stretch of 3 consecutive cytosines shows an elevated baseline fluorescent with clearly distinguishable peaks for each of the three cytosines. Similar to the modeled data the fluorescent trace in the inset of Fig. 1C shows an elevated baseline fluorescence and individual peaks for each member of the homopolymer stretch. The sequence of labeled DNA is as follows (SEQ ID NO: 1):

5'- GCT ATG TGG CGC GGT ATT ATC AAG AAG GAG ACT GAG AGG AGA
GTA GGA GCG AGA AGG AAA CGA GAG TGA GAG GAG AGT AGG AGC
AAG AAG GAA ACG AGA GTG AGA GGA GAG TAG GAG CAA GAA GGA
AAC GAG AGT GAG AGG AGA GTA GGA GCA AGA AGG AAA CTG AGA
GGA GAG TAG GAG TTA CTC TAG CTT CCC GGC AA -3'

[0032] In some embodiments, a nanopore is hybrid nanopore comprising a protein nanopore inserted into a pore of a solid phase membrane, as described more fully below. In hybrid nanopores, a first member of a FRET pair may be attached directly to the protein nanopore, or alternatively, directly to the solid phase membrane using conventional linking chemistries, such as “click” chemistries, e.g. Kolb et al, Angew. Chem. Int. Ed., 4): 2004-2021 (2001), or the like. In one embodiment, a first member of a FRET pair is attached directly or indirectly to the protein nanopore, for example, as discussed in reference to Fig. 2D. In another embodiment, the first member of the FRET pair is a donor and a quantum dot. Quantum dots are typically much larger than acceptors, especially acceptors that are organic dyes, which typically have molecular weights in the range of from 200 to 2000 daltons.

Nanopores and Nanopore Sequencing

[0033] Nanopores used with the invention may be solid-state nanopores, protein nanopores, or hybrid nanopores comprising protein nanopores or organic nanotubes such as carbon nanotubes, configured in a solid-state membrane, or like framework. Important features of nanopores include (i) constraining analytes, particularly polymer analytes, to pass through a detection zone in sequence, or in other words, so that monomers pass the detection zone one at a time, or in single file, (ii) compatibility with a translocating means, that is, whatever method is used to drive an analyte through a nanopore, such as an electric field, and (iii) suppression of fluorescent signals within the lumen, or bore, of the nanopore, for example, by contact quenching, or the like. Nanopores used in connection with the methods and devices of the invention may be used singly or in the form of arrays, either a regular array, such as a rectilinear array of a plurality nanopores in a planar support or membrane, or a random array, for example, where a plurality of nanopores are spaced in accordance with a Poisson distribution in a planar support or membrane.

[0034] Nanopores may be fabricated in a variety of materials including but not limited to, silicon nitride (Si_3N_4), silicon dioxide (SiO_2), and the like. The fabrication and operation of nanopores for analytical applications, such as DNA sequencing, are disclosed in the following exemplary references that are incorporated by reference: Russell, U.S. patent 6,528,258; Feier, U.S. patent 4,161,690; Ling, U.S. patent 7,678,562; Hu et al, U.S. patent 7,397,232; Golovchenko et al, U.S. patent 6,464,842; Chu et al, U.S. patent 5,798,042; Sauer et al, U.S. patent 7,001,792; Su et al, U.S. patent 7,744,816; Church et al, U.S. patent 5,795,782; Bayley et al, U.S. patent 6,426,231; Akesson et al, U.S. patent 7,189,503; Bayley et al, U.S. patent 6,916,665; Akesson et al, U.S. patent 6,267,872; Meller et al, U.S. patent publication 2009/0029477; Howorka et al, International patent publication WO2009/007743; Brown et al, International patent publication WO2011/067559; Meller et al, International patent publication WO2009/020682; Polonsky et al, International patent publication WO2008/092760; Van der Zaag et al, International patent publication WO2010/007537; Yan et al, Nano Letters, 5(6): 1129-1134 (2005); Iqbal et al, Nature Nanotechnology, 2: 243-248 (2007); Wanunu et al, Nano Letters, 7(6): 1580-1585 (2007); Dekker, Nature Nanotechnology, 2: 209-215 (2007); Storm et al, Nature Materials, 2: 537-540 (2003); Wu et al, Electrophoresis, 29(13): 2754-2759 (2008); Nakane et al, Electrophoresis, 23: 2592-2601 (2002); Zhe et al, J. Micromech. Microeng., 17: 304-313 (2007); Henriquez et al, The Analyst, 129: 478-482 (2004); Jagtiani et al, J. Micromech. Microeng., 16: 1530-1539 (2006); Nakane et al, J. Phys. Condens. Matter, 15 R1365-

R1393 (2003); DeBlois et al, Rev. Sci. Instruments, 41(7): 909-916 (1970); Clarke et al, Nature Nanotechnology, 4(4): 265-270 (2009); Bayley et al, U.S. patent publication 2003/0215881; and the like.

[0035] Briefly, in one aspect, a 1-50 nm channel is formed through a substrate, usually a membrane, through which an analyte, such as single stranded DNA, is induced to translocate. The solid-state approach of generating nanopores offers robustness and durability as well as the ability to tune the size and shape of the nanopore, the ability to fabricate high-density arrays of nanopores on a wafer scale, superior mechanical, chemical and thermal characteristics compared with lipid-based systems, and the possibility of integrating with electronic or optical readout techniques. Biological nanopores on the other hand provide reproducible narrow bores, or lumens, especially in the 1-10 nanometer range, as well as techniques for tailoring the physical and/or chemical properties of the nanopore and for directly or indirectly attaching groups or elements, such as fluorescent labels, which may be FRET donors or acceptors, by conventional protein engineering methods. Protein nanopores typically rely on delicate lipid bilayers for mechanical support, and the fabrication of solid-state nanopores with precise dimensions remains challenging. Combining solid-state nanopores with a biological nanopore overcomes some of these shortcomings, especially the precision of a biological pore protein with the stability of a solid state nanopore. For optical read out techniques a hybrid nanopore provides a precise location of the nanopore which simplifies the data acquisition greatly. The lateral diffusion of nanopore proteins inserted in a lipid bilayer makes an optical detection challenging. Since the biological part (i.e. protein nanopore part) of a hybrid nanopore does not rely on the insertion in a lipid bilayer, the degrees of freedom for modifications made to such a protein are greatly increased e.g. a genetically modified nanopore protein that does not spontaneously insert in a lipid bilayer may still be used as a protein component of a hybrid nanopore. Also, bilayer-destabilizing agents such as quantum dots may be used to label a protein component of a hybrid nanopore.

[0036] In one embodiment, the invention is directed to a method for analyzing one or more polymer analytes, such as determining a nucleotide sequence of a polynucleotide, which comprises the following steps: (a) translocating a polymer analyte through a nanopore having a bore and an exit, the polymer analyte comprising a sequence of monomers, wherein substantially each monomer is labeled with a fluorescent label such that fluorescent labels of adjacent monomers are in a quenched state by self-quenching one

another outside of the nanopore and fluorescent labels are in a sterically constrained state and incapable of generating a detectable fluorescent signal inside of the nanopore; (b) exciting each fluorescent label at the exit of the nanopore as it transitions from a sterically constrained state to a quenched state so that a fluorescent signal is generated which is indicative of the monomer to which it is attached; (c) detecting the fluorescent signal to identify the monomer. As used herein, “substantially every”, “substantially all”, or like terms, in reference to labeling monomers, particularly nucleotides, acknowledges that chemical labeling procedures may not result in complete labeling of every monomer; to the extent practicable, the terms comprehend that labeling reactions in connection with the invention are continued to completion; in some embodiments, such completed labeling reactions include labeling at least fifty percent of the monomers; in other embodiments, such labeling reactions include labeling at least eighty percent of the monomers; in other embodiments, such labeling reactions include labeling at least ninety-five percent of the monomers; in other embodiments, such labeling reactions include labeling at least ninety-nine percent of the monomers.

[0037] In another embodiment, the invention is directed to a method for analyzing one or more polymer analytes comprising the following steps: (a) attaching a fluorescent label substantially every monomer of one or more polymer analytes such that fluorescent labels of adjacent monomers are in a quenched state, (b) translocating the polymer analytes through nanopores so that monomers of each polymer analyte traverses the nanopore in single file and wherein each nanopore has a bore and an exit, the bore sterically constraining the fluorescent labels in a constrained state so that no fluorescent signal is generated therefrom inside the bore; (c) exciting during a transition interval each fluorescent label at the exit of the nanopore as each fluorescent label transitions from a sterically constrained state to a quenched state, thereby generating a fluorescent signal that is indicative of the monomer to which it is attached; (c) detecting the fluorescent signal to identify the monomer.

[0038] In another embodiment the invention is directed to a device for analyzing one or more labeled polymer analytes, such as a device for determining a nucleotide sequence of one or more labeled polynucleotide analytes, such device comprising the following elements: (a) a solid phase membrane separating a first chamber and a second chamber, the solid phase membrane having at least one nanopore fluidly connecting the first chamber and the second chamber through a bore or lumen, the bore or lumen having a cross-

sectional dimension such that labels of a labeled polymer translocating therethrough are sterically constrained so that detectable signals are not generated, and so that the labels of adjacent monomers of the labeled polymer are self-quenching; (b) an excitation source for exciting each label when it exits the nanopore and enters the second chamber so that a signal is generated indicative of a monomer to which the label is attached; and (c) a detector for collecting at least a portion of the signal generated by each excited label; and (d) identifying the monomer to which the excited label is attached by the collected signal.

[0039] In another embodiment, the invention is directed to a system for analyzing polymers comprising a polymer comprising monomers that are substantially all labeled with a mutually quenching dye set and a nanopore device for sequentially detecting optical signals from the dyes of the mutually quenching dye set which are attached to the polymer. Such an embodiment for determining a sequence of a polynucleotide may comprise the following elements: (a) a solid phase membrane separating a first chamber and a second chamber, the solid phase membrane having at least one aperture connecting the first chamber and the second chamber, and having a hydrophobic coating on at least one surface; (b) a lipid layer disposed on the hydrophobic coating; (c) a protein nanopore immobilized in the aperture, the protein nanopore having a bore with an exit, and the protein nanopore interacting with the lipid layer to form a seal with the solid phase membrane in the aperture so that fluid communication between the first chamber and the second chamber occurs solely through the bore of the protein nanopore, and the protein nanopore being cross-sectionally dimensioned so that nucleotides of the polynucleotide pass through the exit of the bore in sequence and so that fluorescent labels attached to the polynucleotide are sterically constrained so that generation of fluorescent signal therein is inhibited or prevented; and (d) a first member of the FRET pair attached to the solid phase membrane or the protein nanopore, so that whenever nucleotides of the polynucleotide emerge from the bore, a plurality of the nucleotides are within a FRET distance of the first member of the FRET pair. In some embodiments, the first member of the FRET pair is a quantum dot that functions as a FRET donor.

