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

(21) **International Application Number:**
PCT/US2012/024708

(22) **International Filing Date:**
10 February 2012 (10.02.2012)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
61/441,945   11 February 2011 (11.02.2011)   US

(71) **Applicant** *(for all designated States except US)*: **NABSYS, INC.** [US/US]; 60 Clifford Street, Providence, RI 02903 (US).

(72) **Inventor; and**
(75) **Inventor/Applicant** *(for US only)*: **OLIVER, John, S.** [US/US]; 35 Burton Street, Bristol, RI 02809 (US).

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

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

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title:** ASSAY METHODS USING DNA BINDING PROTEINS

(57) **Abstract:** Assay methods for preparing a biomolecule analyte includes hybridizing a sequence specific oligonucleotide probe to a biomolecule template and reacting the resulting analyte with a binding moiety.

ASSAY METHODS USING DNA BINDING PROTEINS

Cross-reference to Related Application

[0001]   This application claims priority to and the benefit of U.S. Provisional Patent Application No. 61/441,945, filed February 11, 2011, which is incorporated herein by reference in its entirety.

Field of Invention

5   [0002]   The present invention relates generally to assay methods for the analysis of biopolymers. Mapping and sequencing of such biopolymers is contemplated herein.

Background

[0003]   A number of different approaches for sequencing nucleic acids exist. The traditional methods are the dideoxy-chain termination method described by Sanger et al., Proc

10   Natl. Acad. Sci. USA, (1977) 74: 5463-67 and the chemical degradation method described by Maxam et al., Proc. Natl. Acad. Sci. USA, (1977) 74: 560-564. Of these two methods, the Sanger procedure has been the most widely used. The original Sanger method relied on radioactive labeling of the reaction products and separation of the reaction products by slab gel electrophoresis.

15   [0004]   Both the Sanger and Maxam methods are time- and labor-intensive. The start of the Human Genome Project was the impetus for the development of improved, automated systems to perform Sanger sequencing. As a result, detection of fluorescence has replaced autoradiography and capillary electrophoresis has replaced the ultrathin slab gels originally used to separate reaction products. Automated sequencers have been developed and are

20   capable of processing large numbers of samples without operator intervention.

[0005]   The completion of the Human Genome Project has refocused the need for new technologies that are capable of rapidly and inexpensively determining the sequence of human and other genomes. There is has been much discussion in recent years about personalized medicine. The vision of personalized medicine involves each individual having his or her

- 2 -

complete genome sequenced at high accuracy and using this information to guide clinical care, specifically for risk stratification of patients and pharmacogenomics.

[0006]   In recent years, a number of technological advances have been developed enabling a great reduction in the cost of sequencing and substantially increasing the amount of
5   sequence data produced. Most sequencing methods currently available utilize optical detection for the determination of the DNA sequence. The most prevalent sequencing methods are referred to as sequencing by synthesis (SBS).

[0007]   Typical SBS methods involve the stepwise synthesis of a strand of DNA that is complementary to a template sequence from the target genome to be sequenced. The SBS
10   methods can be divided into those that are performed in batch mode and those that are performed in real-time. The batch mode processes rely on the stepwise synthesis of the new DNA strand with the limitation that the synthesis is only allowed to proceed for one nucleotide position, for one nucleotide type, or for the combination of one nucleotide position and one nucleotide type. The incorporation of the nucleotide occurs in parallel for large numbers of
15   templates. Detection is achieved using a variety of methods.

[0008]   A batch mode SBS method utilizing a single nucleotide type is used by Roche for pyrosequencing with the 454 platform. The Roche technology (see, e.g., Margulies et al. (2005) Nature, 437:376-380; U.S. Patent Nos. 6,274,320; 6,258,568; 6,210,891) utilizes pyrosequencing. The method depends on several enzymes and cofactors to produce
20   luminescence when a nucleotide is incorporated. A single nucleotide species is introduced into a large number of small reaction vessels each containing multiple copies of a single template. The incorporation of the nucleotide is accompanied by light emission. When the reaction has run to completion, the reagents are washed from the reaction volumes and a next nucleotide and its required reagents are washed into the reactions. Each template is thus extended in an
25   iterative fashion, one nucleotide at a time. Multiple incorporations of the same nucleotide require the quantitative determination of the amount of light emitted. Homopolymer tracts in templates may be difficult to accurately sequence as the incremental amount of light emitted for each subsequent position in the homopolymer becomes small compared to the total amount emitted.

30   [0009]   In a second SBS method, platforms by Helicos (see, e.g., Quake et al Proc. Nat. Acad. Sci. USA (2003) 100: 3960-3964; U.S. Patent Nos. 6,818,395; 6,911,345; 7,297,518; 7,462,449 and 7,501,245), Illumina (see, e.g., Bennett et al. Pharmacogenomics (2005) 6:373-

382), and Intelligent Bio-Systems (see, e.g., Ju et al. Proc. Nat. Acad. Sci. USA (2006) 103:19635-19640) allow only the incorporation of a single nucleotide at each step. Template strands are attached to a solid support and a primer sequence is annealed. A polymerase is used to extend the primer to make a complement to the template. The nucleotides are derivatized such that after the incorporation of a single nucleotide, the growing strand is incapable of further extension. The nucleotides are further derivatized to make them fluorescent. In the Helicos technology, the four nucleotides are labeled with the same fluorescent tag. This requires that each nucleotide type be added separately. In contrast, the Illumina and Intelligent Bio-Systems technologies utilize four different fluorescent tags so that a mixture of all four derivatized nucleotides may be added at the same time. For both technologies, the incorporation of a nucleotide is accompanied by the appearance of fluorescence in the growing strand. In the case of Illumina, the wavelength of the fluorescence emission indicates the identity of the newly incorporated nucleotide. In the Helicos technology, only a single nucleotide type is added at each cycle. Thus, the appearance of fluorescence at a position on the solid support indicates the incorporation of the added nucleotide for that template. Templates that do not incorporate the nucleotide present in the reaction remain dark.

[0010]   Following the observation of any incorporated fluorescence, the blocking groups and fluorescent tags are removed prior to the next cycle. Multiple cycles result in the acquisition of sequence data for many templates in a single run. The instrumentation typical for these technologies is said to allow for the automated acquisition of sequence information for hundreds of thousands to millions of templates in parallel.

[0011]   SBS methods may also be performed in real-time. In real-time SBS, polymerase is used to incorporate fluorescently labeled nucleotides and the fluorescence is observed during DNA strand synthesis. The four nucleotides are labeled with different fluorescent tags. The fluorescent tags are attached to the terminal phosphate of the nucleotide triphosphate. During incorporation of the nucleotide into the growing strand the fluorophore is released to solution and the growing strand remains non-fluorescent. The identity of the incorporated strand is determined while the nucleotide resides in the active site of the enzyme and before the cleaved diphosphate is released to bulk solution.

[0012]   The fluorescence of the incorporated nucleotide typically is measured in a background fluorescence from a much larger concentration of unincorporated nucleotide. Pacific Biosciences (see, e.g., US Patent Nos. 7,170,050; 7,302,146; 7,315,019; 7,476,503; and

- 4 -

7,476,504) identifies the incorporated nucleotide based on the residence time in the polymerase active site. Fluorescence emission from the active site for an appropriate time indicates incorporation and the emission wavelength determines the identity of the incorporated nucleotide. Polymerase is attached to the bottom of zero-mode waveguides. Zero-mode waveguides are reaction cells whose dimensions limit the passage of light from the excitation sources. Thus, only fluorescent tags close to the bottom surface of the reaction volume are excited.

[0013]   Visigen identifies the incorporated nucleotide through Fluorescent Resonant Energy Transfer (FRET) between an acceptor in the polymerase active site and a fluorescent tag on the nucleotide (see, e.g., US Patent Nos. 7,211,414 and 7,329,492). Only nucleotides held in the active site of the polymerase show fluorescence. Incorporation is identified by the residence time of the fluorescence in the active site and the nucleotide identity is determined by the emission wavelength.

[0014]   Other recently developed methods to sequence DNA rely on hybridization and ligation. Both the SOLiD and Complete Genomics technologies rely on the combination of hybridization and ligation. The SOLiD system (Life Technologies) immobilizes short template strands via an adapter. A primer and a pool of labeled oligonucleotides containing two fixed positions and six degenerate positions is hybridized to the template. The primer hybridizes to the adaptor. Each pool consists of 16,384 different sequences. Four fluorescent dyes are used to label the oligonucleotides in a pool in a fashion that creates four subsets from the sixteen combinations at the two fixed positions. Thus, each fluorescent tag is associated with four of the sixteen possible combinations. Following hybridization, a ligase is added and any probes in the pool that hybridized contiguously with the primer are ligated to the primer. The fluorescence of the hybridized and ligated product is determined. The fluorescence defines which subset of sequences hybridized to the template and ligated to the primer. The terminal three bases and the associated fluorescent tag are cleaved from the hybridized and ligated oligonucleotide. Subsequent rounds of another round of hybridization, ligation, and cleavage are performed. In this first series of reactions, each cycle identifies a subset for the pair of nucleotides in the template that is 5 nucleotides downstream from subset of pairs that were identified in the last cycle. After several cycles, the primer, and the oligonucleotides that have been ligated to it, is washed off the template.

- 5 -

[0015]   The entire procedure is repeated starting with a primer that is one nucleotide shorter than the original primer, then with primers that are two, three, and four nucleotides shorter than the original primer. These subsequent rounds shift the frame of interrogation so that the bases that make-up the template strand can be identified from the union between the two subsets of reaction that overlapped at that position.

[0016]   Complete Genomics technology utilizes a similar hybridization and ligation method (see, e.g., US Patent Application Publication Nos. 20080234136; 20090005252; 20090011943; and 20090176652). In the Complete Genomics technology, a primer is hybridized to an adaptor that is attached to the end of the template. A series of pools of oligonucleotides is constructed. In each pool, the nucleotide at a single position is identified by using four-color fluorescence. The remaining positions are degenerate. The first pool is hybridized to the template. Oligonucleotides that hybridize adjacent to the primer are subsequently ligated. After washing excess oligonucleotides away, the fluorescence of the ligated oligonucleotide identifies the nucleotide at the defined position in that pool. The ligated primer and oligonucleotide are washed off the template and the process is repeated with the next pool of oligonucleotides that probe the next position down from the primer.

[0017]   The SBS and hybridization-ligation methods generate short pieces or reads of DNA sequence. While the short reads can be used to re-sequence human genomes, they are not favorable for the *de novo* assembly of human genomes. With the recent realization that human genomes contain large numbers of inversions, translocations, duplications, and indels (e.g., mutations that include both insertions, deletions, and the combination thereof), the quality of human genome data from short reads is even more suspect. Genetic rearrangements are even more prevalent in cancer.

[0018]   While variations of the short read technologies that incorporate paired-end reads have been proposed and the length of the sequence data from these technologies has increased incrementally over the last two years, it is clear that longer read technologies are necessary for the accurate assembly of human genome data.

[0019]   In addition to the undesirable nature of short reads, all of the DNA sequencing methods described above employ optical detection. The throughput of optical methods limits the ultimate performance characteristics of any of these sequencing technologies. Optical methods are capable of identifying single molecules. However, the time required to observe and accurately identify events is typically too slow to meet the need for higher throughput.

- 6 -

While the current generation of sequencing technologies has lowered the cost of sequencing by orders of magnitude in comparison to the methods used to sequence the first human genomes, the methods remain too slow, costly, and inaccurate for routine analysis of human genomes.

[0020]   A need therefore exists for efficient methods and devices capable of rapid and accurate nucleic acid sequencing for *de novo* assembly of human genomes. It is desirable to have long read lengths and to use as little nucleic acid template as possible. Moreover, single-molecule optical detection of DNA has limitations with respect to sensitivity and speed.

[0021]   Thus, there remains a need for improved methods and devices for the analysis of biopolymers, including methods and devices for mapping and sequencing such biopolymers.

Summary

[0022]   Embodiments of the invention provide assay methods for preparing analyte samples for mapping and sequencing using nanopore, micro-channel or nano-channel analysis devices.

[0023]   In an aspect, an embodiment of the invention includes a method for preparing a biomolecule analyte. The method includes: a) providing a single-stranded DNA or RNA template; b) hybridizing a first plurality of identical, sequence specific oligonucleotide probes to the template, each probe having a 5' end and a 3' end, to thereby form an analyte having at least one single-stranded region and at least two duplex regions; c) conducting a base extension reaction in the at least one single-stranded region from the 3' end of a hybridized probe toward the 5' end of an adjacent hybridized probe; d) terminating the base-extension reaction such that there remains for each single-stranded region a single-stranded portion thereof adjacent to the 5' end of each hybridized probe, and e) reacting the resulting analyte with a binding moiety that selectively binds to the at least one single-stranded portion to thereby prepare the biomolecule analyte.