[0040] In some embodiments, the hydrophobic coating is optional in that the surface of the solid phase membrane is sufficiently hydrophobic itself so that a lipid layer adheres to it stably. The at least one aperture will have an inner surface, or wall, connected to, or contiguous with the surfaces of the solid phase membrane. In some embodiments, the at least one aperture will be a plurality of apertures, and the plurality of apertures may be

arranged as a regular array, such as a rectilinear array of apertures, the spacing of which depending in part on the number and kind of FRET pairs employed and the optical detection system used. Each of the apertures has a diameter, which in some embodiments is such that a protein nanopore is substantially immobilized therein. In some embodiments, substantially immobilized means that a protein nanopore may move no more than 5 nm in the plane of the solid phase membrane relative to the wall of the aperture. In another embodiment, substantially immobilized means that a protein nanopore may move no more than 5 nm in the plane of the solid phase membrane relative to the wall of the aperture. The protein nanopores each have a bore, or passage, or lumen, which permits fluid communication between the first and second chambers when the protein nanopore is immobilized in an aperture. Generally, the bore is coaxially aligned with the aperture. One function of the hydrophobic layer is to provide a surface to retain lipids in and/or immediately adjacent to the at least one aperture. Such lipids, in turn, permit disposition and immobilization of a protein nanopore within an aperture in a functional conformation and in a manner that forms a fluid seal with the wall of the aperture. In some embodiments, such seal also prevents electrical current passing between the first and second chambers around the protein nanopore. In some embodiments, charged analytes are disposed in an electrolyte solution in the first chamber and are translocated through the bore(s) of the protein nanopore(s) into an electrolytic solution in the second chamber by establishing an electrical field across the solid phase membrane. For convenience of manufacture, in some embodiments the hydrophobic coating will be on one surface of the solid phase membrane and the wall(s) of the aperture(s).

[0041] In some embodiments of the devices of the invention, the at least one nanopore in a solid phase membrane is a plurality of nanopores, or a nanopore array; in some embodiments such nanopores are spaced regularly in the solid phase membrane with their bores oriented perpendicularly to the plane of the solid phase membrane. In some embodiments, nanopores are spaced in a rectilinear pattern in the solid phase membrane; in other embodiments, nanopores are spaced in a random pattern in the solid phase membrane; in some embodiments, such random pattern is Poisson distributed. In some embodiments, nanopores are regularly spaced in a solid phase membrane with a minimal inter-nanopore distance of at least 10 nm; in other embodiments, such minimal inter-nanopore distance is 50 nm; in other embodiments, such minimal inter-nanopore distance

is 100 nm; in other embodiments, such minimal inter-nanopore distance is 200 nm; in other embodiments, such minimal inter-nanopore distance is 500 nm.

[0042] In some embodiments, methods and devices of the invention comprise a solid phase membrane, such as a SiN membrane, having an array of apertures therethrough providing communication between a first chamber and a second chamber (also sometimes referred to as a “cis chamber” and a “trans chamber”) and supporting a lipid bilayer on a surface facing the second, or trans, chamber. In some embodiments, diameters of the aperture in such a solid phase membrane may be in the range of 10 to 200 nm, or in the range of 20 to 100 nm. In some embodiments, such solid phase membranes further include protein nanopores inserted into the lipid bilayer in regions where such bilayer spans the apertures on the surface facing the trans chamber. In some embodiments, such protein nanopores are inserted from the cis side of the solid phase membrane using techniques described herein. In some embodiments, such protein nanopores have a structure identical to, or similar to, α -hemolysin in that it comprises a barrel, or bore, along an axis and at one end has a “cap” structure and at the other end has a “stem” structure (using the terminology from Song et al, Science, 274: 1859-1866 (1996)). In some embodiments using such protein nanopores, insertion into the lipid bilayer results in the protein nanopore being oriented so that its cap structure is exposed to the cis chamber and its stem structure is exposed to the trans chamber.

[0043] In some embodiments, methods and devices of the invention comprise droplet interface bilayers, either as single droplets or as arrays droplets, for example, as disclosed in Bayley et al, U.S. patent publication 2014/0356289; Huang et al, Nature Nanotechnology, 10.1038/nnano.2015.189. [Epub ahead of print]; or like reference, which are hereby incorporated by reference. Briefly, protein nanopores (1.2 nM) are placed in a 200–350 nl droplet (for example, 1.32 M KCl, 8.8 mM HEPES, 0.4 mM EDTA, pH 7.0 (α HL) or 8.0 (MspA), and incubated in, for example, 3 mM 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC) in hexadecane to form a lipid monolayer coating. A droplet may then be transferred by pipetting onto a coverslip in a measurement chamber, for example, that permits application of voltages to move analytes and optical detection, for example, by TIRF. The coverslip may be spin coated (3,000 r.p.m., 30 s) with a thin layer (~200 nm) of agarose (0.66 M CaCl₂, 8.8 mM HEPES, pH 7.0 (α HL)/8.0 (MspA)) and subsequently incubated with 3 mM DPhPC in hexadecane. On contact with the monolayer on the agarose, a lipid coated droplet spontaneously forms a droplet interface bilayer. A ground

electrode (Ag/AgCl) may be inserted into the droplet, with a corresponding active electrode (Ag/AgCl) in the substrate agarose. Voltage protocols may be applied with a patch clamp amplifier (for example, Axopatch 200B, Molecular Devices). Nanopores present in the droplet spontaneously insert into the droplet interface bilayer, and the ion flux may be detected both electrically and/or optically (for example, by way of an ion-sensitive dye, such as Fluo-8, or the like).

[0044] In some embodiments, the solid phase membrane may be treated with a low energy ion beam to bleach its autofluorescence, e.g. as described in Huber et al, U.S. patent publication 2013/0203050, which is incorporated herein by reference.

[0045] Figs. 2A-2C are diagrams of hybrid biosensors. A nanometer sized hole (102) is drilled into a solid-state substrate, or solid phase membrane, (103) which separates two chambers, or compartments cis (101) and trans (107). A protein biosensor (e.g a protein nanopore) (104) attached to a charged polymer (105), such as a single stranded DNA, is embedded into the solid-state nanohole by electrophoretic transport. In Fig. 1C the protein biosensor is inserted. In a nanometer sized hole which surface has a hydrophobic coating (106) and a lipid layer (109) attached thereto. A nanopore may have two sides, or orifices. 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.

[0046] Fig. 2D shows protein nanopore (104) inserted into an aperture drilled in a solid state membrane (103). Attached to the protein nanopore (104) is an oligonucleotide (108) to which a complementary secondary oligonucleotide (111) is hybridized. Said secondary oligonucleotide (111) has one or more second members of a FRET pair (110) attached to it. Alternatively, a member of a FRET pair may be directly attached to an amino acid of a protein nanopore. For example, a hemolysin subunit may be modified by conventional genetic engineering techniques to substitute a cysteine for a suitably located amino acid adjacent to the exit of the nanopore, e.g. the threonine 129. An oligonucleotide or members of a FRET pair may be attached via the thio group of the cysteine using conventional linker chemistries, e.g. Hermanson (cited above).

[0047] In some embodiments, the present invention employs a hybrid nanopore, particularly for optical-based nanopore sequencing of polynucleotides. Such embodiments comprise a solid-state orifice, or aperture, into which a protein biosensor, such as a protein nanopore, is stably inserted. A protein nanopore (e.g. alpha hemolysin) may be attached to a charged polymer (e.g. double stranded DNA) which serves as a drag force in an applied electric field, and which may be used to guide a protein nanopore into an aperture in a solid-state membrane. In some embodiments, the aperture in the solid-state substrate is selected to be slightly smaller than the protein, thereby preventing it from translocating through the aperture. Instead, the protein will be embedded into the solid-state orifice. The solid-state substrate can be modified to generate active sites on the surface that allow the covalent attachment of the plugged-in protein biosensor resulting in a stable hybrid biosensor.

[0048] The polymer attachment site in the biosensor can be generated by protein engineering e.g. a mutant protein can be constructed that will allow the specific binding of the polymer. As an example, a cysteine residue may be inserted at the desired position of the protein. The cysteine can either replace a natural occurring amino acid or can be incorporated as an addition amino acid. Care must be taken not to disrupt the biological function of the protein. The terminal primary amine group of a polymer (i.e. DNA) is then activated using a hetero-bifunctional crosslinker (e.g. SMCC). Subsequently, the activated polymer is covalently attached to the cysteine residue of the protein biosensor. In some embodiments, the attachment of the polymer to the biosensor is reversible. By implementing a cleavable crosslinker, an easily breakable chemical bond (e.g. an S-S bond) is introduced and the charged polymer may be removed after insertion of the biosensor into the solid-state aperture.

[0049] For someone skilled in the art it is obvious that a wide variety of different approaches for covalent or non-covalent attachment methods of a charged polymer to the protein biosensor are possible and the above described approach merely serves as an example. The skilled artisan will also realize that a variety of different polymers may be used as a drag force, including, but not limited to, single or double stranded DNA, polyethyleneglycol (PEG), polyvinylpyrrolidone (PVP), poly-L-lysine, linear polysaccharides etc. It is also obvious that these polymers may exhibit either a negative (-) or positive (+) charge at a given pH and that the polarity of the electric field may be adjusted accordingly to pull the polymer-biosensor complex into a solid-state aperture.

[0050] In some embodiments, a donor fluorophore is attached to the protein nanopore. This complex is then inserted into a solid-state aperture or nanohole (for example, 3-10 nm in diameter) by applying an electric field across the solid state nanohole until the protein nanopore is transported into the solid-state nanohole to form a hybrid nanopore. The formation of the hybrid nanopore can be verified by (a) the inserting protein nanopore causing a drop in current based on a partial blockage of the solid-state nanohole and by (b) the optical detection of the donor fluorophore.

[0051] Once stable hybrid nanopores have formed single stranded, fluorescently labeled (or acceptor labeled) DNA is added to the cis chamber (the chamber with the (+) electrode). The applied electric field forces the negatively charged ssDNA to translocate through the hybrid nanopore during which the labeled nucleotides get in close vicinity of the donor fluorophore.

[0052] Solid state, or synthetic, nanopores may be prepared in a variety of ways, as exemplified in the references cited above. In some embodiments a helium ion microscope may be used to drill the synthetic nanopores in a variety of materials, e.g. as disclosed by Yang et al, Nanotechnology, 22: 285310 (2011), which is incorporated herein by reference. A chip that supports one or more regions of a thin-film material, e.g. silicon nitride, that has been processed to be a free-standing membrane is introduced to the helium ion microscope (HIM) chamber. HIM motor controls are used to bring a free-standing membrane into the path of the ion beam while the microscope is set for low magnification. Beam parameters including focus and stigmation are adjusted at a region adjacent to the free-standing membrane, but on the solid substrate. Once the parameters have been properly fixed, the chip position is moved such that the free-standing membrane region is centered on the ion beam scan region and the beam is blanked. The HIM field of view is set to a dimension (in μm) that is sufficient to contain the entire anticipated nanopore pattern and sufficient to be useful in future optical readout (i.e. dependent on optical magnification, camera resolution, etc.). The ion beam is then rastered once through the entire field of view at a pixel dwell time that results in a total ion dose sufficient to remove all or most of the membrane autofluorescence. The field of view is then set to the proper value (smaller than that used above) to perform lithographically-defined milling of either a single nanopore or an array of nanopores. The pixel dwell time of the pattern is set to result in nanopores of one or more predetermined diameters, determined through the use of a

calibration sample prior to sample processing. This entire process is repeated for each desired region on a single chip and/or for each chip introduced into the HIM chamber.

[0053] In some embodiments, the solid-state substrate may be modified to generate active sites on the surface that allow the covalent attachment of the plugged in protein biosensor or to modify the surface properties in a way to make it more suitable for a given application. Such modifications may be of covalent or non-covalent nature. A covalent surface modification includes a silanization step where an organosilane compound binds to silanol groups on the solid surface. For instance, the alkoxy groups of an alkoxy silane are hydrolyzed to form silanol-containing species. Reaction of these silanes involves four steps. Initially, hydrolysis of the labile groups occurs. Condensation to oligomers follows. The oligomers then hydrogen bond with hydroxyl groups of the substrate. Finally, during drying or curing, a covalent linkage is formed with the substrate with concomitant loss of water. For covalent attachment organosilanes with active side groups may be employed. Such side groups consist of, but are not limited to epoxy side chain, aldehydes, isocyanates, isothiocyanates, azides or alkynes (click chemistry) to name a few. For someone skilled in the art it is obvious that multiple ways of covalently attaching a protein to a surface are possible. For instance, certain side groups on an organosilane may need to be activated before being capable of binding a protein (e.g. primary amines or carboxyl side groups activated with an N-hydroxysuccinimide ester). Another way of attaching a protein to the solid surface may be achieved through affinity binding by having one affinity partner attached to the protein and the second affinity partner being located on the solid surface. Such affinity pairs consist of the group of, but are not limited to biotin-streptavidin, antigen-antibody and aptamers and the corresponding target molecules. In a preferred embodiment the surface modification of the solid state nanopore includes treatment with an organosilane that renders the surface hydrophobic. Such organosilanes include but are not limited to, alkanesilanes (e.g. octadecyldimethylchlorosilane) or modified alkanesilanes such as fluorinated alkanesilanes with an alkane chain length of 5 to 30 carbons. The hydrophobic surface is then treated with a dilute solution of a lipid in pentane. After drying of the solvent and immersing the surface in an aqueous solution the lipid will spontaneously form a layer on the surface. A layer of lipid on the solid surface might prove beneficial for the formation of a hybrid nanopore. The lipid layer on the solid phase might reduce the leak current between protein and solid state nanopore and it might increase the stability of the inserted protein pore. Combining a low capacitance solid substrate as well

as a lipid coating of said substrate may render the hybrid nanopore system amenable to an electrical readout based on current fluctuations generated by translocation of DNA through the hybrid nanopore. To achieve electrical read out with such a system a means of decreasing the translocation speed of unmodified DNA must be combined with a lipid coated hybrid nanopore. Molecular motors such as polymerases or helicases may be combined with a hybrid nanopore and effectively reduce the translocation speed of DNA through the hybrid nanopore. The lipids used for coating the surface are from the group of sphingolipids, phospholipids or sterols. 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 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.

[0054]

[0055] 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. The labeling of a pore structure may be achieved via covalent or non-covalent interactions.

[0056] A donor label (also sometimes referred to as a “pore label”) may be placed as close as possible to the aperture, for example, the exit, of a nanopore without causing an occlusion that impairs translocation of a nucleic acid through the nanopore. 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, such as an acceptor 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 to 10^9 excitation and energy transfer cycles.

Labels for Nanopores and Analytes

[0057] In some embodiments, a nanopore may be labeled with one or more quantum dots. In particular, in some embodiments, one or more quantum dots may be attached to a nanopore, or attached to a solid phase support adjacent to (and within a FRET distance of an entrance or exit of a nanopore), and employed as donors in FRET reactions with acceptors on analytes. Such uses of quantum dots are well known and are described widely in the scientific and patent literature, such as, in U.S. patents 6,252,303; 6,855,551; 7,235,361; and the like, which are incorporated herein by reference.

[0058] One example of a Quantum dot which may be utilized as a pore label is a CdTe quantum dot which can be synthesized in an aqueous solution. 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.

[0059] For example, the primary amine of the lysine 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. Alternatively, amino acid 129 (threonine) may be exchanged into cysteine. Since there is no other cysteine residue in the natural alpha hemolysin protein the thiol side group of the newly inserted cysteine may be used to covalently attach other chemical moieties.

[0060] 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 cysteine 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.

[0061] 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, as discussed below. 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.

[0062] 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.

[0063] 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.

[0064] 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.

[0065] 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

steric hindrance of the labels during the polymerization process into newly synthesized DNA.

[0066] In some embodiments, 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 (Hoevel, 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). Yet another example are azide or alkyne modified nucleotides which are readily incorporated by polymerases (Gierlich et al., *Chem. Eur. J.*, 2007, 13, 9486-0404). The alkyne or azide modified polynucleotide is subsequently labeled with an azide or alkyne modified fluorophore following well established click chemistry protocols.

[0067] As mentioned above, in some embodiments, DNA may be labeled using "click chemistry," e.g. using commercially available kits (such as "Click-It" from Life Technologies, Carlsbad, CA). Click chemistry in general refers to a synthetic process in which two molecules are linked together by a highly efficient chemical reaction, one which is essentially irreversible, in which the yield is nearly 100%, and which produces few or no reaction byproducts.. More recently, the meaning has come to refer to the cyclization reaction of a substituted alkyne with a substituted azide to form a 1,2,3-triazole bearing the two substituents. When catalyzed by copper at room temperature the reaction is known as the Huisgen cycloaddition, and it fully satisfies the requirements for click chemistry in that no other chemical functionality on the two molecules is affected during the reaction. Thus the coupling reaction has found broad application in bioconjugate chemistry, for example, in dye labeling of DNA or proteins, where many amine, hydroxy, or thiol groups may be found. The key requirement is that an alkyne group and an azide can easily be introduced into the molecules to be coupled. For example, in the coupling of a fluorescent dye to a DNA oligonucleotide, the azide group is typically introduced synthetically into the dye, while the alkyne group is incorporated into the DNA during oligonucleotide synthesis. Upon mixing in the presence of Cu^+ the two components are quickly coupled to form the triazole, in this case bearing the oligonucleotide as one substituent and the dye as the other. Another more recent advance provides the alkyne component within a strained ring structure. In this case the reaction with an azide does not require the copper catalyst, being

driven by release of the ring strain energy as the triazole is formed. This is better known as the copper-free click reaction. Guidance for applying click chemistry to methods of the invention may be found in the following references which are incorporated by reference: Rostovtsev VV, Green LG; Fokin, Valery V, Sharpless KB (2002). "A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes". *Angewandte Chemie International Edition* **41** (14): 2596--2599. Moses JE and Moorhouse AD (2007). "The growing applications of click chemistry". *Chem. Soc. Rev.* **36** (8): 1249-1262.

[0068] Whenever two or more mutually quenching dyes are used, such dyes may be attached to DNA using orthogonal attachment chemistries. For instance NHS esters can be used to react very specifically with primary amines or maleimides will react with thiol groups. Either primary amines (NH₂) or thiol (SH) modified nucleotides are commercially available. These relatively small modifications are readily incorporated in a polymerase mediated DNA synthesis and can be used for subsequent labeling reactions using either NHS or maleimide modified dyes. Guidance for selecting and using such orthogonal linker chemistries may be found in Hermanson (cited above).

[0069] Additional orthogonal attachment chemistries are shown in Figs. 3A-3H. Fig. 3A shows typical attachment positions of linking moieties on nucleoside bases. Fig. 3B shows a reaction diagram for Huisgen-type cycloaddition for a copper-catalyzed reaction and an uncatalyzed reaction, e.g. disclosed in the references cited above. Fig. 3C shows a reaction diagram for alkene plus nitrile oxide cycloaddition, e.g. as disclosed in Gutmiedl et al, *Org. Lett.*, 11: 2405-2408 (2009). Fig. 3D shows a reaction diagram for Diels-Alder cycloaddition, e.g. disclosed in Seelig et al, *Tetrahedron Lett.*, 38: 7729-7732 (1997). Fig. 3E shows a reaction diagram for carbonyl ligation, e.g. as disclosed in Casi et al, *J. Am. Chem. Soc.*, 134: 5887-5892 (2012); Shao et al *J. Am. Chem. Soc.*, 117: 3893-3899 (1995); Rideout, *Science*, 233: 561-563 (1986); or the like. Fig. 3F shows a reaction diagram for Michael addition, e.g. disclosed in Brinkley, *Bioconjugate Chemistry*, 3: 2-13 (1992). Fig. 3G shows a reaction diagram for native chemical ligation, e.g. disclosed in Schuler et al, *Bioconjugate Chemistry*, 13: 1039-1043 (2002); Dawson et al, *Science*, 266: 776-779 (1994); or the like. Fig. 3H shows a reaction diagram for amide formation via an active ester, e.g. disclosed in Hermanson (cited above).

[0070] A nucleic acid molecule may be directly modified with N-Bromosuccinimide which upon reacting with the nucleic acid will result in 5-Bromocystein, 8-Bromoadenine 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, cited above).

[0071] 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.

[0072] 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. In some

embodiments, fluorescent labels or donor molecules are excited in a TIRF system with an evanescent wave, sometimes referred to herein as “evanescent wave excitation.”

[0073] 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.

[0074] 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.

[0075] 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.

[0076] 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 the exit of 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.

[0077] 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.

[0078] 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.

Translocation Speed

[0079] A major obstacle associated with nanopore based sequencing approaches is the high translocation velocity of nucleic acid through a nanopore (~500,000 – 1,000,000 nucleotides/sec) which doesn't allow for direct sequence readout due to the limited bandwidth of the recording equipment. A way of slowing down the nucleic acid translocation with two different nanopore proteins was recently shown by Cherf et al. (Nat Biotechnol. 2012 Feb 14; 30(4):344-8) and Manrao et al. (Nat Biotechnol. 2012 Mar 25; 30(4):349-53) and are incorporated herein by reference. Both groups used a DNA polymerase to synthesize a complementary strand from a target template which resulted in the step-wise translocation of the template DNA through the nanopore. Hence, the synthesis speed of the nucleic acid polymerase (10-500 nucleotides/sec) determined the translocation speed of the DNA and since it's roughly 3-4 orders of magnitude slower than direct nucleic acid translocation the analysis of single nucleotides became feasible. However, the polymerase-aided translocation requires significant sample preparation to generate a binding site for the polymerase and the nucleic acid synthesis has to be blocked in bulk and can only start once the nucleic acid-polymerase complex is captured by the nanopore protein. This results in a rather complex set-up which might prevent the implementation in a commercial setting. Furthermore, fluctuation in polymerase synthesis

reactions such as a stalled polymerization as well as the dissociation of the polymerase from the nucleic acid may hamper the sequence read-out resulting in a high error rate and reduced read-length, respectively. In some embodiments, a target nucleic acid is enzymatically copied by incorporating fluorescent modified nucleotides. In other embodiments, modified nucleotides with reactive groups are incorporated which can be labeled in a post-extension reaction. The resulting labeled nucleic acid has an increased nominal diameter which results in a decreased translocation velocity when pulled through a nanopore. The preferred translocation rate for optical sequencing lies in the range of 1-1000 nucleotides per second with a more preferred range of 200-800 nucleotides per second and a most preferred translocation rate of 200-600 nucleotides per second.