[0024]   One or more of the following features may be included. The probes may include single-stranded DNA and/or RNA. The base extension reaction may be performed by a DNA or RNA polymerase. The binding moiety may include a protein, such as RecA, T4 gene 32 protein, f1 geneV protein, human replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding protein, and/or *E. coli* single-stranded binding protein. A length of each probe may be selected from a range of 4 to 12 bases.

- 7 -

[0025]   Steps a – e may be performed sequentially. Steps a – e may be repeated sequentially by replacing the first plurality of probes with a subsequent plurality of different unique probes. At least a portion of the probes in the first plurality of probes has attached thereto a detectable tag.

[0026]   The biomolecule analyte may be configured for detection of positional information in a nanopore system. Changes in an electrical property across a nanopore may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety. Regions of the biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

[0027]   The biomolecule analyte may be configured for detection of positional information in a fluidic channel system. The fluidic channel system may include a micro-channel or a nano-channel. Changes in an electrical property across a fluidic channel may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety. Regions of the biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

[0028]   A sequence of at least a portion of the single-stranded DNA or RNA template may be determined according to the determined binding moiety locations.

[0029]   Sequence-specific oligonucleotide analog probes such as LNAs, PNAs or 2'-methoxy nucleotide analogs may be substituted for the sequence specific oligonucleotide probes. The first plurality of identical, sequence specific oligonucleotide probes may be replaced by a pool of sequence specific oligonucleotide probes including at least a first plurality of identical, sequence specific oligonucleotide probes and a second plurality of identical, sequence specific oligonucleotide probes, with the probes of the second plurality being different from the probes of the first plurality.

[0030]   In another aspect, embodiments of the invention includes a method for preparing a biomolecule analyte by: a) providing a single-stranded DNA template; b) hybridizing a first plurality of identical, sequence specific RNA probes to the template, each probe having a 5' end and a 3' end, to thereby form an analyte having at least one single-stranded region and at

- 8 -

least one duplex region; c) conducting a base extension reaction in the at least one single-stranded region from the 3' end of a hybridized probe; d) allowing the base-extension reaction to fill each single-stranded region on the analyte; e) removing the RNA probes to provide the analyte with at least one single-stranded segment in the region to which an RNA probe had been hybridized; and f) reacting the resulting analyte with a binding moiety that selectively binds to the single-stranded segment, to thereby prepare the biomolecule analyte.

[0031]   One or more of the following features may be included. The base extension reaction may be performed by a DNA or RNA polymerase. The binding moiety may include a protein, such as RecA, T4 gene 32 protein, f1 geneV protein, human replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding protein, and/or *E. coli* single-stranded binding protein. A length of each probe may be selected from a range of 4 to 12 bases.

[0032]   Steps a – f may be performed sequentially. Steps a – f may be repeated sequentially by replacing the first plurality of probes with a subsequent plurality of different unique probes.

[0033]   The RNA probes may be removed by reacting the analyte with hydroxyl ions.

[0034]   The biomolecule analyte may be configured for detection of positional information in a nanopore system. Changes in an electrical property across a nanopore may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety. Regions of the biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

[0035]   The biomolecule analyte may be configured for detection of positional information in a fluidic channel system. The fluidic channel system may include a micro-channel or a nano-channel. Changes in an electrical property across a fluidic channel may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety. Regions of the biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

[0036]   A sequence of at least a portion of the single-stranded DNA or RNA template may be determined according to the determined binding moiety locations.

[0037]   The first plurality of identical, sequence specific RNA probes may be replaced by a pool of sequence specific RNA probes including at least a first plurality of identical, sequence specific RNA probes and a second plurality of identical, sequence specific RNA probes, the probes of the second plurality being different from the probes of the first plurality.

[0038]   In yet another aspect, embodiments of the invention include a method for preparing a biomolecule analyte. The method includes: a) providing a single-stranded DNA or RNA template; b) hybridizing a first plurality of identical, sequence specific oligonucleotide probes to the template, to thereby form an analyte having at least one single-stranded region and at least one duplex region; and c) reacting the resulting analyte with a binding moiety that selectively binds to the at least one duplex region to thereby prepare the biomolecule analyte.

[0039]   One or more of the following features may be included. The probes may include single-stranded DNA and/or RNA. The binding moiety may include a protein, such as RecA, T4 gene 32 protein, f1 geneV protein, human replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding protein, and/or *E. coli* single-stranded binding protein. A length of each probe may be selected from a range of 4 to 12 bases.

[0040]   Steps a – c may be performed sequentially. Steps a – c may be repeated sequentially by replacing the first plurality of probes with a subsequent plurality of different unique probes. At least a portion of the probes in the first plurality of probes has attached thereto a detectable tag.

[0041]   The biomolecule analyte may be configured for detection of positional information in a nanopore system. Changes in an electrical property across a nanopore may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety. Regions of the biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

[0042]   The biomolecule analyte may be configured for detection of positional information in a fluidic channel system. The fluidic channel system may include a micro-channel or a nano-channel. Changes in an electrical property across a fluidic channel may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical

- 10 -

property being indicative of regions including or lacking the binding moiety. Regions of the biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

5   **[0043]**   A sequence of at least a portion of the single-stranded DNA or RNA template may be determined according to the determined binding moiety locations.

[0044]   Sequence specific oligonucleotide analog probes, such as LNAs, PNAs and 2'-methoxy nucleotide analogs, may be substituted for the sequence specific oligonucleotide probes. The first plurality of identical, sequence specific oligonucleotide probes may be

10 replaced by a pool of sequence specific oligonucleotide probes including at least a first plurality of identical, sequence specific oligonucleotide probes and a second plurality of identical, sequence specific oligonucleotide probes, the probes of the second plurality being different from the probes of the first plurality.

[0045]   In another aspect, embodiments of the invention include a method for preparing a

15 biomolecule analyte. The method includes: a) providing a single-stranded DNA or RNA template; b) hybridizing a first plurality of identical, sequence specific oligonucleotide probes to the template, to thereby form an analyte having at least one single-stranded region and at least one duplex region; and c) reacting the resulting analyte with a binding moiety that selectively binds to the at least one single-stranded region, to thereby prepare the biomolecule

20 analyte.

[0046]   One or more of the following features may be included. The probes may include single-stranded DNA and/or RNA. The binding moiety may include a protein, such as RecA, T4 gene 32 protein, f1 geneV protein, human replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding protein, and/or *E. coli* single-stranded binding protein. A

25 length of each probe may be selected from a range of 4 to 12 bases.

[0047]   Steps a – c may be performed sequentially. Steps a – c may be repeated sequentially by replacing the first plurality of probes with a subsequent plurality of different unique probes. At least a portion of the probes in the first plurality of probes has attached thereto a detectable tag.

30 **[0048]**   The biomolecule analyte may be configured for detection of positional information in a nanopore system. Changes in an electrical property across a nanopore may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical

- 11 -

property being indicative of regions including or lacking the binding moiety. Regions of the biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

5   [0049]   The biomolecule analyte may be configured for detection of positional information in a fluidic channel system. The fluidic channel system may include a micro-channel or a nano-channel. Changes in an electrical property across a fluidic channel may be monitored as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety. Regions of the

10   biomolecule analyte including or lacking the binding moiety may be differentiated based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

[0050]   A sequence of at least a portion of the single-stranded DNA or RNA template may be determined according to the determined binding moiety locations.

15   [0051]   Sequence specific oligonucleotide analog probes such as LNAs, PNAs and 2'-methoxy nucleotide analogs may be substituted for the sequence specific oligonucleotide probes. The first plurality of identical, sequence specific oligonucleotide probes may be replaced by a pool of sequence specific oligonucleotide probes including at least a first plurality of identical, sequence specific oligonucleotide probes and a second plurality of identical,

20   sequence specific oligonucleotide probes, the probes of the second plurality being different from the probes of the first plurality.

[0052]   In some embodiments, the probes are single-stranded DNA or RNA, and they may optionally be provided with tags that enhance detection in analysis devices. Based on the nature of the tags, the binding moiety may bind to the tags as well as to regions of the analyte,

25   further enhancing detection. Alternatively, the probes may include oligonucleotide analogs such as Locked Nucleic Acids (LNAs), Peptide Nucleic Acids (PNA s) or 2'-methoxy nucleotide analogs. Base extension reactions are performed using DNA polymerase or RNA polymerase, as appropriate.

[0053]   In any of the above-described embodiments, one or more of the features

30   described in this paragraph and in the paragraphs that follow may be included. The process steps may be repeated using subsequent pluralities of matching probes having a known binding selectivity different from the known binding selectivity of the first plurality of probes. The first

- 12 -

plurality of probes may be hybridizing oligonucleotides having n number of bases, with n preferably ranging from 4 to 12. Each process may be repeated sequentially by replacing the first plurality of probes with a subsequent plurality of each of the different unique probes within the entire library of 4^n n-mer probes. A portion of the library of 4^n n-mer probes may be

5   used, such as about 85%, 75%, 65%, 55%, 45%, or 33% of the library. The sequential repetition of the process may be conducted in a linear or parallel series of reactions.

[0054]   Alternatively, rather than employing a single plurality of identical, sequence specific oligonucleotide probes, a pool of probes comprising a first plurality of identical, sequence specific oligonucleotide probes and a second plurality of identical, sequence specific

10   oligonucleotide probes, different from the first plurality, may be employed.

[0055]   The biomolecule to be analyzed may be DNA or RNA. The binding moiety may be a protein. Examples of suitable proteins include RecA, T4 gene 32 protein, f1 geneV protein, human replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding protein, and *E. coli* single-stranded binding protein.

15   [0056]   The products of the assay methods described herein may be analyzed in an apparatus using a nanopore or a fluidic channel such as a micro-channel or nano-channel to detect probes or probe positions on the biomolecule being analyzed. Suitable apparatus are described, for example, in U.S. Patent Application Serial Number 12/789,817, published as U.S. Publication No. 2010/0310421, incorporated herein by reference in its entirety. Such

20   systems use electrical detection methods to determine the presence and/or the relative position of the oligonucleotide probes. The products of the assay methods described herein may be sequenced in accordance with sequencing methods such as those described in U.S. Patent Application Serial Number 13/292,415, incorporated herein by reference in its entirety.

[0057]   An electrical property across a nanopore or fluidic channel as the analyte

25   translocates therethrough may be monitored, with changes in the electrical property being indicative of regions on the analyte including or lacking the binding moiety. Furthermore, differentiating between regions on the analyte including or lacking the binding moiety, allows a determination of binding moiety locations on the analyte.

Brief Description of the Figures

30   [0058]   FIG. 1 is a schematic depiction of a DNA molecule.

[0059]   FIG. 2 is a schematic depiction of an RNA molecule.

- 13 -

[0060]   FIG. 3 is a schematic depiction of a hybridizing oligonucleotide probe.

[0061]   FIG. 4 is a schematic depiction of a single-stranded DNA molecule hybridized with two identical probes.

[0062]   FIGs. 5(a) – 5(d) are a schematic depiction of an assay method in accordance with an embodiment of the invention in which single-stranded DNA ("ssDNA") probes are bound to a single-stranded DNA or RNA analyte, a base extension reaction is carried out, and a binding moiety which binds to single-stranded portions of the resulting analyte is employed.

[0063]   FIGs. 6(a) – 6(d) are a schematic depiction of an assay method in accordance with an embodiment of the invention in which tagged ssDNA probes are bound to a single-stranded DNA or RNA analyte, a base extension reaction is carried out, and a binding moiety which binds to single-stranded portions of the resulting analyte is employed.

[0064]   FIGs. 7(a) – 7(f) are a schematic depiction of an assay method in accordance with an embodiment of the invention in which RNA probes are bound to a single-stranded DNA analyte, a base extension reaction is carried out, the RNA probes are removed, and a binding moiety which binds to single-stranded portions of the resulting analyte is employed.

[0065]   FIGs. 8(a) – 8(c) are a schematic depiction of an assay method in accordance with an embodiment of the invention in which oligonucleotide probes are bound to a single-stranded DNA or RNA analyte, and a binding moiety is bound to duplex structures formed by the probes.

[0066]   FIGs. 9(a) – 9(c) are a schematic depiction of an assay method in accordance with an embodiment of the invention in which tagged oligonucleotide probes are bound to a single-stranded DNA or RNA analyte, and a binding moiety is bound to duplex structures formed by the probes.

[0067]   FIGs. 10(a) – 10(c) are a schematic depiction of an assay method in accordance with an embodiment of the invention in which oligonucleotide probes are bound to a single-stranded DNA or RNA analyte, and a binding moiety is bound to the analyte in gaps between the probes.