[0080] Alternatively, translocation speed of a polynucleotide, especially a single stranded polynucleotide, may be controlled by employing a nanopore dimensioned so that adducts and/or labels, e.g. organic dyes attached to bases, inhibit but do not prevent polynucleotide translocation. A translocation speed may be selected by attaching labels and/or adducts at a predetermined density. Such labels and/or adducts may have regular spaced attachments, e.g. every third nucleotide or the like, or they may have random, or pseudorandom attachments, e.g. every C may be labeled. In some embodiments, a selected number of different nucleotides may be labeled, e.g. every A and C, or every A and G, or every A and T, or every C, or the like, that results in an average translocation speed. Such average speed may be decreased by attaching adducts to unlabeled nucleotides. Adducts include any molecule, usually and organic molecule, that may be attached to a nucleotide using conventional chemistries. Typically adducts have a molecular weight in the same range as common organic dyes, e.g. fluorescein, Cy3, or the like. Adducts may or may not be capable of generating signals, that is, serving as a label. In some embodiments, adducts and/or labels are attached to bases of nucleotides. In other embodiments, labels and/or adducts may be attached to linkages between nucleosides in a polynucleotide. In one aspect, a method of controlling translocation velocity of a single stranded polynucleotide through a nanopore comprises the step of attaching adducts to the polynucleotide at a density, wherein translocation velocity of the single stranded polynucleotide monotonically decreases with a larger number of adducts attached, or with the density of adducts attached. In some embodiments, not every kind of nucleotide of a polynucleotide is labeled. For example, four different sets of a polynucleotide may be produced where nucleotides of each set are labeled with the same molecule, e.g. a fluorescent organic dye acceptor, but in

each set a different kind of nucleotide will be labeled. Thus, in set 1 only A's may be labeled; in set 2 only C's may be labeled; in set 3 only G's may be labeled; and so on. After such labeling, the four sets of polynucleotides may then be analyzed separately in accordance with the invention and a nucleotide sequence of the polynucleotide determined from the data generated in the four analysis. In such embodiments, and similar embodiments, e.g. two labels are used, where some of the nucleotides of a polynucleotide are not labeled, translocation speed through a nanopore will be affected by the distribution of label along the polynucleotide. To prevent such variability in translocation speed, in some embodiments, nucleotides that are not labeled with an acceptor or donor for generating signals to determine nucleotide sequence, may be modified by attaching a non-signal-producing adduct that has substantially the same effect on translocation speed as the signal-producing labels.

[0081] 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.

Kits

[0082] The invention may include kits for carrying out the methods of the invention. In some embodiments, kits include reagents for adding reactive groups to target polynucleotides. For example, a target polynucleotide for analysis in accordance with the invention may be obtained by transcribing its complement from a sample using a nucleic acid polymerase in the presence of nucleoside triphosphate analogs that include reactive groups, such as amines or thiols. Thus, in some embodiments kits comprise one or more nucleoside triphosphate analogs with reactive groups. Kits may further comprise one or more mutually quenching fluorescent labels with complementary functionalities to the reactive groups. Kits may further comprise a nucleic acid polymerase for incorporating nucleoside triphosphates into a target polynucleotide. Nucleic acid polymerases may include a reverse transcriptase when mRNA is used to produce target polynucleotides, or nucleic acid polymerases may include a DNA polymerase when genomic DNA is used to produce target polynucleotides. Kits may further comprise buffers, co-factors and like reagents for carrying out polymerase reactions. Likewise, kits may further include buffers

and other reaction components for carrying out reactions between reactive groups and complementary functionalities on mutually quenching fluorescent labels in order to produce a labeled target polynucleotide. Kits may further include solid phase membranes and protein nanopores for assembly into an operable nanopore array. Such latter kits may further include donor members of a FRET pair for attachment to protein nanopores or to a solid phase membrane. Kits may include assembled nanopore arrays comprising a solid phase membrane including incorporated protein nanopores and donor members of a FRET pair.

Definitions

[0083] “FRET” or “Forrester, or fluorescence, resonant energy transfer” means 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). “FRET distance” means a distance between a FRET donor and a FRET acceptor over which a FRET interaction can take place and a detectable FRET signal produced by the FRET acceptor.

[0084] “Kit” refers to any delivery system for delivering materials or reagents for carrying out a method of the invention. In the context of reaction assays, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (e.g., fluorescent labels, such as mutually quenching fluorescent labels, fluorescent label linking agents, enzymes, etc. in the appropriate containers) and/or supporting materials (e.g., buffers, written instructions for performing the assay etc.) from one location to another. For example, kits include one or more enclosures (e.g., boxes) containing the relevant reaction reagents and/or supporting materials. Such contents may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second or more containers contain mutually quenching fluorescent labels.

[0085] “Microfluidics” or “nanofluidics” device means an integrated system of one or more chambers, ports, and channels that are interconnected and in fluid communication and designed for carrying out an analytical reaction or process, either alone or in cooperation with an appliance or instrument that provides support functions, such as sample introduction, fluid and/or reagent driving means, temperature control, detection

systems, data collection and/or integration systems, and the like. Microfluidics and nanofluidics devices may further include valves, pumps, filters and specialized functional coatings on interior walls, e.g. to prevent adsorption of sample components or reactants, facilitate reagent movement by electroosmosis, or the like. Such devices are usually fabricated in or as a solid substrate, which may be glass, plastic, or other solid polymeric materials, and may have a planar format for ease of detecting and monitoring sample and reagent movement, especially via optical or electrochemical methods. In some embodiments, such devices are disposable after a single use. Features of a microfluidic device usually have cross-sectional dimensions of less than a few hundred square micrometers and passages typically have capillary dimensions, e.g. having maximal cross-sectional dimensions of from about 500 μm to about 0.1 μm . Microfluidics devices typically have volume capacities in the range of from 1 μL to a few nL, e.g. 10-100 nL. Dimensions of corresponding structures in nanofluidics devices are typically from 1 to 3 orders of magnitude less than those for microfluidics devices. The fabrication and operation of microfluidics and nanofluidics devices are well-known in the art as exemplified by the following references that are incorporated by reference: Ramsey, U.S. patents 6,001,229; 5,858,195; 6,010,607; and 6,033,546; Soane et al, U.S. patents 5,126,022 and 6,054,034; Nelson et al, U.S. patent 6,613,525; Maher et al, U.S. patent 6,399,952; Ricco et al, International patent publication WO 02/24322; Bjornson et al, International patent publication WO 99/19717; Wilding et al, U.S. patents 5,587,128; 5,498,392; Sia et al, Electrophoresis, 24: 3563-3576 (2003); Unger et al, Science, 288: 113-116 (2000); Enzelberger et al, U.S. patent 6,960,437; Cao, "Nanostructures & Nanomaterials: Synthesis, Properties & Applications," (Imperial College Press, London, 2004).

[0086] "Nanopore" means any opening positioned in a substrate that allows the passage of analytes through the substrate in a predetermined or discernable order, or in the case of polymer analytes, passage of their monomeric units through the substrate in a predetermined or discernible order. In the latter case, a predetermined or discernible order may be the primary sequence of monomeric units in the polymer. Examples of nanopores include proteinaceous or protein based nanopores, synthetic or solid state nanopores, and hybrid nanopores comprising a solid state nanopore having a protein nanopore embedded therein. A nanopore may have an inner diameter of 1-10 nm or 1-5 nm or 1-3 nm. Examples of protein nanopores include but are not limited to, alpha-hemolysin, voltage-

dependent mitochondrial porin (VDAC), OmpF, OmpC, MspA and LamB (malto porin), e.g. disclosed in Rhee, M. et al., Trends in Biotechnology, 25(4) (2007): 174-181; Bayley et al (cited above); Gundlach et al, U.S. patent publication 2012/0055792; and the like, which are incorporated herein by reference. Any protein pore that allows the translocation of single nucleic acid molecules may be employed. A nanopore 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. Pore proteins are chosen from a group of proteins such as, but not limited to, alpha-hemolysin, MspA, voltage-dependent mitochondrial porin (VDAC), Anthrax porin, OmpF, OmpC and LamB (malto porin). Integration of the pore protein into the solid state hole is accomplished by attaching a charged polymer to the pore protein. After applying an electric field the charged complex is electrophoretically pulled into the solid state hole. A synthetic nanopore, or solid-state nanopore, 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. 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). The substrate surface of a synthetic nanopore may be chemically modified to allow for covalent attachment of the protein pore or to render the surface properties suitable for optical nanopore sequencing. Such surface modifications can be covalent or non-covalent. Most covalent modification include an organosilane deposition for which the most common protocols are described: 1) Deposition from aqueous alcohol. This is the most facile method for preparing silylated surfaces. A 95% ethanol-5% water solution is adjusted to pH 4.5-5.5 with acetic acid. Silane is added with stirring to yield a 2% final concentration. After hydrolysis and silanol group formation the substrate is added for 2-5min. After rinsed free of excess materials by dipping briefly in ethanol. Cure of the silane layer is for 5-10 min at 110 degrees Celsius. 2) Vapor Phase Deposition. Silanes can be applied to substrates under dry aprotic conditions by chemical

vapor deposition methods. These methods favor monolayer deposition. In closed chamber designs, substrates are heated to sufficient temperature to achieve 5mm vapor pressure. Alternatively, vacuum can be applied until silane evaporation is observed. 3) Spin-on deposition. Spin-on applications can be made under hydrolytic conditions which favor maximum functionalization and polylayer deposition or dry conditions which favor monolayer deposition. In some embodiments, single nanopores are employed with methods of the invention. In other embodiments, a plurality of nanopores are employed. In some of the latter embodiments, a plurality of nanopores is employed as an array of nanopores, usually disposed in a planar substrate, such as a solid phase membrane. Nanopores of a nanopore array may be spaced regularly, for example, in a rectilinear pattern, or may be spaced randomly. In a preferred embodiment, nanopores are spaced regularly in a rectilinear pattern in a planar solid phase substrate.

[0087] “Peptide,” “peptide fragment,” “polypeptide,” “oligopeptide,” or “fragment” in reference to a peptide are used synonymously herein and refer to a compound made up of a single unbranched chain of amino acid residues linked by peptide bonds. Amino acids in a peptide or polypeptide may be derivatized with various moieties, including but not limited to, polyethylene glycol, dyes, biotin, haptens, or like moieties. The number of amino acid residues in a peptide varies widely; however, preferably, peptides or oligopeptides referred to herein usually have from 2 to 70 amino acid residues; and more preferably, they have from 2 to 50 amino acid residues. Polypeptides and peptide fragments referred to herein usually have from a few tens of amino acid residues, e.g. 20, to up to a few hundred amino acid residues, e.g. 200, or more.

[0088] “Polymer” means a plurality of monomers connected into a linear chain. Usually, polymers comprise more than one type of monomer, for example, as a polynucleotide comprising A’s, C’s, G’s and T’s, or a polypeptide comprising more than one kind of amino acid. Monomers may include without limitation nucleosides and derivatives or analogs thereof and amino acids and derivatives and analogs thereof. In some embodiments, polymers are polynucleotides, whereby nucleoside monomers are connected by phosphodiester linkages, or analogs thereof.

[0089] “Polynucleotide” or “oligonucleotide” are used interchangeably and each mean a linear polymer of nucleotide monomers. Whenever the use of an oligonucleotide or polynucleotide requires enzymatic processing, such as extension by a polymerase, ligation by a ligase, or the like, one of ordinary skill would understand that oligonucleotides or

polynucleotides in those instances would not contain certain analogs of internucleosidic linkages, sugar moieties, or bases at any or some positions. Polynucleotides typically range in size from a few monomeric units, e.g. 5-40, when they are usually referred to as "oligonucleotides," to several thousand monomeric units. Whenever a polynucleotide or oligonucleotide is represented by a sequence of letters (upper or lower case), such as "ATGCCTG," it will be understood that the nucleotides are in 5'→3' order from left to right and that "A" denotes deoxyadenosine, "C" denotes deoxycytidine, "G" denotes deoxyguanosine, and "T" denotes thymidine, "I" denotes deoxyinosine, "U" denotes uridine, unless otherwise indicated or obvious from context. Unless otherwise noted the terminology and atom numbering conventions will follow those disclosed in Strachan and Read, Human Molecular Genetics 2 (Wiley-Liss, New York, 1999). Usually polynucleotides comprise the four natural nucleosides (e.g. deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine for DNA or their ribose counterparts for RNA) linked by phosphodiester linkages; however, they may also comprise non-natural nucleotide analogs, e.g. including modified bases, sugars, or internucleosidic linkages. It is clear to those skilled in the art that where an enzyme has specific oligonucleotide or polynucleotide substrate requirements for activity, e.g. single stranded DNA, RNA/DNA duplex, or the like, then selection of appropriate composition for the oligonucleotide or polynucleotide substrates is well within the knowledge of one of ordinary skill, especially with guidance from treatises, such as Sambrook et al, Molecular Cloning, Second Edition (Cold Spring Harbor Laboratory, New York, 1989), and like references. Likewise, the oligonucleotide and polynucleotide may refer to either a single stranded form or a double stranded form (i.e. duplexes of an oligonucleotide or polynucleotide and its respective complement). It will be clear to one of ordinary skill which form or whether both forms are intended from the context of the terms usage.

[0090] "Sequence determination", "sequencing" or "determining a nucleotide sequence" or like terms in reference to polymers, such as polynucleotides, includes determination of partial as well as full sequence information of the polymer. In some embodiments, sequence determination may include detection or measurement of an identifying characteristic, or fingerprint, of a polymer, such as a unique, or substantially unique, sequence of signals that is correlated to a particular polymer sequence. In some embodiments, such correlation is a one-to-one correspondence. In other embodiments, such correlation may not be unique. In other embodiments, such correlation permits

identification of a polymer with a particular sequence with a probability of greater than ninety percent; in other embodiments, such identification can be made with a probability of greater than ninety-nine percent. In the case of polynucleotides, the above terms include identifying sequences of subsets of the full set of four natural nucleotides, A, C, G and T, such as, for example, a sequence of just A's and C's of a target polynucleotide. That is, the terms include the determination of the identities, ordering, and locations of one, two, three or all of the four types of nucleotides within a target polynucleotide. In some embodiments, the terms include the determination of the identities, ordering, and locations of two, three or all of the four types of nucleotides within a target polynucleotide. In some embodiments sequence determination may be accomplished by identifying the ordering and locations of a single type of nucleotide, e.g. cytosines, within the target polynucleotide "catgcgc . . ." so that its sequence is represented as a binary code, e.g. "100101 . . ." representing "c-(not c)(not c)c-(not c)-c . . ." and the like. In some embodiments, the terms may also include subsequences of a target polynucleotide that serve as a fingerprint for the target polynucleotide; that is, subsequences that uniquely identify a target polynucleotide within a set of polynucleotides, e.g. all different RNA sequences expressed by a cell.

CLAIMS

1. A method of determining a nucleotide sequence of at least one polynucleotide, the method comprising the steps of:

translocating at least one single stranded polynucleotide through a nanopore, wherein different kinds of nucleotides of the single stranded polynucleotide are labeled with different fluorescent labels from a mutually quenching set and wherein the nanopore has a bore dimensioned to force the fluorescent labels within the nanopore into a constrained state wherein substantially no detectable signal is generated and wherein each fluorescent label of the mutually quenching set (i) quenches fluorescence of every other fluorescent label of the same set whenever such labels are attached to adjacent nucleotides of a polynucleotide in free solution, and (ii) generates a distinct fluorescent signal in a non-quenched state;

exciting the fluorescent label of each nucleotide upon exiting the nanopore and during its transition from a constrained state to formation of a quenched state with an adjacent nucleotide;

measuring a fluorescent signal generated by the exiting fluorescent label to identify the nucleotide to which the fluorescent label is attached; and

determining a nucleotide sequence of the polynucleotide from a sequence of fluorescent signals.

2. The method of claim 1 wherein nucleotides of said polynucleotide are labeled with second members of a FRET pair, each second member producing a FRET signal indicative of the nucleotide to which it is attached, and wherein nucleotides of said polynucleotide pass in sequence by a first member of the FRET pair positioned adjacent to said nanopore so that each second member upon exiting said nanopore passes within a FRET distance of the first member of the FRET pair; and

wherein said step of exciting includes exposing the first member to a light beam of a first wavelength so that FRET occurs between the first and second members of the FRET pair within the FRET distance to generate a FRET signal of a second wavelength indicative of the nucleotide exiting said nanopore.

3. The method of claim 2 wherein said nanopore is disposed in a solid phase membrane and wherein said first member of said FRET pair is attached to the solid phase membrane adjacent to said nanopore.

4. The method of claim 2 wherein said nanopore is a protein nanopore and wherein said first member of said FRET pair is attached to the protein nanopore.
5. The method of claim 2 wherein said first member of said FRET pair is a donor and said second member of said FRET pair is an acceptor.
6. The method of claim 5 wherein said donor is a quantum dot.
7. The method of claim 5 wherein said acceptor is a fluorescent organic dye.
8. The method of claim 5 wherein said step of exposing includes exposing said donor to an evanescent wave from a total internally reflected excitation.
9. The method of claim 1 wherein said nanopore is in a nanopore array comprising a plurality of substantially identical nanopores.
10. The method of claim 9 wherein said nanopore array comprises a solid phase membrane separating a first chamber from a second chamber, the solid phase membrane comprising a planar array of apertures each having said protein nanopore immobilized therein.
11. A method of determining a nucleotide sequence of at least one polynucleotide, the method comprising the steps of:
 - translocating at least one single stranded polynucleotide through a nanopore, wherein substantially all different kinds of nucleotides of the single stranded polynucleotide are labeled with adducts and/or different fluorescent labels from a mutually quenching set and wherein the nanopore has a bore dimensioned to force the fluorescent labels within the nanopore into a constrained state so that substantially no detectable signal is generated and wherein the adducts and fluorescent labels are selected so that translocation speed through the nanopore is reduced to permit detection of individual nucleotides and wherein each adduct is an organic molecule having a molecular weight in the range of that of an organic dye and wherein each fluorescent label of the mutually quenching set (i) quenches fluorescence of every other fluorescent label of the same set whenever such labels are attached to adjacent

nucleotides of a polynucleotide in free solution, and (ii) generates a distinct fluorescent signal in a non-quenched state;

exciting the fluorescent label of each nucleotide upon exiting the nanopore and during its transition from a constrained state to formation of a quenched state with an adjacent nucleotide;

measuring a fluorescent signal generated by the exiting fluorescent label to identify the nucleotide to which the fluorescent label is attached; and

determining a nucleotide sequence of the polynucleotide from a sequence of fluorescent signals.

12. The method of claim 11 wherein said adducts and said fluorescent labels are selected so that said translocation speed is in the range of 1 to 1000 nucleotides per second.

13. The method of claim 12 wherein said nanopore is a protein nanopore.

14. The method of claim 13 wherein said nanopore is a hemolysin or a mutant thereof.

15. The method of claim 12 wherein said single stranded polynucleotides are labeled by polymerase extension.

16. The method of claim 12 wherein one or more of said adducts and/or fluorescent labels are selected from among rhodamine dyes, fluorescein dyes and cyanine dyes.

17. The method of claim 1 wherein said nanopore is a protein nanopore.

18. The method of claim 17 wherein said protein nanopore is a hemolysin or a variant thereof.

19. The method of claim 1 wherein said single stranded polynucleotides are labeled by polymerase extension.

20. The method of claim 1 wherein one or more of said fluorescein labels are selected from among rhodamine dyes, fluorescein dyes and cyanine dyes.