[0068]   FIG. 11(a) is a schematic depiction of an assay method in accordance with an embodiment of the invention showing a DNA molecule having a binding moiety region in a nanopore apparatus.

- 14 -

[0069]   FIG. 11(b) is a schematic depiction of an assay method in accordance with an embodiment of the invention showing a current measurement waveform as a DNA molecule having a binding moiety region translocates through the nanopore apparatus of FIG. 11(a).

[0070]   FIG. 12 is a schematic depiction of an assay method in accordance with an embodiment of the invention showing a nano-channel or micro-channel apparatus useful for mapping the analytes of the present invention.

[0071]   FIG. 13(a) is a schematic depiction of an assay method in accordance with an embodiment of the invention showing an electrical potential measurement as a DNA molecule having a binding moiety region enters a detection volume in the apparatus of FIG. 12.

[0072]   FIG. 13(b) is a schematic depiction of an assay method in accordance with an embodiment of the invention showing an electrical potential measurement as a binding moiety region on a DNA molecule enters a detection volume in the apparatus of FIG. 12.

[0073]   FIG. 13(c) is a schematic depiction of an assay method in accordance with an embodiment of the invention showing an electrical potential measurement as a binding moiety region on a DNA molecule exits a detection volume in the apparatus of FIG. 12.

[0074]   FIG. 13(d) is a schematic depiction of an assay method in accordance with an embodiment of the invention showing an electrical potential measurement as a DNA molecule having a binding moiety region exits a detection volume in the apparatus of FIG. 12.

[0075]   FIG. 14 is a schematic depiction of an assay method in accordance with an embodiment of the invention showing a nano-channel or micro-channel apparatus having multiple detection volumes.

Detailed Description

[0076]   As used in this description and the accompanying claims, the following terms shall have the meanings given, unless the context indicates otherwise:

[0077]   A "biomolecule analyte" or "target analyte" means a biomolecule, for example, having sequence information that is to be determined using embodiments of the present invention. The target may be a biomolecule such as, for example, deoxyribonucleic acid, a ribonucleic acid, a protein, or a polypeptide. The target may be single-stranded or double-stranded.

[0078]   A "probe" means any molecule or assembly of molecules capable of sequence-specific covalent or non-covalent binding to an analyte.

- 15 -

[0079]   A "tag" means a moiety that is attached to a probe in order to make the probe more visible to a detector. These tags may be, for example, proteins, double-stranded DNA, single-stranded DNA or other molecules. Exemplary tags that will bind RecA include single-stranded or double-stranded DNA. Another useful tag may be a dendrimer, for example. Optionally, tags have either a larger volume than the probe or a different charge so that they slow translocation of the biomolecule through a nanopore, or through a nano-channel or micro-channel.

[0080]   Disclosed herein are methods that increase the signal-to-noise ratio in translocation detection of biomolecules that have been hybridized to probes. In one embodiment, a single-stranded biomolecule may be hybridized with a probe. The hybridized biomolecule may then be incubated with a protein or enzyme that binds to the biomolecule and forms at least a partial coating along the biomolecule.

[0081]   Coated biomolecules typically have greater uniformity in their translocation rates, which leads to a decrease in positional error and thus more accurate sequencing. Due to its increased diameter, a coated biomolecule generally translocates through a sequencing system at a slower speed than a non-coated biomolecule. The translocation is preferably slow enough so that a signal can be detected during its passage from a first chamber into a second chamber. The translocation rate or frequency may be regulated by introducing a salt gradient between the chambers. Exemplary salt concentration ratios of the cis to the trans side of the chamber may include, but are not limited to, 1:2, 1:4, 1:6, and 1:8. For example, salt concentrations may range from about 0.5 M KCl to about 1 M KCl on the cis side and from about 1 M KCl to about 4 M KCl on the trans side. The signal is preferably strong enough to be detected using known methods or methods described herein. Exemplary signal-to-noise ratios include, but are not limited to, 2:1, 5:1, 10:1, 15:1, 20:1, 50:1, 100:1, and 200:1. With a higher signal-to-noise ratio, a lower voltage may be used to effect translocation.

[0082]   In one embodiment, a biomolecule of interest is hybridized with the entire library of probes of a given length. For example, the biomolecule of interest can be hybridized with the entire universe of 4096 (i.e., 4^6) possible six-mers. The hybridization can be done sequentially (i.e., one probe after another) or in parallel (i.e., a plurality of biomolecules of interest are each separately hybridized simultaneously with each of the possible probes.) Alternatively, the probes can be separated from each other in both space and time.

- 16 -

Additionally, more than one probe type may be hybridized to the same biomolecule of interest at the same time.

[0083]   The set of probes used to perform the sequencing may be a subset of the complete library of probes of a given length, such as about 85%, 75%, 65%, 55% , 45%, or 33% of the library. For instance, if sequencing is performed on a biomolecule that starts as double-stranded DNA, then only one-half of the probes that make up a library may be needed. Other subsets of the library may be designed to allow sequencing as well. If some information concerning the target sequence is known prior to performing the sequencing reaction, it may be possible to use a small subset of the total library. For instance, if the sequencing reaction is being performed to determine if single nucleotide polymorphisms are present with respect to a reference sequence, then a small number of probes with respect to the complete library may be used. Alternatively, the set of probes may not all be the same length. In an embodiment, a set of at least two probes may be used for hybridization, rather than an entire library of probes or subset thereof. In another embodiment, probes may be separated by (GC) content or other determinants of probe binding strength, in order to allow for optimization of reaction conditions. By separating the probes based on relative properties, multiple probes may be incorporated into a single hybridization reaction. Further, the probes may be grouped based on their related optimum reaction environment preferences. In yet another embodiment, pools of probes may be simultaneously hybridized to a biomolecule of interest. A pool of probes is a group of probes of different composition, each of which may likely be present in many copies. The composition of the probes may be chosen so as to reduce the chance of competitive binding to the biomolecule of interest. Alternatively, the composition of multiple pools may be chosen so that the same competitive binding is not present in all pools occupied by a single probe.

[0084]   In still another embodiment, the probes may be attached to tags, making electrical fluctuations more noticeable as the hybridized probes translocate through the sequencing system. In addition, different tags may be used to help distinguish among the different probes. These tags may be proteins, double-stranded DNA, single-stranded DNA or other molecules.

[0085]   It should be understood that the invention is not intended to be limited strictly to DNA and RNA oligonucleotide probes. Rather, it is envisioned that oligonucleotide analog probes such as those comprising LNAs, PNAs and 2'-methoxy nucleotide analogs may be used as well.

- 17 -

[0086]   In embodiments of the present invention, the biomolecule may be hybridized with sequence-specific probes prior to being reacted with a binding moiety such as a protein. The probes may or may not have tags attached to them. If the probe has an attached tag composed of single- or double-stranded DNA, the binding moiety, such as the protein RecA, may coat (i) the single-stranded target, (ii) the double-stranded regions where hybridization between the single-stranded target and (iii) the probe has occurred, and the tag attached to the probe. Alternatively, the bound probes and associated tags may have a different affinity for the binding moiety than for the biomolecule. If the tags have an essentially equal affinity for the binding moiety, then both the tag and the target may be coated. If the tag or probe has a greater affinity for the binding moiety, selective coating may be achieved. If the tag or probe has a lower affinity for the binding moiety, it may selectively coat regions of the biomolecule that do not have probe bound. Since any region with bound protein will have a larger signal, differentiation of the hybridized and non-hybridized regions allows for greater accuracy in determining the position of hybridization.

[0087]   The translocation of biomolecule / protein complexes through a nanopore, a nano-channel or a micro-channel sequencing system may include detecting an electrical signal indicative of the passage of coated regions. In one embodiment, the signal detected may be formed by passage of the tagged region of the biomolecule through the sequencing system. The time for translocation may be indicative of the length of the biomolecule. The detection step may discriminate between coated, uncoated, or multiply coated regions, as a coated region may have a signal about ten times that of an uncoated region. Increased signal-to-noise may increase confidence for the detection of the probes. Positional information of probe binding to target biomolecule allows for the mapping or sequencing of the biomolecule analyte.

[0088]   In one embodiment, the process of sequencing a biomolecule such as single-strands of DNA or RNA using one or more probes may performed as follows. Suitable processes are also described in U.S. Serial No. 11/538,189, published as U.S. Publication No. 2007/0190542, incorporated herein by reference in its entirety. Referring to Figure 1, a DNA molecule 1 is schematically depicted and is structured in two strands 2, 4 positioned in anti-parallel relation to one another. Each of the two opposing strands 2, 4 may be sequentially formed from repeating groups of nucleotides 6 where each nucleotide 6 consists of a phosphate group, 2-deoxyribose sugar and one of four nitrogen-containing bases. The nitrogen-containing bases include cytosine (C), adenine (A), guanine (G) and thymine (T). DNA strands 2, 4 are

- 18 -

read in a particular direction, from the top (called the 5' or "five prime" end) to the bottom (called the 3' or "three prime" end). Similarly, RNA molecules 8, as schematically depicted in FIG. 2, are polynucleotide chains, which differ from those of DNA 1 by having ribose sugar instead of deoxyribose and uracil bases (U) instead of thymine bases (T).

5   [0089]   Traditionally, in determining the particular arrangement of the bases 6 and thereby the sequences of the molecules, a process called hybridization may be utilized. The hybridization process is the coming together, or binding, of two genetic sequences with one another. This process is predictable because the bases 6 in the molecules do not share an equal affinity for one another. T (or U) bases favor binding with A bases while C bases favor binding
10   with G bases. Binding is mediated via hydrogen bonds that exist between the opposing base pairs. For example, A binds to T (or U) using two hydrogen bonds, while C binds to G using three hydrogen bonds.

[0090]   A hybridizing oligonucleotide, i.e., a probe, may be used to determine and identify the sequence of bases in the molecule of interest. Figure 3 illustrates a probe 10 that is
15   a short DNA sequence having a known composition. Probes 10 may be of any length depending on the number of bases 12 that they include. For example, a probe 10 that includes six bases 12 is referred to as a six-mer probe wherein each of the six bases 12 in the probe 10 may be any one of the known four natural base types A, T(U), C or G. Alternately, the probe may include non-natural bases.

20   [0091]   In this regard, the total number of unique probes 10 in a library is dependent upon the number of bases 12 contained within each probe 10 and the number of different types of bases in the probes. If only the four natural bases are used in probe 10, the total number of probes in the library is determined by the formula 4^n (four raised to the n power) where n is equal to the total number of bases 12 in each probe 10. Formulas for other arrangements or
25   types of bases are well known in the art. Accordingly, the size of the probe library can be expressed as 4^n n-mer probes 10. For the purpose of illustration, in the context of a six-mer probe, the total number of possible unique, identifiable probe combinations includes 4^6 (four raised to the sixth power) or 4096 unique six-mer probes 10. The inclusion of non-natural bases allows for the creation of probes that have spaces or wildcards therein in a manner that
30   expands the versatility of the library, while reducing the number of probes that may be needed to reach the final sequence result. Probes that include universal bases organized into patterns

- 19 -

with natural bases may also be used, for example those described in U.S. Patent Nos. 7,071,324, 7,034,143, and 6,689,563, incorporated herein by reference in their entireties.

[0092]   The process of hybridization using probes 10, as depicted in Figure 4, may begin by denaturing a double-stranded biomolecule, or by starting with a single-stranded

5   biomolecule. Denaturing is accomplished usually through the application of heat or chemicals, such that the hydrogen bonds between adjacent portions of the biomolecule are broken. For example, the hydrogen bonds between the two halves of an original double-stranded DNA may be broken, leaving two single strands of DNA whose bases are now available for hydrogen bonding. After the biomolecule 14 has been denatured, a single-stranded probe 10 may be

10   introduced to the biomolecule 14 to locate portions of the biomolecule 14 that have a base sequence that correlates to the sequence that is found in the probe 10. In order to hybridize the biomolecule 14 with the probe 10, the denatured biomolecule 14 and a plurality of the probes 10 having a known sequence are both introduced into a solution. The solution may be an ionic solution, such as a salt-containing solution. The mixture may be mixed to facilitate binding of

15   the probes 10 to the biomolecule 14 strand along portions thereof that have a matched complementary sequence. Hybridization of the biomolecule 14 using the probe 10 may be accomplished before the biomolecule 14 is introduced into a nanopore sequencing apparatus or after the denatured biomolecule 14 has been placed into the cis chamber of such an apparatus. In this case, after the denatured biomolecule has been added to the cis chamber, buffer solution

20   containing probes 10 with a known sequence is also added to the cis chamber and allowed to hybridize with the biomolecule 14 before the hybridized biomolecule is translocated.

[0093]   Embodiments of the present invention relate to improved methods for the preparation of biomolecule strands for analysis. In one embodiment, shown in FIGs. 5(a) – 5(d), a denatured biomolecule analyte 15 is formed from a single-stranded DNA (ssDNA) or

25   RNA fragment 20 exposed to ssDNA probes 10. As stated above, each probe 10 is a short ssDNA sequence of a known sequence. The probes 10 may be of any length depending on the number of bases that they include. As before, each of the probes is preferably of an identical sequence, thereby causing the probes to selectively hybridize only to portions of the biomolecule fragment 20 that have a complementary sequence. The fragment 20 and probes 10

30   are depicted prior to hybridization in FIG. 5(a). For purposes of clarity in Figures 5-12, probes 10 are shown having a small dot at the 3' end. This dot is not intended to signify a physical structure; rather, it is included in the Figures simply to designate the 3' end of the probe.

[0094]   The biomolecule analyte 15 is shown in FIG. 5(b) following hybridization of the probes 10 to the biomolecule fragment 20. The resulting structure is a biomolecule fragment having, where hybridization has occurred, double-stranded, i.e., duplex, domains. The duplex domains are of a length corresponding to the length of the probes. Thus, in a case where a 6-mer probe is employed, the analyte 10 will comprise a single-stranded biomolecule fragment having a plurality of 6-mer duplex regions, formed by the hybridized probes.

[0095]   Following the hybridizing step, a base extension reaction, such as a primer extension reaction, utilizing for example, a polymerase and one or more nucleotides, is performed as depicted in FIG. 5(c). In such reactions, which form a nucleic acid complementary to a nucleic acid template, a primer complementary to a single-stranded DNA template is typically employed. Starting at the primer, a DNA polymerase may be used to add mononucleotides complementary to the template at the 3' end of the primer. Various base extension reactions will be familiar to those of ordinary skill in the art. Note that if the template includes RNA, an RNA dependent DNA polymerase is employed. Specifically, the hybridized probes 10 are extended from their 3' ends along the biomolecule fragment 20 to create duplex regions 40 on the analyte in gaps that had previously existed between the probe locations. Note, however, that the base extension reaction is intended to be limited in scope. Rather than extending from the 3' end of each probe to the 5' end of an adjacent probe, the base extension reaction may be terminated such that single-stranded segments remain on the biomolecule analyte 15 before the 5' end of each probe 10. The resulting analyte 15 structures, like those of FIG. 5(b), comprise duplex regions alternating with single-stranded regions. However, unlike the structures of FIG. 5(b) that may have relatively large single-stranded gaps between the bound probes 10, the resulting structures are characterized as being primarily duplexes with small single-stranded gaps.

[0096]   Finally, as depicted in FIG. 5(d), a binding moiety 50 which is selective to the single-stranded regions, (i.e., a protein such as E. coli single-stranded DNA binding protein), is reacted with the biomolecule analyte 15 in a manner such that the binding moiety fills the single-stranded gaps in the analyte.

[0097]   The protein RecA from *E. coli* typically binds single- or double-stranded DNA in a cooperative fashion to form filaments containing the DNA in a core and an external sheath of protein (McEntee, K.; Weinstock, G.M.; Lehman, I.R. Binding of the RecA Protein of *Escherichia coli* to Single- and Double-Stranded DNA. *J. Biol. Chem.* **1981**, *256*, 8835,

- 21 -

incorporated herein by reference in its entirety). DNA has a diameter of about 2 nm, while

DNA coated with RecA has a diameter of about 10 nm. The persistence length of the DNA

increases to around 950 nm, in contrast to 0.75 nm for single-stranded DNA or 50 nm for

double-stranded DNA. T4 gene 32 protein is known to cooperatively bind single-stranded

5   DNA (Alberts, B.M.; Frey, L. T4 Bacteriophage Gene32: A Structural Protein in the

Replication and Recombination of DNA. *Nature*, **1970**, *227*, 1313-1318, incorporated herein by

reference in its entirety). *E. coli* single-stranded binding protein binds single-stranded DNA in

several forms depending on salt and magnesium concentrations (Lohman, T.M.; Ferrari, M.E.

Escherichia Coli Single-Stranded DNA-Binding Protein: Multiple DNA-Binding Modes and

10   Cooperativities. *Ann. Rev. Biochem.* **1994**, *63*, 527-570, incorporated herein by reference in its

entirety). The E. coli single-stranded binding protein may form a varied coating on the

biomolecule. The f1 geneV protein is known to coat single-stranded DNA (Terwilliger, T.C.

Gene V Protein Dimerization and Cooperativity of Binding of poly(dA). *Biochemistry* **1996,**

35, 16652, incorporated herein by reference in its entirety), as is human replication protein A

15   (Kim, C.; Snyder, R.O.; Wold, M.S. Binding properties of replication protein A from human

and yeast cells. *Mol. Cell Biol.* **1992**, *12*, 3050, incorporated herein by reference in its entirety),

Pf3 single-stranded binding protein (Powell, M.D.; Gray, D.M. Characterization of the Pf3

single-strand DNA binding protein by circular dichroism spectroscopy. *Biochemistry* **1993**, *32*,

12538, incorporated herein by reference in its entirety), and adenovirus DNA binding protein

20   (Tucker, P.A.; Tsernoglou, D.; Tucker, A.D.; Coenjaerts, F.E.J.; Leenders, H.; Vliet, P.C.

Crystal structure of the adenovirus DNA binding protein reveals a hook-on model for

cooperative DNA binding. *EMBO J.* **1994**, *13*, 2994, incorporated herein by reference in its

entirety). The protein-coated DNA may then be translocated through a nanopore as has been

demonstrated with RecA bound to double-stranded DNA (Smeets, R.M.M.; Kowalczyk, S.W.;

25   Hall, A.R.; Dekker, N.H.; Dekker, C. Translocation of RecA-Coated Double-Stranded DNA

through Solid-State Nanopores. Nano Lett. **2009,** incorporated herein by reference).

Translocation of protein bound to single-stranded DNA is contemplated. The protein coating

functions in the same manner for single-stranded DNA and double-stranded DNA.

[0098]   It is important that the binding moiety 50 be distinguishable from duplex regions

30   when the analyte is introduced to a nanopore or micro- or nanochannel sequencing system. As

such, the locations of the binding moiety 50 on each fragment may be identified, leading to the

location of each probe, since the binding moiety will end adjacent to the 5' end of each probe.

- 22 -

Since the positions of the probes may then be readily determined, and as the analyte portions containing the probes will have regions complementary to the known probe sequences, one may determine the sequence and location of specific domains on the analyte. Numerous maps may be created, corresponding to the positions of different probes. The resulting maps may be
5   combined and utilized to determine broader sequence information for the analyte.

[0099]   A similar embodiment is depicted in FIGs. 6(a) – 6(d). The analyte 60 is formed from a single-stranded DNA (ssDNA) or RNA fragment 20 exposed to tagged oligonucleotide probes 70. As stated above, each tagged probe 70 may include a short segment of ssDNA 10 of a known sequence. The probing segments 10 may be of any length depending on the number
10   of bases that they include. As before, each of the probing sequences is preferably of an identical sequence, thereby causing the probes to selectively hybridize only to portions of the biomolecule fragment 20 that have a complementary sequence. Unlike the embodiment of FIG. 5, however, in this embodiment, each probe 70 includes a probe sequence 10 and a tag 80 connected to the 5' end of each probe sequence 10 by a linker 90. In the embodiment shown,
15   the tag 80 may comprise a dsDNA segment, however, any of a wide variety of tags known to those skilled in the art may be employed. The tags make current fluctuations in sequencing systems more noticeable as the hybridized probes translocate through the system. In addition, different tags may be used to help distinguish among the different probes. These tags may be proteins, double-stranded DNA, single-stranded DNA or other molecules.

20   [0101]   The biomolecule fragment 20 to be analyzed and probes 70 are depicted prior to hybridization in FIG. 6(a).

[0102]   The analyte 60 is shown in FIG. 6(b) following hybridization of the probes 70 to the biomolecule fragment 20. The resulting structure is a biomolecule fragment having, where hybridization has occurred, duplex domains. The double stranded domains are of a length
25   corresponding to the length of the probe sequences 10. Thus, in a case where a 6-mer probe sequence is employed, the analyte 60 will comprise a single-stranded structure having a plurality of 6-mer duplex regions, formed by the hybridized probes.

[0103]   Following the hybridization step, a base extension reaction is performed as depicted in FIG. 6(c). As in the prior example, the hybridized probes 70 are extended from the
30   3' ends of the probe sequences 10 along the biomolecule fragment 20 to create duplex regions 40 on the analyte in single-stranded gaps of that had previously existed between the probe locations. Note, however, that as before, the base extension reaction is intended to be limited in

- 23 -

scope. Rather than extending from the 3' end of each probe to the 5' end of an adjacent probe, the base extension reaction is intended to be terminated such that single-stranded segments remain on the analyte 60 before the 5' end of each probe sequence 10. The resulting analyte 60 structures, like those of FIG. 6(b), comprise duplex regions alternating with single-stranded regions. However, unlike the structures of FIG. 6(b), the resulting structures are characterized as being primarily duplexes with small single-strand gaps.

[0104]   Finally, as depicted in FIG. 6(d), a binding moiety 50, which is selective to the single-strand regions, is reacted with the analyte 60 in a manner such that the binding moiety fills the single-stranded gaps in the analyte. Although it is preferred that the binding moiety 50 be distinguishable from duplex regions when the analyte is introduced to a nanopore or micro- or nanochannel sequencing system, since this embodiment uses tags to enhance detection of the probes, the requirement is less stringent than in that of FIG. 5. As before, by detecting the binding moiety and/or the tags, the locations of the probes on the analyte may be determined. Since the positions of the probes may be readily determined, and as the analyte portions containing the probes will have regions complementary to the known probe sequences, one may map the sequence and location of specific domains on the analyte. The resulting maps may be combined and utilized to determine broader sequence information for the analyte.

[0105]   Another embodiment is depicted in FIGs. 7(a) – 7(f). In FIG. 7(a), the analyte 25 comprising a single-stranded DNA fragment 20 is exposed to RNA probes 35. Each probe 35 is a short RNA fragment of a known sequence. The probes 35 may be of any length depending on the number of bases that they include. As before, each of the probes is preferably of an identical sequence, although probe mixtures of two or more sequences are envisioned as well. The probes selectively hybridize only to portions of the ssDNA fragment 20 that have a complementary sequence. The ssDNA fragment 20 and RNA probes 35 are depicted prior to hybridization in FIG. 7(a).

[0106]   The analyte 25 is shown in FIG. 7(b) following hybridization of the RNA probes 35 to the biomolecule fragment 20. The resulting structure is a ssDNA fragment having duplex domains where hybridization has occurred. The duplex domains are of a length corresponding to the length of the RNA probes 35. Thus, in a case where a 6-mer probe is employed, the analyte 25 will include a plurality of 6-mer duplex regions, formed by the hybridized probes.

[0107]   Following the hybridizing step, a DNA base extension reaction is performed as depicted in FIG. 7(c). Specifically, the extension runs from the 3' ends of the RNA probes 35

- 24 -

along ssDNA fragment 20 to create duplex regions 45 on the analyte in gaps that had previously existed between the probe locations. Unlike the base extension reaction of FIGs. 5 and 6 above, however, the base extension reaction may be allowed to proceed from the 3' end of each probe to the 5' end of an adjacent probe as shown in FIG 7(d). The resulting analyte 25 structure includes duplex regions of double-stranded DNA (dsDNA) 45 resulting from the base extension reaction, alternating with regions of RNA/DNA duplexes 35 in regions where the RNA probes hybridized to the ssDNA analyte.

[0108] The analyte 25 of FIG. 7(d) is then reacted with a hydroxyl ion (-OH) which denatures and decomposes the RNA probes. The resulting analyte shown in FIG. 7(e) includes dsDNA regions 45 separated by gaps formed where the RNA probes had previously hybridized to the ssDNA analyte.

[0109] Finally, as depicted in FIG. 7(f), a binding moiety 50 which is selective to ssDNA regions, is reacted with the analyte 25 in a manner such that the binding moiety fills the ssDNA gaps in the analyte. Thus, the RNA probes are removed to allow the reaction of the binding moiety 50 with the analyte 25, to thereby prepare the biomolecule analyte, i.e., the target analyte. The binding moiety is distinguishable from the double-stranded regions, and is typically easier to detect than a single-stranded region. Moreover, the binding moiety 50 is typically easier to detect than an RNA probe 35. The described method, therefore, facilitates the detection of sites to which the RNA probes 35 hybridize.