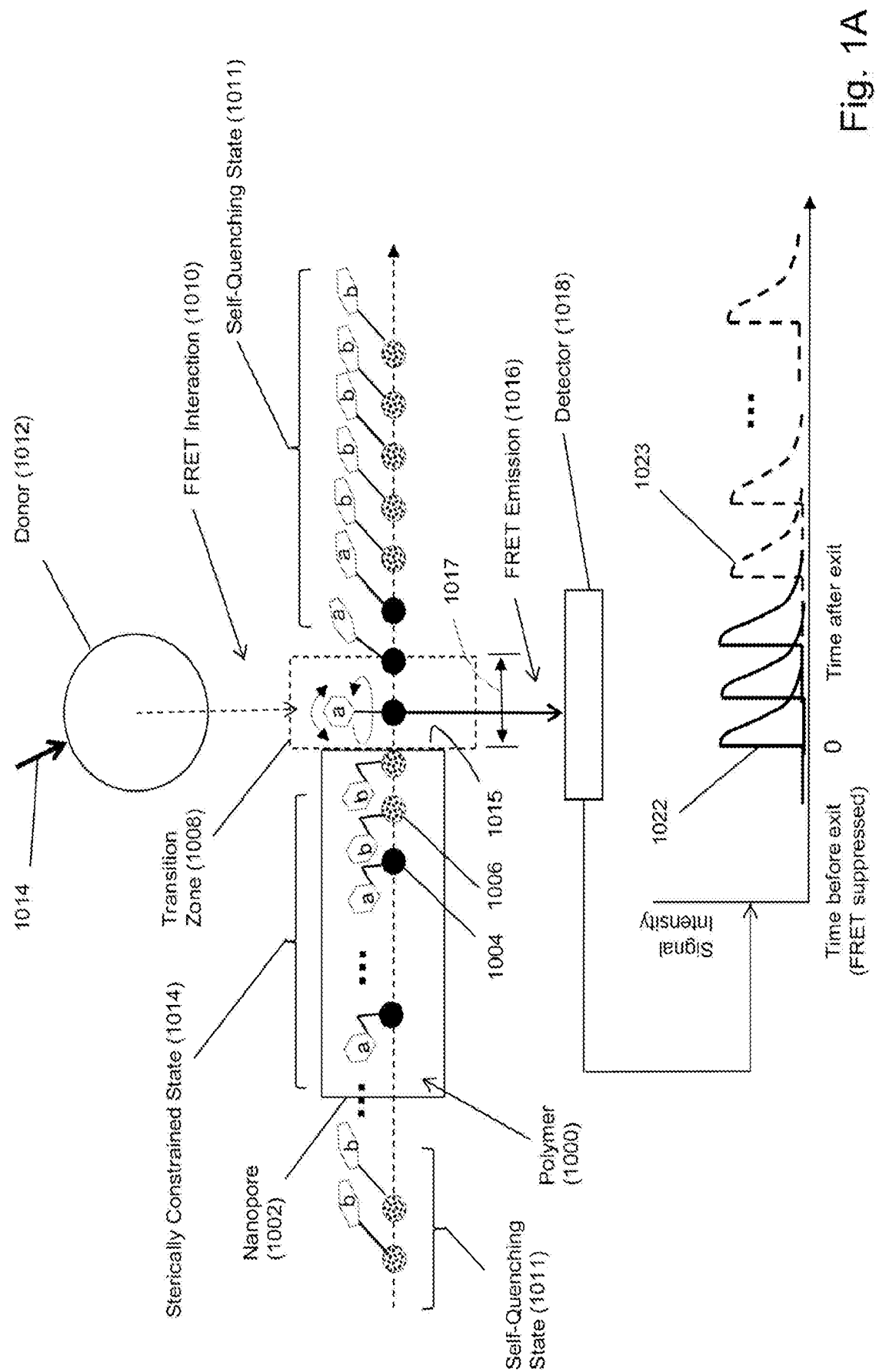


Fig. 1A

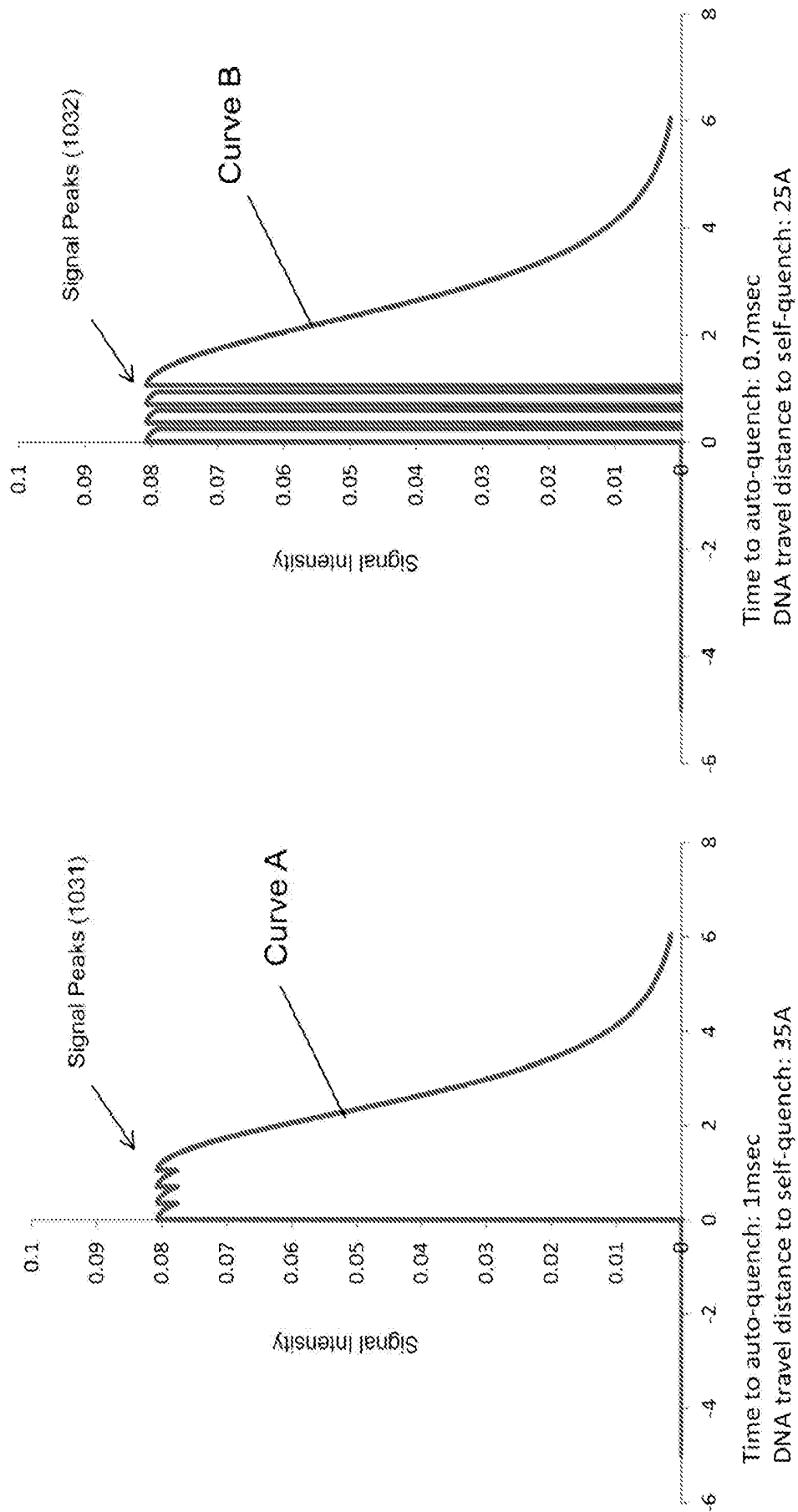
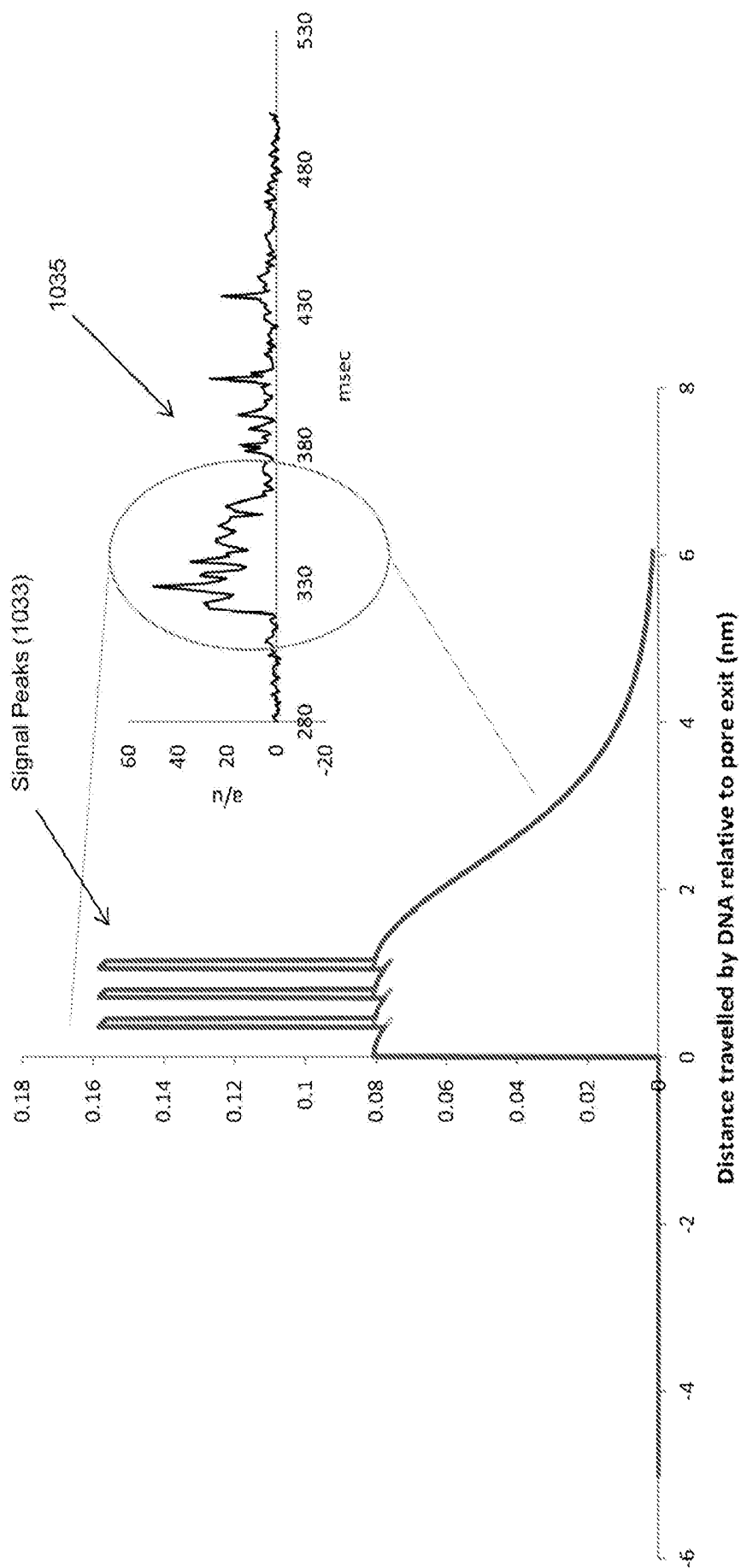