[0110] Another embodiment is depicted in FIGs. 8(a) – 8(c), where the analyte 100 is formed from a ssDNA or RNA biomolecule fragment 20 exposed to single-stranded oligonucleotide probes 10. As stated above, each probe 10 is a short ssDNA or RNA segment of a known sequence. The probes 10 may be of any length depending on the number of bases that they include. As before, each of the probes is preferably of an identical sequence, thereby causing the probes to selectively hybridize only to portions of the biomolecule fragment 20 that have a complementary sequence. The biomolecule fragment 20 and probes 10 are depicted prior to hybridization in FIG. 8(a).

[0111] The analyte 100 is shown in FIG. 8(b) following hybridization of the probes 10 to the biomolecule fragment 20. The resulting structure is a single-stranded biomolecule fragment having, where hybridization has occurred, duplex domains. The duplex domains are of a length corresponding to the length of the probes. Thus, in a case where a 6-mer probe is employed,

- 25 -

the analyte 100 will include a plurality of 6-mer duplex regions, formed by the hybridized probes.

[0112]   Next, unlike the embodiments of FIGs. 5–7, no base extension reaction is used. Rather, a binding moiety 110 with a selectivity for duplex domains is employed. Suitable compositions include proteins such as enzymes that lack a cofactor. Examples include restriction enzymes that are added in the absence of Mg^{2+}. Lacking the Mg ion, the protein will bind to duplex domains but be unable to cut. In other embodiments, polymerases can be used in the absence of Mg^{2+} or nucleoside triphosphates may be used. Likewise, topoisomerases could be added in the absence of Mg^{2+}. Regardless of the specific duplex selective entity, as shown in FIG. 8(c), the binding moiety 110 selectively binds to the analyte 100 only at duplex regions on the analyte 100, thereby enhancing detection of these regions when the analyte is introduced to a nanopore or micro- or nanochannel sequencing system. As before, this allows determination of the locations of probe binding, and provides both location and sequence information for the analyte.

[0113]   Similarly, in FIGs. 9(a) – 9(c), the analyte 120 is formed from a single-stranded biomolecule fragment 20 exposed to tagged single-stranded oligonucleotide probes 70. As in FIG. 6, each tagged probe 70 is a short segment of ssDNA or RNA, (i.e., a probe or probing segment), 10 of a known sequence. The probing segments 10 may be of any length depending on the number of bases that they include. As before, each of the probing sequences is preferably of an identical sequence, thereby causing the probes to selectively hybridize only to portions of the biomolecule fragment 20 that have a complementary sequence. Each probe includes a probe sequence 10 and a tag 80 connected the 5' end of the probe sequence 10 by a linker 90. As before, the tag 80 may comprise a dsDNA segment, however, any of a wide variety of tags known to those skilled in the art may be employed. The biomolecule fragment 20 to be analyzed and probes 70 are depicted prior to hybridization in FIG. 9(a).

[0114]   The analyte 120 is shown in FIG. 9(b) following hybridization of the probes 70 to the biomolecule fragment 20. The resulting structure is a biomolecule fragment having, where hybridization has occurred, tagged duplex domains thereon. The duplex domains are of a length corresponding to the length of the probe sequences 10. Thus, in a case where a 6-mer probe sequence is employed, the analyte 120 will include a plurality of tagged 6-mer duplex regions, formed by the hybridized tagged probes.

- 26 -

[0115]   As is shown in FIG. 9(c), the analyte 120 has been reacted with a binding moiety 110 with a selectivity for duplex structures. The binding moiety 110 selectively binds to the analyte 120 only at regions hybridized to probe sequences 10, and, if duplex tags are employed, at the tags, thereby enhancing detection of these regions when the analyte is introduced to a nanopore or micro- or nanochannel sequencing system. As before, this allows determination of the locations of probe binding, and provides both location and sequence information for the analyte.

[0116]   Another embodiment, depicted in FIGs. 10(a) – 10(c), is similar to that shown in FIG. 5, but without the use of a base extension reaction. In FIG. 10, the analyte 130 is formed from an ssDNA or RNA fragment 20 exposed to single-stranded oligonucleotide probes 10. As stated above, each probe 10 is of a known sequence. As before, each of the probes is preferably of an identical sequence, thereby causing the probes to selectively hybridize only to portions of the biomolecule fragment 20 that have a complementary sequence. The biomolecule fragment 20 and probes 10 are depicted prior to hybridization in FIG. 10(a).

[0117]   The analyte 130 is shown in FIG. 10(b) following hybridization of the probes 10 to the biomolecule fragment 20. The resulting structure is a single-stranded fragment having, where hybridization has occurred, duplex domains. The duplex domains are of a length corresponding to the length of the probes. Thus, in a case where a 6-mer probe is employed, the analyte 130 will comprise a single-stranded biomolecule fragment having a plurality of 6-mer duplex regions, formed by the hybridized probes.

[0118]   In FIG. 10(c), a binding moiety 50 with a selectivity for duplex regions is employed as described above. A base extension reaction is not used. Suitable binding moieties have been described previously. As is shown in FIG. 10(c), the binding moiety 50 selectively binds to the analyte 130 only at single-stranded regions on the analyte 130, thereby enhancing detection of these regions when the analyte is introduced to a nanopore or micro- or nanochannel sequencing system. As before, this allows determination of the locations of probe binding, and provides both location and sequence information for the analyte.

[0119]   Mapping of target analytes prepared using the methods of embodiments of the present invention may be carried out using electrical detection methods employing nanopores, nano-channels or micro-channels using the methods described in U.S. Patent Application Serial No. 12/789,817.

[0120]   In one embodiment, current is measured during translocation of a DNA strand through a nanopore as shown in FIG. 11(a). When used in embodiments of the present invention, nanopores may have a diameter selected from a range of about 1 nm to about 1 μm. More preferably the nanopore has a diameter that is between about 2.3 nm and about 100 nm. Even more preferably the nanopore has a diameter that is between about 2.3 nm and about 50 nm.

[0121]   Specifically, for nanopore 150, a measurable current produced by electrodes 152, 154 runs parallel 160 to the movement of the target analyte 25. In this example, the target analyte 25, i.e., biomolecule analyte, is made using the method depicted in FIGS. 7(a) – 7(f), however any of the methods and analytes of embodiments of the present invention could be used. The target analyte 25 includes a single-stranded DNA fragment 20 having a duplex region 45 and a binding moiety 50 present at non-duplexed regions. Variations in current are a result of the relative diameter of the target analyte 25 as it passes through the nanopore 150. A relative increase in volume of the target analyte 25 passing through the nanopore 150 causes a temporary interruption or decrease in the current flow through the nanopore, resulting in a measurable current variation. Portions of the target analyte 25 including the binding moiety 50 are larger in diameter than portions of the target analyte that do not include the binding moiety 50. As a result, when the binding moiety 50 passes through the nanopore 150, further interruptions or decreases in the current flow between electrodes 152, 154 occurs. This is depicted in the waveform 200 in FIG. 11(b).

[0122]   In FIG. 11(b), the waveform 200 may be interpreted as follows. Current measurement 210 represents measured current prior to passage of the target analyte 25, i.e., biomolecule analyte, through the nanopore 150 from the *cis* side to the *trans* side. As the target analyte 25 enters the nanopore 150, from the *cis* side of the nanopore, the current is partially interrupted forming a first trough 220 in the recorded current. Once the binding moiety 50 on the target analyte 25 enters the nanopore 150, a further decrease in current occurs, causing a deeper, second trough 230 in the current measurement. Upon passage of the binding moiety 50 entirely through the nanopore 150, a distal portion of the target analyte 25 may remain in the nanopore. This causes the measured current 240 to rise to approximately the level of the first trough 220. Finally, once the entire target analyte has passed completely through the nanopore 150 to the *trans* side, the measured current 250 returns to a level approximating that of the initial level 210. The current variation measurements are recorded as a function of time. As a

- 28 -

result, the periodic variations in current indicate where, as a function of relative or absolute position, the regions of binding moiety 50 are formed on the target analyte 25. Since the binding moiety is present at recognition sites for specific probes, the relative or absolute position of the specific sequences associated with each probe may be determined. This allows mapping of those specific sequences on the target analyte. Multiple maps and sequence information produced using multiple probes may be generated.

[0123]   In another embodiment, an electrical property such as electrical potential or current is measured during translocation of a DNA strand through a nano-channel or micro-channel as shown in FIGs. 12 through 14. One embodiment of a fluidic channel apparatus is shown schematically in FIG. 12. In FIG. 12, the apparatus 300 includes a fluidic micro-channel or nano-channel 302. The fluidic channel may be a micro-channel having a width selected from a range of about 1 μm to about 25 μm or a nano-channel having a width selected from a range of about 10 nm to about 1 μm. In the case of a micro-channel, the depth may be selected from a range of about 200 nm to about 5 μm, whereas in the case of a nano-channel, the depth may be selected from a range of about 10 nm to about 1 μm. In either case, the channel may have a length selected from a range of about 1 μm to about 10 cm.

[0124]   A first pair of electrodes 304, 304' is connected to a current source 306 and positioned in a spaced apart relationship in the channel. These electrodes provide an electrical current along the channel and may be used to provide or enhance a driving force 308 to a target analyte 25, i.e., biomolecule analyte, in the channel. Other driving forces such as pressure or chemical gradients are contemplated as well. A second pair of electrodes 312, 312', i.e., detector electrodes, is positioned preferably substantially perpendicular to the channel in a spaced apart relationship to define a detection volume 314. The second pair of electrodes 312, 312', is connected to a detector 316, such as a voltmeter, which monitors an electrical property in the detection volume 314. In the embodiment where the detector 316 is a voltmeter, a drop in electrical potential, produced by the first pair of electrodes 304, 304', is measured across the detection volume 314. As in the case of the nanopore apparatus, in this example, a target analyte 25 is made using the method depicted in FIGS. 7(a) – 7(f), however any of the methods and analytes of embodiments of the present invention could be used.

[0125]   The operation of the device is depicted schematically in FIGs. 13(a) – 13(d). In FIGs. 13(a) – 13(d), the first pair of electrodes 304, 304' and the current source 306 have been omitted for clarity. In FIG. 13(a), the fluidic channel 302 contains a target analyte 25, i.e.,

- 29 -

biomolecule analyte, traveling therethrough. An electrical property, in this case electrical potential, is measured and recorded across the detection volume 314 by the detector electrodes 312, 312' and the detector 316. In the embodiment depicted, the target analyte 25 is that of FIGs. 7(f) and 14; i.e., it includes a single-stranded DNA fragment having a duplex region and

5   a binding moiety present at non-duplexed regions.

[0126]   Prior to the entry of the analyte 25 into the detection volume 314, a substantially constant voltage 322 is measured across the detection volume. This voltage is shown in the waveform 320 of FIG. 13(a). As the analyte 25 enters the detection volume 314, it causes an interruption or decrease in the electrical property measured in the detection volume. This

10   interruption or decrease causes a first trough 324 to be exhibited in the waveform 320.

[0127]   FIG. 13(b) shows the device and waveform 320 once the portion of the analyte 25 including the binding moiety has entered the detection volume 314. Due to its increased volume, entry of the binding moiety into the detection volume 314 causes a further interruption or decrease in the electrical property measured in the detection volume. This further

15   interruption or decrease causes a second trough 326 to be exhibited in the waveform 320.

[0128]   In FIG. 13(c), the portion of the analyte 25 containing the binding moiety has exited the detection volume 314; however, a distal portion of the analyte 25 may still be present in the detection volume. As a result, the waveform 320 has returned to a level 328 approximating that detected when the initial portion of the analyte first entered the detection

20   volume.

[0129]   Finally, as shown in FIG. 13(d), the analyte 25 has fully exited the detection volume 314. As a result, the waveform 320 has returned to a level 330 approximating that detected prior to initial entry of the analyte into the detection volume.

[0130]   Another embodiment of a fluidic channel apparatus is shown in FIG. 14. In FIG.

25   14, the apparatus 400 includes a fluidic micro-channel or nano-channel 402. As before, the fluidic channel may be a micro-channel having a width selected from a range of about 1 μm to about 25 μm or a nano-channel having a width selected from a range of about 10 nm to about 1 μm. In the case of a micro-channel, the depth may be selected from a range of about 200 nm to about 5 μm, whereas in the case of a nano-channel, the depth may be selected from a range of

30   about 10 nm to about 1 μm. In either case, the channel may have a length selected from a range of about 1 μm to about 10 cm.