Fig. 1B



Time to auto-quench: 1.3msec
DNA travel distance to self-quench: 45A

Fig. 1C

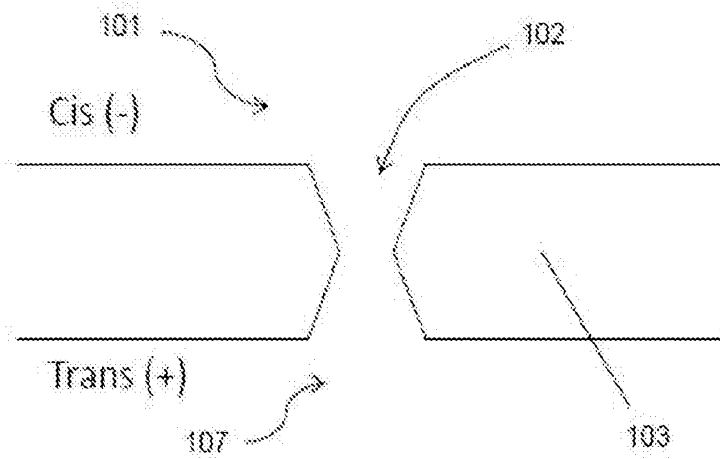


Fig. 2A

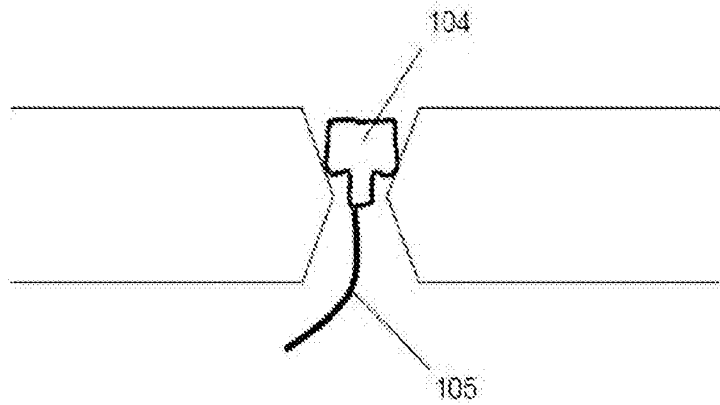


Fig. 2B

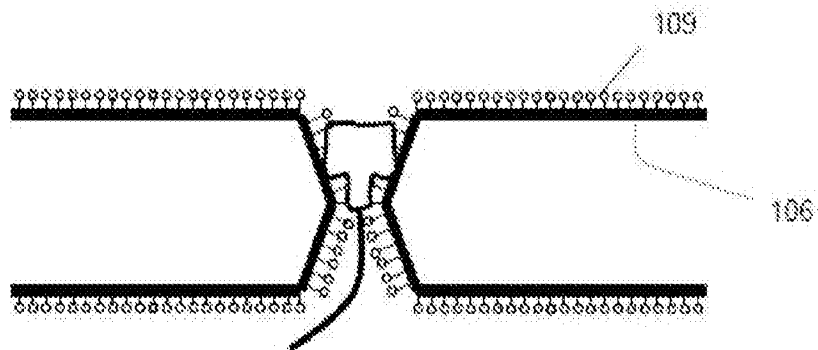


Fig. 2C

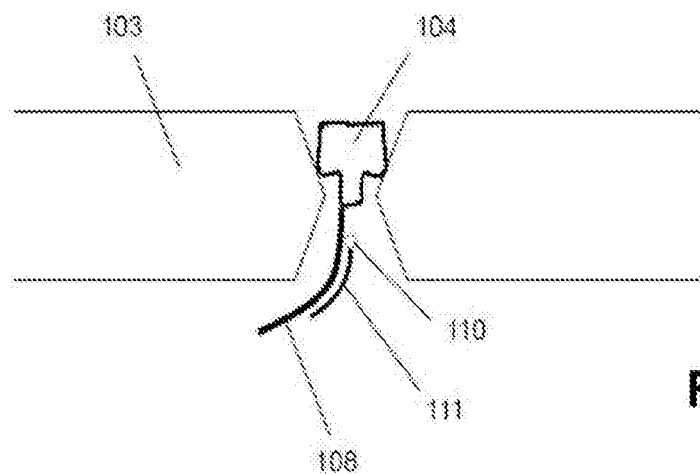


Fig. 2D

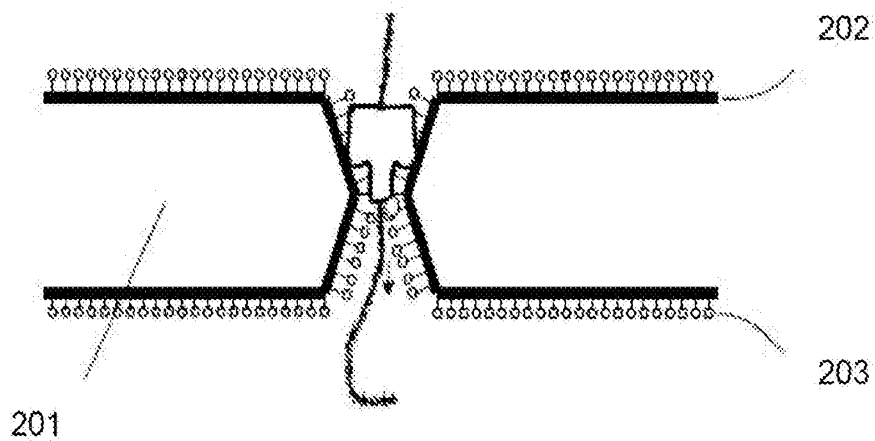


Fig. 2E

7/9

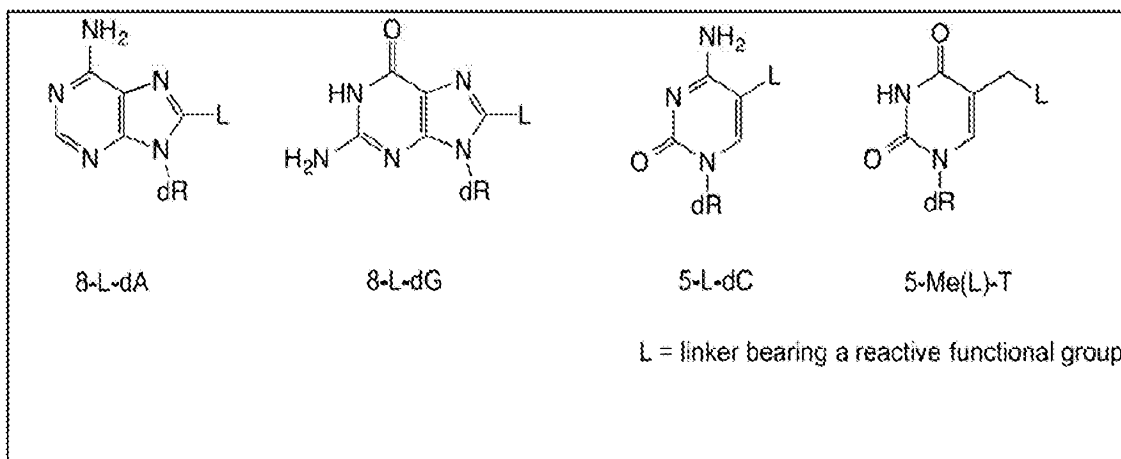
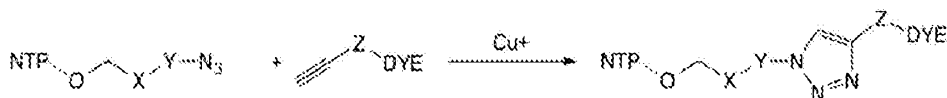


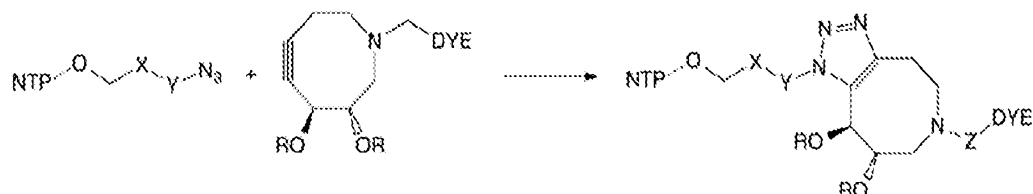
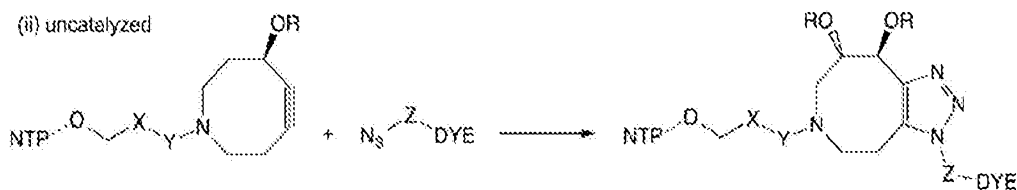
Fig. 3A

Huisgen-type cycloaddition ("click chemistry")

(i) copper-catalyzed



(ii) uncatalyzed



X, Y, Z (independently) = $(\text{CH}_2)_n$, $(\text{OCH}_2\text{CH}_2)_n$, CHR, CR_1R_2 , $(\text{CH}_2)_n\text{NHCO}(\text{CH}_2)_m$, $(\text{CH}_2)_n\text{NHCO}(\text{CH}_2)_m$, aryl
 $n = 0-10$