[0131]   A first pair of electrodes 404, 404' is connected to a current source 406 and positioned in a spaced apart relationship in the channel. These electrodes provide an electrical current along the channel and may be used to provide or enhance a driving force 408 to an analyte 25 in the channel. Other driving forces such as pressure or chemical gradients are contemplated as well. Multiple detector electrodes 412, 414, 416, 418, are positioned preferably perpendicular to the channel in a spaced apart relationship to define a plurality of detection volumes between adjacent detector electrodes. Thus, as seen in FIG. 16, detector electrodes 412 and 414 define detection volume 420, detector electrodes 414 and 416 define detection volume 422, and detector electrodes 416 and 418 define detection volume 424. The detector electrodes are each connected to detectors 426, 428, 430 such as voltmeters, which monitor an electrical property in each detection volume. In the embodiment where the detectors are voltmeters, a drop in electrical potential is measured across each detection volume. Operation of the apparatus is similar to that of the system of FIG. 12, with the exception that additional waveforms are generated due to the presence of additional detection volumes. The additional waveforms may be combined to further improve the quality of the data being generated by the device.

[0132]   It should be understood that number of detector electrodes and detection volumes is not intended to be limited to those depicted in FIG. 14. Rather, any number of detection volumes may be included along the length of the fluidic channel. Further, the detector electrodes and detection volumes need not be evenly spaced, evenly sized or directly adjacent to one another. Various detection volume sizes, spacing and configurations are contemplated.

[0133]   Both the nanopore apparatus and the fluidic channel apparatus allow detection of an analyte as well as detection of a binding moiety present on the analyte. Furthermore, relative or absolute positional information of the binding moiety may be obtained. Since, in each of the embodiments depicted herein, the relative or absolute position of the binding moiety is, at the very least, suggestive of the relative or absolute position of a known probe, embodiments of the invention allow determination of the location of the known recognition sequence of the probe. This in turn, allows the biomolecule to be mapped. The repeated use of different probes allows greater complexity, i.e., multiple recognition sequences, to be combined. Multiple maps and sequence information produced using multiple probes may be generated.

- 31 -

Equivalents

[0134]   Those skilled in the art will readily appreciate that all parameters listed herein are meant to be exemplary and actual parameters depend upon the specific application for which the methods and materials of embodiments of the present invention are used. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that,

5   within the scope of the appended claims and equivalents thereto, the invention may be practiced otherwise than as specifically described.

[0135]   The described embodiments of the invention are intended to be merely exemplary and numerous variations and modifications will be apparent to those skilled in the art. All such variations and modifications are intended to be within the scope of the present invention as

10   defined in the appended claims.

[0136]   What is claimed is:

- 32 -

1   1. A method for preparing a biomolecule analyte, the method comprising:
2   a. providing a single-stranded DNA or RNA template;
3   b. hybridizing a first plurality of identical, sequence specific oligonucleotide
4   probes to the template, each probe having a 5' end and a 3' end, to thereby
5   form an analyte having at least one single-stranded region and at least two
6   duplex regions;
7   c. conducting a base extension reaction in the at least one single-stranded region
8   from the 3' end of a hybridized probe toward the 5' end of an adjacent
9   hybridized probe;
10   d. terminating the base-extension reaction such that there remains for each single-
11   stranded region a single-stranded portion thereof adjacent to the 5' end of each
12   hybridized probe; and
13   e. reacting the resulting analyte with a binding moiety that selectively binds to the
14   at least one single-stranded portion to thereby prepare the biomolecule analyte.

1   2. The method of Claim 1, wherein the probes comprise single-stranded DNA.

1   3. The method of Claim 1, wherein the probes comprise RNA.

1   4. The method of any one of Claims 1-3, wherein the base extension reaction is performed
2   by a DNA or RNA polymerase.

1   5. The method of any one of Claims 1-4, wherein the binding moiety comprises a protein.

1   6. The method of Claim 5, wherein said protein comprises one or more proteins selected
2   from the group consisting of RecA, T4 gene 32 protein, f1 geneV protein, human
3   replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding
4   protein, and *E. coli* single-stranded binding protein.

1   7. The method of any one of Claims 1-6, wherein a length of each probe is selected from a
2   range of 4 to 12 bases.

1   8. The method of any one of Claims 1-7, wherein the method comprises performing steps
2   a – e sequentially.

1   9. The method of any one of Claims 1-8, further comprising repeating steps a – e
2   sequentially by replacing said first plurality of probes with a subsequent plurality of
3   different unique probes.

1   10. The method of any one of Claims 1-9, wherein at least a portion of the probes in the
2   first plurality of probes has attached thereto a detectable tag.

1   11. The method of any one of Claims 1-10, wherein the biomolecule analyte is configured
2   for detection of positional information in a nanopore system.

1   12. The method of Claim 11, further comprising:
2   monitoring changes in an electrical property across a nanopore as the biomolecule
3   analyte is translocated therethrough, the changes in the electrical property being indicative
4   of regions including or lacking the binding moiety.

1   13. The method of Claim 12, further comprising:
2   differentiating between regions of the biomolecule analyte including or lacking the
3   binding moiety based, at least in part, on the changes in the electrical property, to thereby
4   determine binding moiety locations.

1   14. The method of any one of Claims 1-10, wherein the biomolecule analyte is configured
2   for detection of positional information in a fluidic channel system.

1   15. The method of Claim 14, wherein the fluidic channel system comprises a micro-channel
2   or a nano-channel.

1   16. The method of Claim 14, further comprising:
2   monitoring changes in an electrical property across a fluidic channel as the biomolecule
3   analyte is translocated therethrough, the changes in the electrical property being indicative
4   of regions including or lacking the binding moiety.

1   17. The method of Claim 16, further comprising:
2   differentiating between regions of the biomolecule analyte including or lacking the
3   binding moiety based, at least in part, on the changes in the electrical property, to thereby
4   determine binding moiety locations.

1   18. The method of Claim 13 or 17, further comprising:
2   determining a sequence of at least a portion of the single-stranded DNA or RNA
3   template according to the determined binding moiety locations.

1   19. The method of Claim 1, wherein sequence specific oligonucleotide analog probes
2   selected from the group consisting of LNAs, PNAs and 2'-methoxy nucleotide analogs are
3   substituted for the sequence specific oligonucleotide probes.

1   20. The method of Claim 1, wherein the first plurality of identical, sequence specific
2   oligonucleotide probes is replaced by a pool of sequence specific oligonucleotide probes

3   comprising at least a first plurality of identical, sequence specific oligonucleotide probes

4   and a second plurality of identical, sequence specific oligonucleotide probes, wherein the

5   probes of the second plurality are different from the probes of the first plurality.

1   21. A method for preparing a biomolecule analyte, the method comprising:

2   a.   providing a single-stranded DNA template,

3   b.   hybridizing a first plurality of identical, sequence specific RNA probes to the

4   template, each probe having a 5' end and a 3' end, to thereby form an analyte

5   having at least one single-stranded region and at least one duplex region,

6   c.   conducting a base extension reaction in the at least one single-stranded region

7   from the 3' end of a hybridized probe,

8   d.   allowing the base-extension reaction to fill each single-stranded region on the

9   analyte,

10   e.   removing the RNA probes to provide the analyte with at least one single-

11   stranded segment in the region to which an RNA probe had been hybridized, and

12   f.   reacting the resulting analyte with a binding moiety that selectively binds to the

13   single-stranded segment, to thereby prepare the biomolecule analyte.

1   22. The method of Claim 21, wherein the base extension reaction is performed by a DNA or

2   RNA polymerase.

1   23. The method of Claim 21 or 22, wherein the binding moiety comprises a protein.

1   24. The method of Claim 23, wherein said protein comprises one or more proteins selected

2   from the group consisting of RecA, T4 gene 32 protein, f1 geneV protein, human

3   replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding

4   protein, and *E. coli* single-stranded binding protein.

1   25.   The method of any one of Claims 21-24, wherein a length of each probe is selected

2   from a range of 4 to 12 bases.

1   26.   The method of any one of claims 21-25, wherein the method comprises performing

2   steps a – f sequentially.

1   27.   The method of any one of Claim 21-26, further comprising repeating steps a - f

2   sequentially by replacing said first plurality of probes with a subsequent plurality of

3   different unique probes.

1   28. The method of any one of Claims 21-27, wherein removing the RNA probes comprises

2   reacting the analyte with a hydroxyl ion.

29. The method of any one of Claims 21-28, wherein the biomolecule analyte is configured for detection of positional information in a nanopore system.

30. The method of Claim 29, further comprising:

monitoring changes in an electrical property across a nanopore as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety.

31. The method of Claim 30, further comprising:

differentiating between regions of the biomolecule analyte including or lacking the binding moiety based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

32. The method of any of Claims 21-28, wherein the biomolecule analyte is configured for detection of positional information in a fluidic channel system.

33. The method of Claim 32, wherein the fluidic channel system comprises a micro-channel or a nano-channel.

34. The method of Claim 32, further comprising:

monitoring changes in an electrical property across a fluidic channel as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety.

35. The method of Claim 34, further comprising:

differentiating between regions of the biomolecule analyte including or lacking the binding moiety based, at least in part, on the changes in the electrical property, to thereby determine binding moiety locations.

36. The method of Claim 31 or 35, further comprising:

determining a sequence of at least a portion of the single-stranded DNA or RNA template according to the determined binding moiety locations.

37. The method of Claim 21, wherein the first plurality of identical, sequence specific RNA probes is replaced by a pool of sequence specific RNA probes comprising at least a first plurality of identical, sequence specific RNA probes and a second plurality of identical, sequence specific RNA probes, wherein the probes of the second plurality are different from the probes of the first plurality.

38. A method for preparing a biomolecule analyte, the method comprising:

a. providing a single-stranded DNA or RNA template,

- 36 -

b. hybridizing a first plurality of identical, sequence specific oligonucleotide
    probes to the template, to thereby form an analyte having at least one single-
    stranded region and at least one duplex region, and

c. reacting the resulting analyte with a binding moiety that selectively binds to the
    at least one duplex region to thereby prepare the biomolecule analyte.

39. The method of Claim 38, wherein the probes comprise single-stranded DNA.

40. The method of Claim 38, wherein the probes comprise RNA.

41. The method of any of Claims 38-40, wherein the binding moiety comprises a protein.

42. The method of Claim 41, wherein said protein comprises one or more proteins selected
from the group consisting of RecA, T4 gene 32 protein, f1 geneV protein, human
replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding
protein, and *E. coli* single-stranded binding protein.

43.   The method of any of Claims 38-42, wherein a length of each probe is selected from a
range of 4 to 12 bases.

44.   The method of any of Claims 38-43, wherein the method comprises performing steps a
– c sequentially.

45.   The method of any of Claims 38-44, wherein steps a - c are repeated sequentially by
replacing said first plurality of probes with a subsequent plurality of different unique
probes.

46. The method of any of Claims 38-45, wherein at least a portion of the probes in the first
plurality of probes has attached thereto a detectable tag.

47. The method of any of Claims 38-46, wherein the biomolecule analyte is configured for
detection of positional information in a nanopore system.

48. The method of Claim 47, further comprising:
    monitoring changes in an electrical property across a nanopore as the biomolecule
analyte is translocated therethrough, the changes in the electrical property being indicative
of regions including or lacking the binding moiety.

49. The method of Claim 48, further comprising:
    differentiating between regions of the biomolecule analyte including or lacking the
binding moiety based, at least in part, on the changes in the electrical property, to thereby
determine binding moiety locations.

1   50. The method of any of Claims 38-46, wherein the biomolecule analyte is configured for

2   detection of positional information in a fluidic channel system.

1   51. The method of Claim 50, wherein the fluidic channel system comprises a micro-channel

2   or a nano-channel.

1   52. The method of Claim 50, further comprising:

2   monitoring changes in an electrical property across a fluidic channel as the biomolecule

3   analyte is translocated therethrough, the changes in the electrical property being indicative

4   of regions including or lacking the binding moiety.

1   53. The method of Claim 52, further comprising:

2   differentiating between regions of the biomolecule analyte including or lacking the

3   binding moiety based, at least in part, on the detected changes in the electrical property, to

4   thereby determine binding moiety locations.

1   54. The method of Claim 49 or 53, further comprising:

2   determining a sequence of at least a portion of the single-stranded DNA or RNA

3   template according to the determined binding moiety locations.

1   55. The method of Claim 38, wherein sequence specific oligonucleotide analog probes

2   selected from the group consisting of LNAs, PNAs and 2'-methoxy nucleotide analogs are

3   substituted for the sequence specific oligonucleotide probes.

1   56. The method of Claim 38, wherein the first plurality of identical, sequence specific

2   oligonucleotide probes is replaced by a pool of sequence specific oligonucleotide probes

3   comprising at least a first plurality of identical, sequence specific oligonucleotide probes

4   and a second plurality of identical, sequence specific oligonucleotide probes, wherein the

5   probes of the second plurality are different from the probes of the first plurality.