R = alkyl, aralkyl, PEG

Fig. 3B

8/9

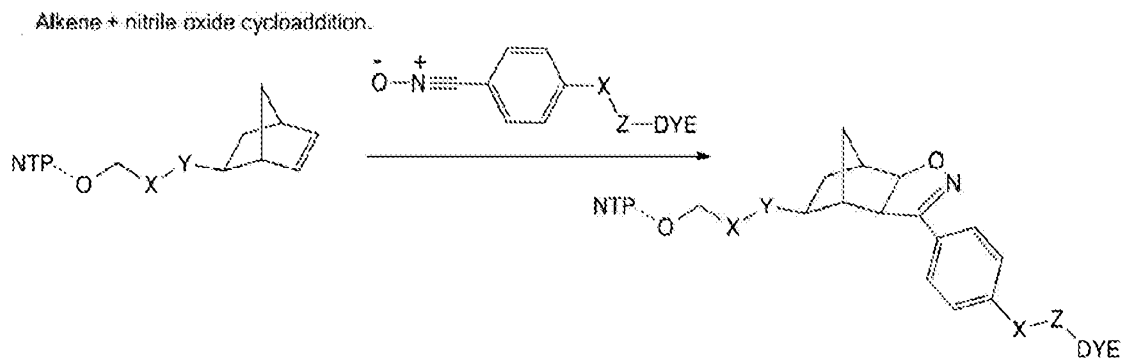


Fig. 3C

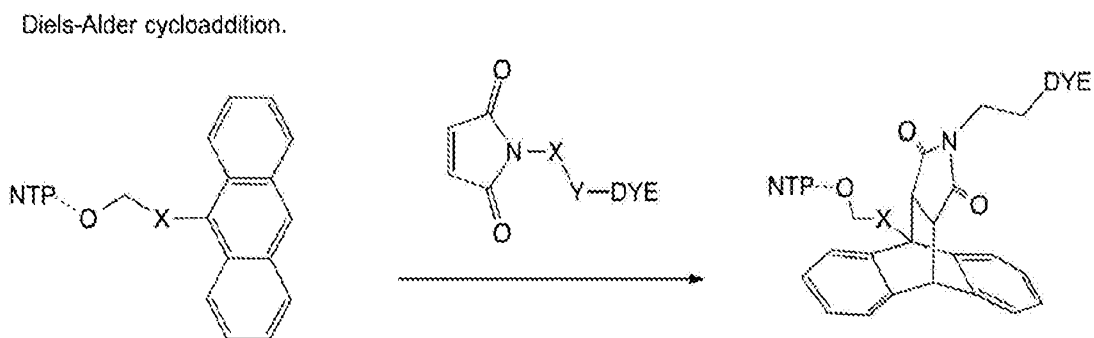


Fig. 3D

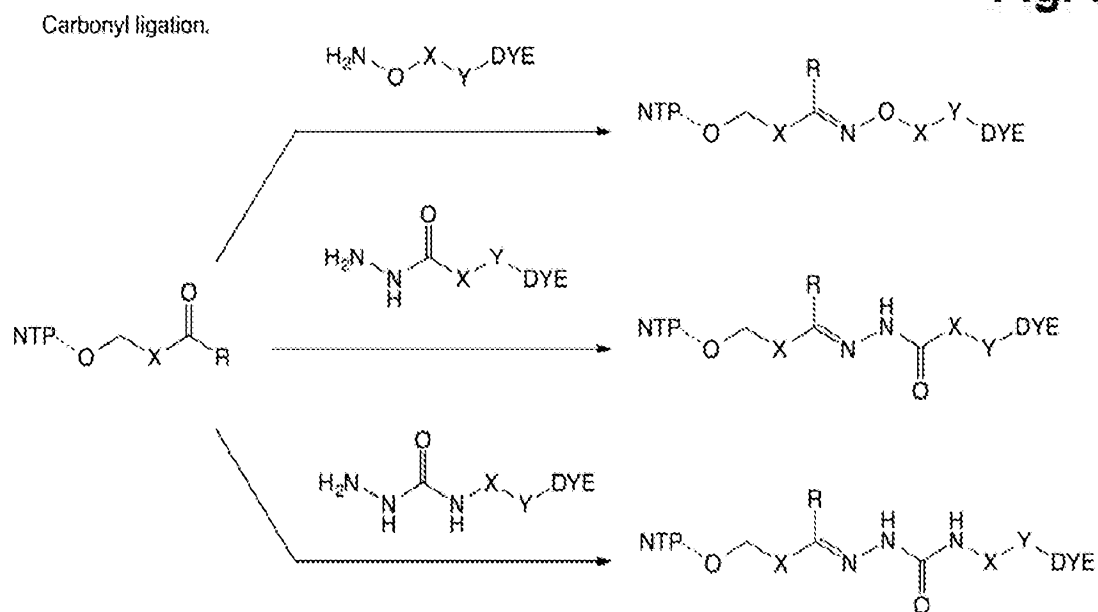


Fig. 3E

Michael addition.

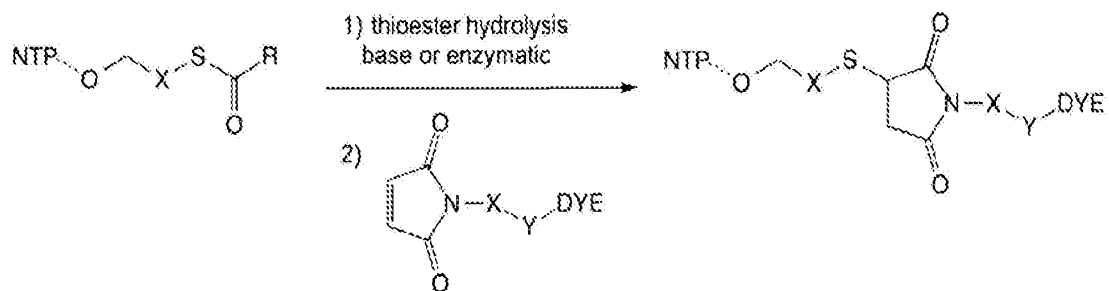


Fig. 3F

Native chemical ligation.

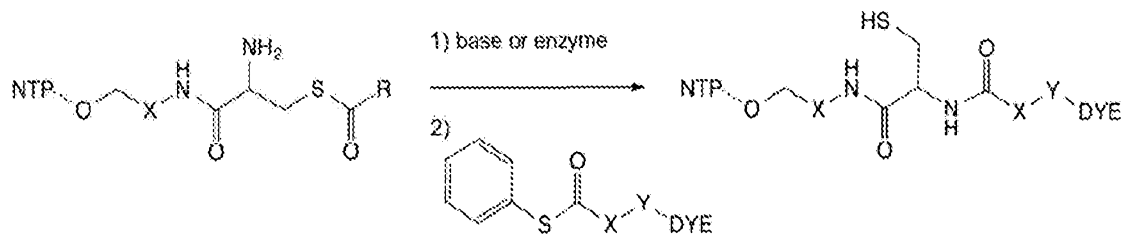


Fig. 3G

Amide formation via active ester.

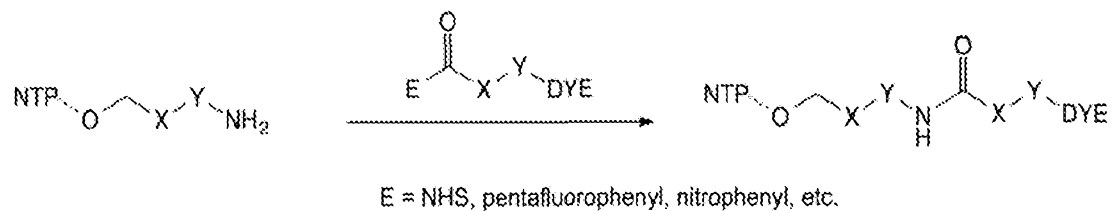


Fig. 3H

eolf-seq1
SEQUENCE LISTING

<110> FUNDACIO INSTITUT D'INVESTIGACIO EN CIENCIES DE LA SALUT
GERMANS TRIAS I PUJOL

<120> HF prognostic marker

<130> P3130PC00

<150> EP14182846.7

<151> 2014-08-29

<160> 1

<170> PatentIn version 3.5

<210> 1

<211> 698

<212> PRT

<213> Homo sapiens

<400> 1

Asp Asp Gly Ile Cys Lys Ser Ser Asp Cys Ile Lys Ser Ala Ala Arg
1 5 10 15

Leu Ile Gln Asn Met Asp Ala Thr Thr Glu Pro Cys Thr Asp Phe Phe
20 25 30

Lys Tyr Ala Cys Gly Gly Trp Leu Lys Arg Asn Val Ile Pro Glu Thr
35 40 45

Ser Ser Arg Tyr Gly Asn Phe Asp Ile Leu Arg Asp Glu Leu Glu Val
50 55 60

Val Leu Lys Asp Val Leu Gln Glu Pro Lys Thr Glu Asp Ile Val Ala
65 70 75 80

Val Gln Lys Ala Lys Ala Leu Tyr Arg Ser Cys Ile Asn Glu Ser Ala
85 90 95

Ile Asp Ser Arg Gly Gly Glu Pro Leu Leu Lys Leu Leu Pro Asp Ile
100 105 110

Tyr Gly Trp Pro Val Ala Thr Glu Asn Trp Glu Gln Lys Tyr Gly Ala
115 120 125

Ser Trp Thr Ala Glu Lys Ala Ile Ala Gln Leu Asn Ser Lys Tyr Gly
130 135 140

Lys Lys Val Leu Ile Asn Leu Phe Val Gly Thr Asp Asp Lys Asn Ser
145 150 155 160

Val Asn His Val Ile His Ile Asp Gln Pro Arg Leu Gly Leu Pro Ser
165 170 175

Arg Asp Tyr Tyr Glu Cys Thr Gly Ile Tyr Lys Glu Ala Cys Thr Ala
Page 1

eolf-seql

180	185	190
Tyr Val Asp Phe Met Ile Ser Val Ala Arg Leu Ile Arg Gln Glu Glu	195	200
Arg Leu Pro Ile Asp Glu Asn Gln Leu Ala Leu Glu Met Asn Lys Val	210	215
Met Glu Leu Glu Lys Glu Ile Ala Asn Ala Thr Ala Lys Pro Glu Asp	225	230
Arg Asn Asp Pro Met Leu Leu Tyr Asn Lys Met Thr Leu Ala Gln Ile	245	250
Gln Asn Asn Phe Ser Leu Glu Ile Asn Gly Lys Pro Phe Ser Trp Leu	260	265
Asn Phe Thr Asn Glu Ile Met Ser Thr Val Asn Ile Ser Ile Thr Asn	275	280
Glu Glu Asp Val Val Val Tyr Ala Pro Glu Tyr Leu Thr Lys Leu Lys	290	295
Pro Ile Leu Thr Lys Tyr Ser Ala Arg Asp Leu Gln Asn Leu Met Ser	305	310
Trp Arg Phe Ile Met Asp Leu Val Ser Ser Leu Ser Arg Thr Tyr Lys	325	330
Glu Ser Arg Asn Ala Phe Arg Lys Ala Leu Tyr Gly Thr Thr Ser Glu	340	345
Thr Ala Thr Trp Arg Arg Cys Ala Asn Tyr Val Asn Gly Asn Met Glu	355	360
Asn Ala Val Gly Arg Leu Tyr Val Glu Ala Ala Phe Ala Gly Glu Ser	370	375
Lys His Val Val Glu Asp Leu Ile Ala Gln Ile Arg Glu Val Phe Ile	385	390
Gln Thr Leu Asp Asp Leu Thr Trp Met Asp Ala Glu Thr Lys Lys Arg	405	410
Ala Glu Glu Lys Ala Leu Ala Ile Lys Glu Arg Ile Gly Tyr Pro Asp	420	425
Asp Ile Val Ser Asn Asp Asn Lys Leu Asn Asn Glu Tyr Leu Glu Leu	435	440
Asn Tyr Lys Glu Asp Glu Tyr Phe Glu Asn Ile Ile Gln Asn Leu Lys		

eolf-seql

450 455 460

Phe Ser Gln Ser Lys Gln Leu Lys Lys Leu Arg Glu Lys Val Asp Lys
465 470 475 480

Asp Glu Trp Ile Ser Gly Ala Ala Val Val Asn Ala Phe Tyr Ser Ser
 485 490 495

Gly Arg Asn Gln Ile Val Phe Pro Ala Gly Ile Leu Gln Pro Pro Phe
 500 505 510

Phe Ser Ala Gln Gln Ser Asn Ser Leu Asn Tyr Gly Gly Ile Gly Met
 515 520 525

Val Ile Gly His Glu Ile Thr His Gly Phe Asp Asp Asn Gly Arg Asn
 530 535 540

Phe Asn Lys Asp Gly Asp Leu Val Asp Trp Trp Thr Gln Gln Ser Ala
545 550 555 560

Ser Asn Phe Lys Glu Gln Ser Gln Cys Met Val Tyr Gln Tyr Gly Asn
 565 570 575

Phe Ser Trp Asp Leu Ala Gly Gly Gln His Leu Asn Gly Ile Asn Thr
 580 585 590

Leu Gly Glu Asn Ile Ala Asp Asn Gly Gly Leu Gly Gln Ala Tyr Arg
 595 600 605

Ala Tyr Gln Asn Tyr Ile Lys Lys Asn Gly Glu Glu Lys Leu Leu Pro
610 615 620

Gly Leu Asp Leu Asn His Lys Gln Leu Phe Phe Leu Asn Phe Ala Gln
625 630 635 640

Val Trp Cys Gly Thr Tyr Arg Pro Glu Tyr Ala Val Asn Ser Ile Lys
 645 650 655

Thr Asp Val His Ser Pro Gly Asn Phe Arg Ile Ile Gly Thr Leu Gln
 660 665 670

Asn Ser Ala Glu Phe Ser Glu Ala Phe His Cys Arg Lys Asn Ser Tyr
 675 680 685

Met Asn Pro Glu Lys Lys Cys Arg Val Trp
690 695