1   57. A method for preparing a biomolecule analyte, the method comprising:

2   a.   providing a single-stranded DNA or RNA template,

3   b.   hybridizing a first plurality of identical, sequence specific oligonucleotide

4   probes to the template, to thereby form an analyte having at least one single-

5   stranded region and at least one duplex region, and

6   c.   reacting the resulting analyte with a binding moiety that selectively binds to the

7   at least one single-stranded region, to thereby prepare the biomolecule analyte.

1   58. The method of Claim 57, wherein the probes comprise single-stranded DNA.

1   59. The method of Claim 57, wherein the probes comprise RNA.

- 38 -

60. The method of any of Claims 57-59, wherein the binding moiety comprises a protein.

61. The method of Claim 60, wherein said protein comprises one or more proteins selected from the group consisting of RecA, T4 gene 32 protein, f1 geneV protein, human replication protein A, Pf3 single-stranded binding protein, adenovirus DNA binding protein, and *E. coli* single-stranded binding protein.

62. The method of any of Claims 57-61, wherein a length of each probe is selected from a range of 4 to 12 bases.

63. The method of any of Claims 57-62, wherein the method comprises performing steps a – c sequentially.

64. The method of any of Claims 57-63, wherein steps a - c are repeated sequentially by replacing said first plurality of probes with a subsequent plurality of different unique probes.

65. The method of any of Claim 57-64 wherein at least a portion of the probes in the first plurality of probes has attached thereto a detectable tag.

66. The method of any of Claims 57-65, wherein the biomolecule analyte is configured for detection of positional information in a nanopore system.

67. The method of Claim 66, further comprising:

monitoring changes in an electrical property across a nanopore as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety.

68. The method of Claim 67, further comprising:

differentiating between regions of the biomolecule analyte including or lacking the binding moiety based, at least in part, on the detected changes in the electrical property, to thereby determine binding moiety locations.

69. The method of any of Claims 57-65, wherein the biomolecule analyte is configured for detection of positional information in a fluidic channel system.

70. The method of claim 69, wherein the fluidic channel system comprises a micro-channel or a nano-channel.

71. The method of Claim 69, further comprising:

monitoring changes in an electrical property across a fluidic channel as the biomolecule analyte is translocated therethrough, the changes in the electrical property being indicative of regions including or lacking the binding moiety.

- 39 -

1 72. The method of Claim 71, further comprising:

2 differentiating between regions of the biomolecule analyte including or lacking the

3 binding moiety based, at least in part, on the detected changes in the electrical property, to

4 thereby determine binding moiety locations.

1 73. The method of Claim 68 or 71, further comprising:

2 determining a sequence of at least a portion of the single-stranded DNA or RNA

3 template according to the determined binding moiety locations.

1 74. The method of Claim 57, wherein sequence specific oligonucleotide analog probes

2 selected from the group consisting of LNAs, PNAs and 2'-methoxy nucleotide analogs are

3 substituted for the sequence specific oligonucleotide probes.

1 75. The method of Claim 57, wherein the first plurality of identical, sequence specific

2 oligonucleotide probes is replaced by a pool of sequence specific oligonucleotide probes

3 comprising at least a first plurality of identical, sequence specific oligonucleotide probes

4 and a second plurality of identical, sequence specific oligonucleotide probes, wherein the

5 probes of the second plurality are different from the probes of the first plurality.

FIG. 1

FIG. 2

12

TTAAG

FIG. 3

10

14

CTT CTA GATTC TAG

AAGAT   AAGAT

FIG. 4   10   10

FIG. 5a

FIG. 5b

FIG. 5c

FIG. 5d

FIG. 6a

FIG. 6b

FIG. 6c

FIG. 6d

FIG. 7a

FIG. 7b

FIG. 7c

FIG. 7d

FIG. 7e

FIG. 7f

FIG. 8a

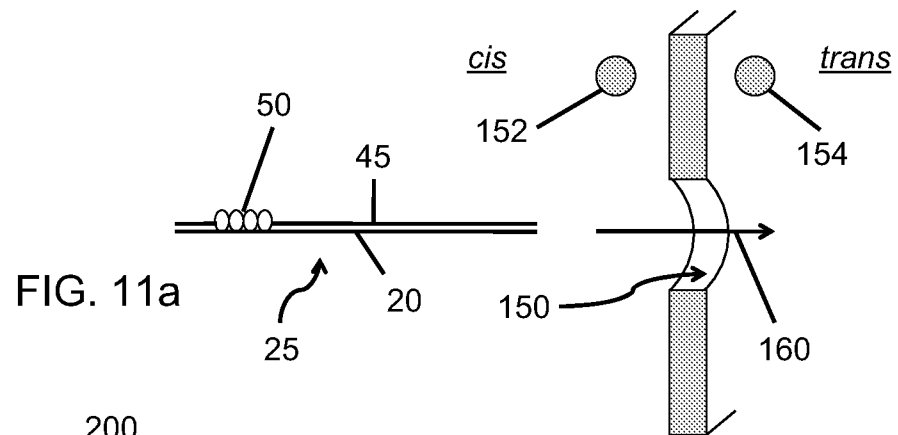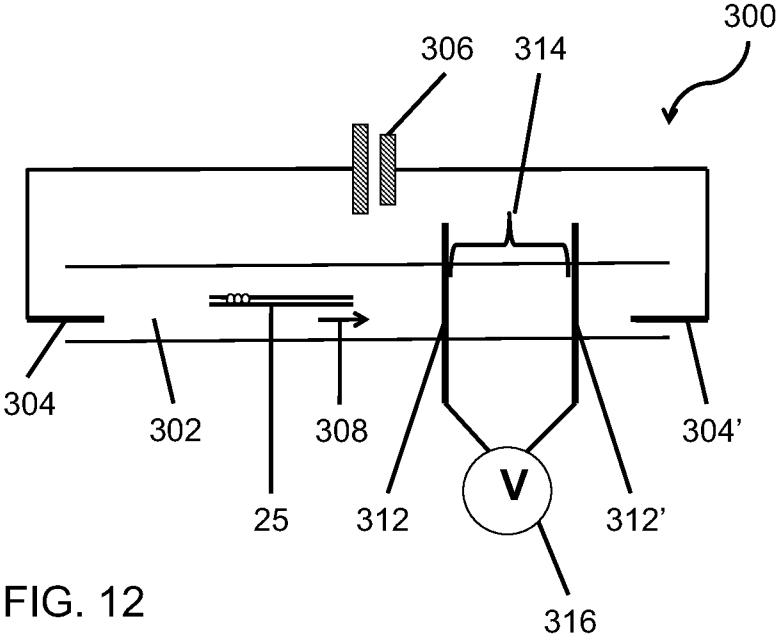FIG. 8b

FIG. 8c

FIG.9a

FIG.9b
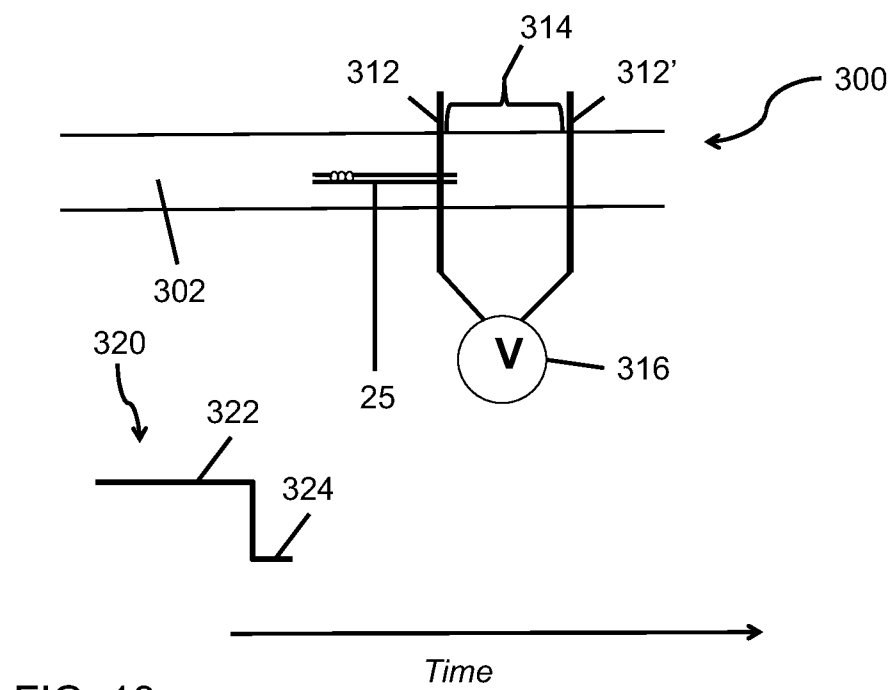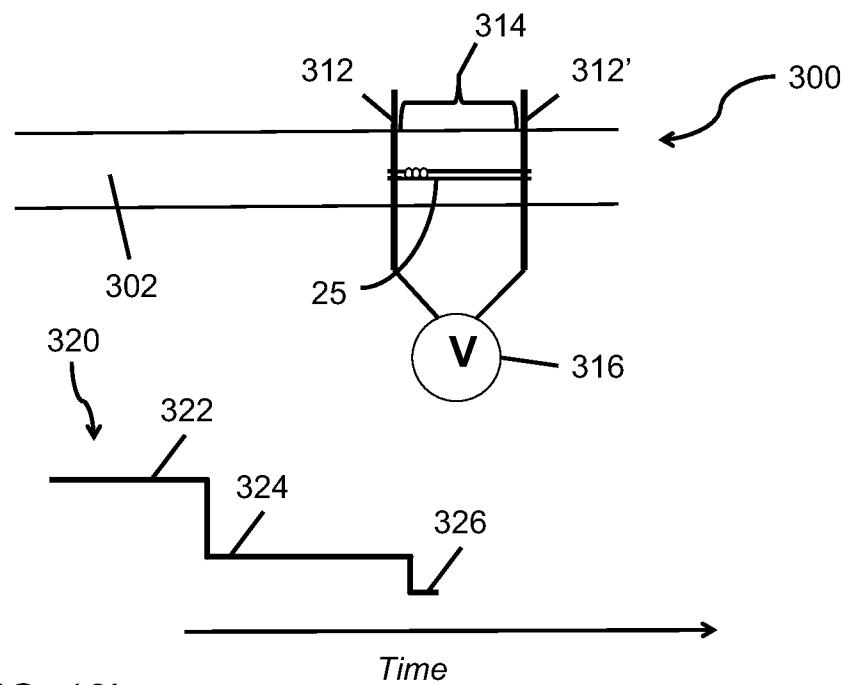
FIG.9c

FIG.10a

FIG.10b

FIG.10c

FIG. 11a

FIG. 11b

Time
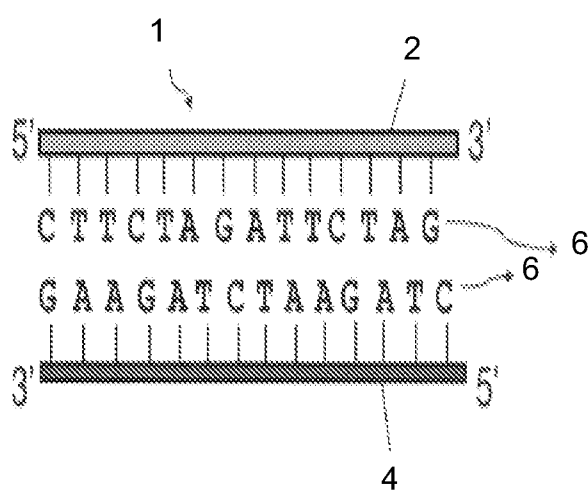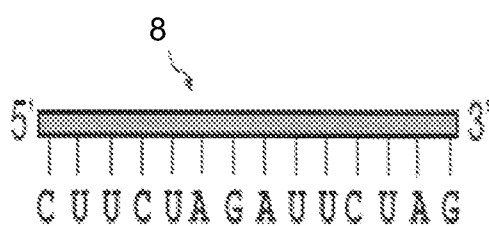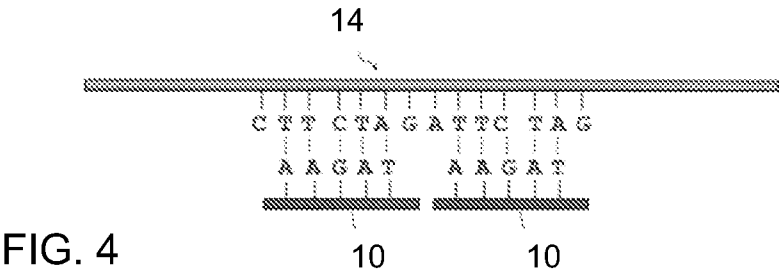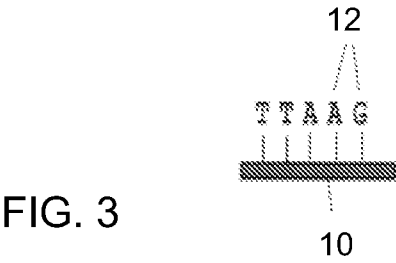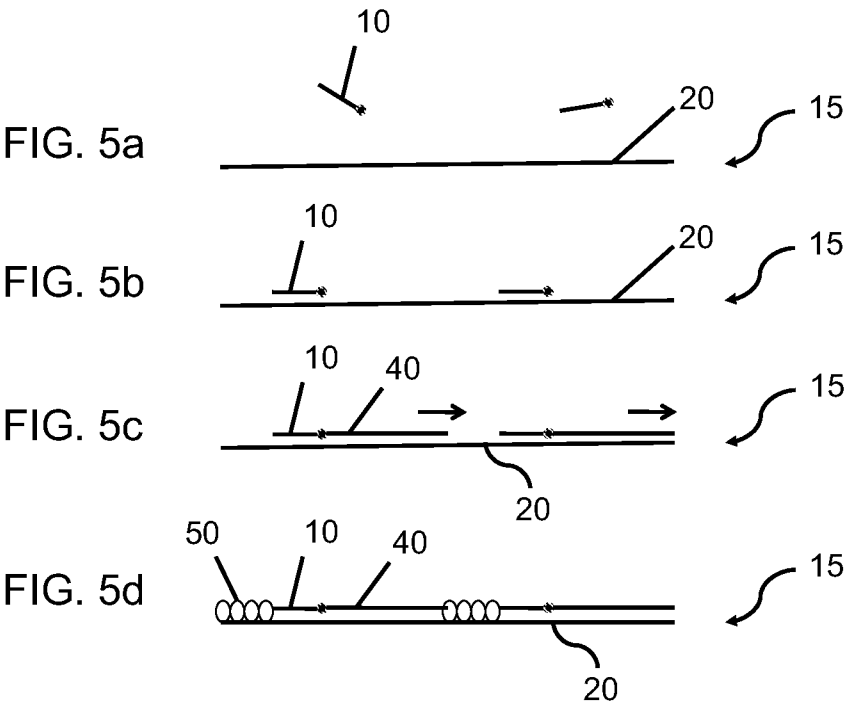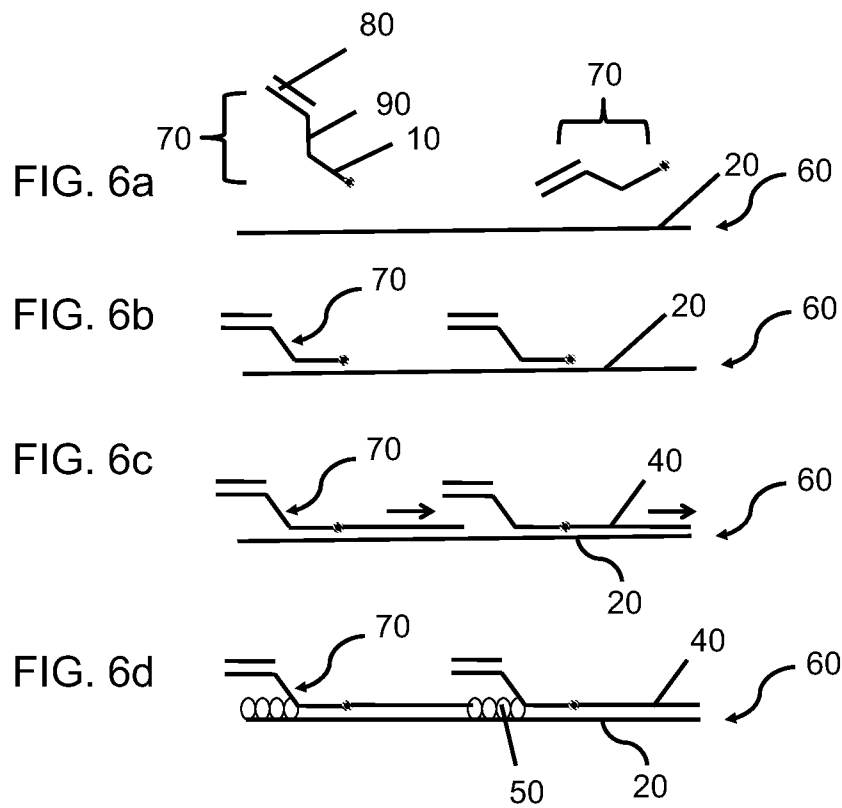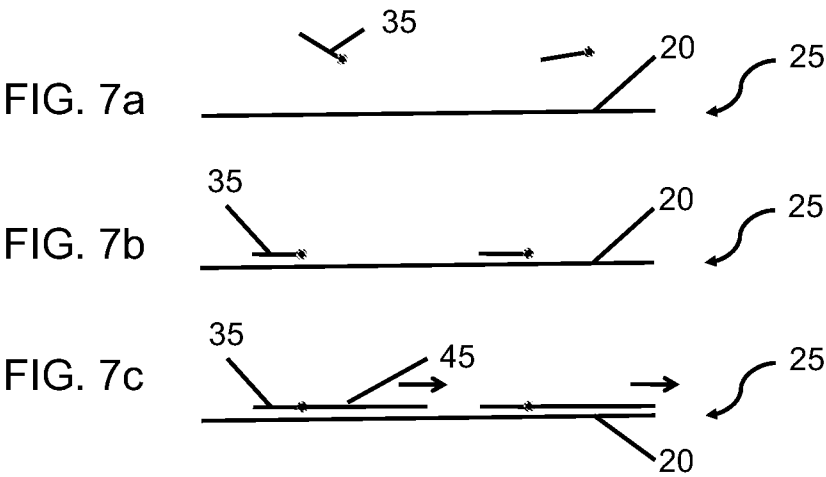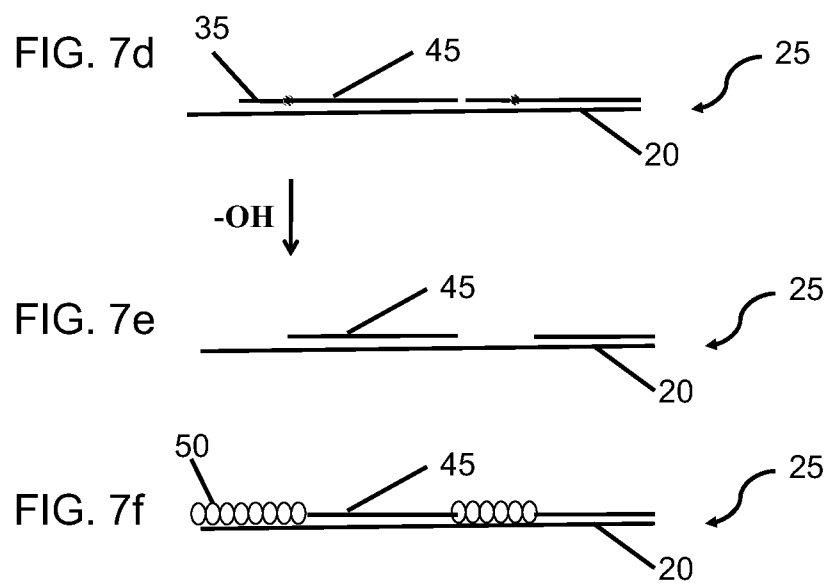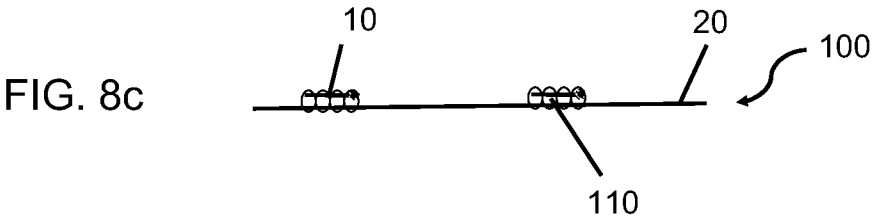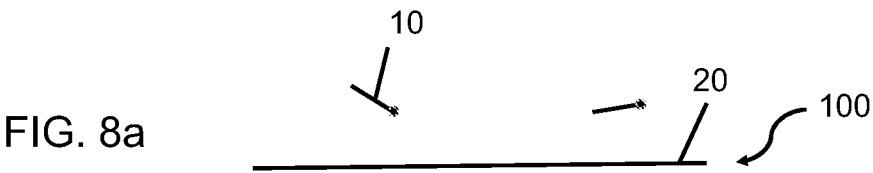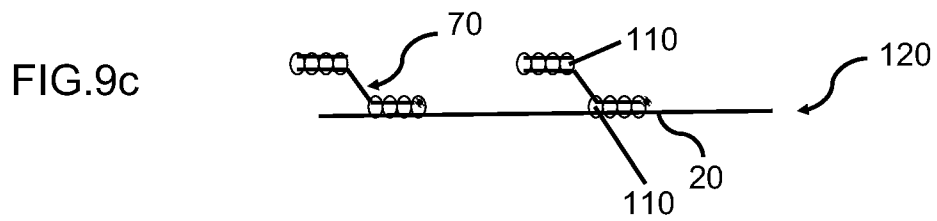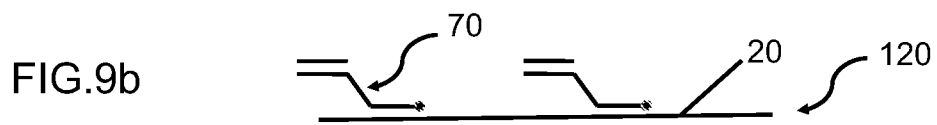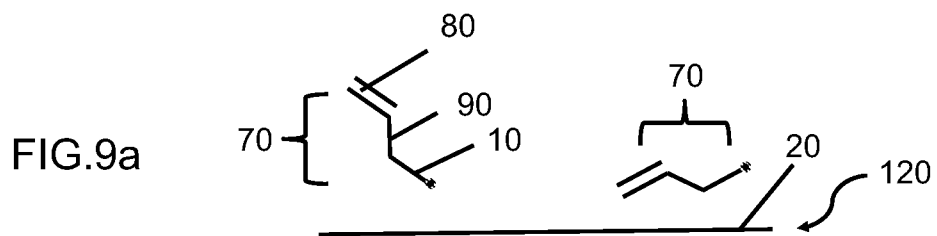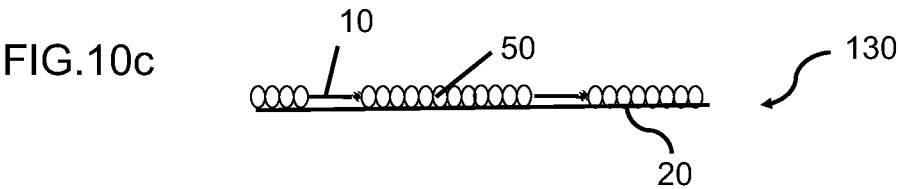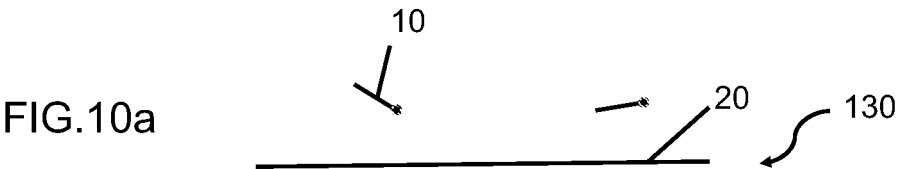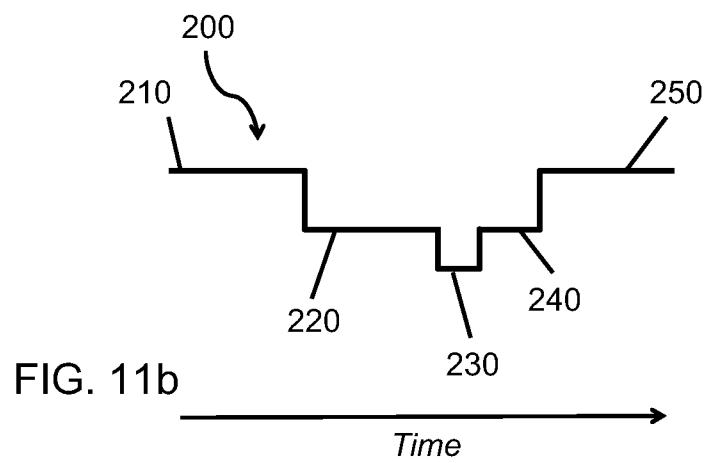
FIG. 12

FIG. 13a

FIG. 13b

FIG. 13c

FIG. 13d

FIG. 14