

SEQUENCE-DETECTION SYSTEM

BACKGROUND

[0001] Many different types of sequences of objects or entities are encountered in industrial and research settings. DNA-sequence determination and RNA-sequence determination are employed in many different research, diagnostic, forensic, and pharmaceutical processes and applications. Many sequence-detection systems use multiple types of detectors and multiple types of signals output by the detectors to identify the types of entities and objects within an input target sequence. Often, electromagnetic sensors are used to generate electrical signals that vary with the dimensional and electrical properties of the objects and entities that pass by the electromagnetic sensors. Optical sensors are used, in certain sequence-detection systems, to identify and classify entities and objects within target sequences by the wavelength of light absorbed by the objects or entities or by computational processing of optical images generated from the objects or entities. Additional detectors include weight detectors, volume detectors, and magnetic-susceptibility detectors.

[0002] Sequence detectors are often characterized by the reliability and accuracy with which they determine and report the sequences of objects or entities of input targets, by the speed with which input target sequences are identified, and by the cost and complexity of the sequence detectors. Often, increases in reliability and accuracy are accompanied with increases in cost and complexity. In order to optimally employ sequence detectors in real-world applications, systems and process engineers seek to evaluate and compare a variety of different types of sequence detectors for use in particular applications. For this reason, systems and process engineers, researchers, diagnosticians, and other users of sequence detectors continuously seek new and different types of sequence detectors to facilitate identifying and deploying specific sequence-detection processes and systems that best meet sequence-detection parameters and goals for specific applications.

SUMMARY

[0003] The current document discusses a detection system comprising a mechanical-change sensor that exhibits one or more mechanical changes when specifically interacting with entities within a target, each entity having a type, a mechanical-change-to-signal transducer that

transduces the one or more mechanical changes into a signal, and an analysis subsystem that determines the types of entities within the target using the signal. The current document additionally discusses a sequence-detection system comprising a nucleic-acid-polymerase mechanical-change component that exhibits mechanical changes when specifically associating with nucleotide polyphosphates within an active site; a mechanical-change-to-signal transducer that transduces mechanical changes in the nucleic-acid-polymerase mechanical-change component into an output signal, and an analysis subsystem that determines a sequence of monomer types within a nucleic-acid-polymer target using the output signal. The current document additionally discusses a method for determining a monomer sequence from a signal output by a sequence-detection system, the method comprising identifying portions of the signal that each corresponds to a different monomer in a sequence of monomers, for each signal portion, computing n derived values from the signal portion, wherein n is an integer greater than or equal to 2, mapping the n derived values to an n -dimensional range volume corresponding to a particular type of monomer, and assigning the particular type of monomer to the signal portion, and generating and storing a symbolic representation of a sequence of monomer types complementary to the monomer types assigned to the signal portions. In certain cases, a single derived value may be sufficient to identify the different monomers in a sequence of monomers.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] Figure 1 illustrates, in one example, a target input to, and a determined sequence representation output by, a first type of mechanical-change-based sequence detector.

[0005] Figure 2 illustrates one implementation of a sequence-detection system based on the first type of mechanical-change-based sequence detector.

[0006] Figures 3A-3F illustrates, in one example, the mechanical-change sensor component of the sequence-detection system shown in Figure 2.

[0007] Figure 4 illustrates, in one example, generation of a voltage signal from changes in the shape of the mechanical-change sensor component by a mechanical-change-detection subsystem within the sequence-detection system shown in Figure 2.

[0008] Figure 5 illustrates, in one example, the internal components of the mechanical-change-based sequence detector 204 shown in Figure 2.

[0009] Figures 6A-6D illustrate, in one example, the voltage signals produced by each of the four different types of objects that occur within targets.

[0010] Figure 7 shows hypothetical analytical results produced by the sequence-detection system from a test target that includes 10 objects of each of the four object types *a*, *b*, *c*, and *d*.

[0011] Figure 8 shows a plot of the mean-voltage-magnitude data contained in the data tables shown in Figure 7.

[0012] Figure 9 shows a two-dimensional plot of the data contained in the tables shown in Figure 7.

[0013] Figure 10 illustrates, in one example, use of three derived values for determination of object types.

[0014] Figure 11 summarizes the sequence detection system discussed above with reference to Figures 1-10.

[0015] Figure 12 provides a table that compares an example first sequence-detection system, discussed above with reference to Figures 1-11, and an example second sequence-detection system, discussed in the current subsection of the current document.

[0016] Figure 13 illustrates, in one example, a mechanical-change-based sequence detector that is included in the second sequence-detection system.

[0017] Figures 14A-14B illustrate two different current-to-voltage converter circuits that are used separately or together in various implementations of the second sequence-detection system.

[0018] Figure 15 illustrates, in one example, an array of cells that provides for parallel sequence determination.

[0019] Figure 16 illustrates a short DNA polymer.

[0020] Figures 17A-17B illustrate the hydrogen bonding between the purine and pyrimidine bases of two anti-parallel DNA strands.

[0021] Figures 18A-18B illustrate double-stranded DNA.

[0022] Figure 19 shows the names and chemical structures of the 20 common amino acids. Amino acids are polymerized in a ribosome-mediated translation process to form proteins.

[0023] Figure 20 illustrates the polymerization reaction catalyzed by the Klenow fragment of *E. coli* DNA polymerase I, used as the mechanical-change sensor component in the second sequence-detection system.

[0024] Figures 21A-21E illustrate copy-strand extension catalyzed by the Klenow fragment of E. coli DNA polymerase I.

[0025] Figures 22A-22B illustrate, in one example, the variable-resistance component of the second sequence-detection system.

[0026] Figures 23A-23B illustrates one method for attaching a DNA-polymer tether to the DNA-polymerase mechanical-change sensor component 1304.

[0027] Figure 24 illustrates, in one example, several features of the DNA-polymer tether that mechanically couples the DNA polymerase to the variable-resistance component and that additionally forms a portion of the variable-resistance component.

[0028] Figure 25 shows numerous examples of low-current, high-resistance four-deoxynucleotide-monomer sequences and high-current, low-resistance four-deoxynucleotide-monomer sequences.

[0029] Figures 26A-26B illustrate, in one example, tailoring the responsiveness of the mechanical-to-electrical signal transduction through DNA-polymer-tether-sequence design.

[0030] Figure 27 illustrates, in one example, a DNA-polymer tether that features a repetitive deoxynucleotide sequence.

[0031] Figures 28A-28C illustrate, in one example, an approach used to distinguish the different deoxynucleotides added to the copy strand by the DNA polymerase in certain implementations of the second sequence-detection system.

[0032] Figures 29A-29C illustrate, in one example, use of a locking oligonucleotide to securely hold the DNA-polymer tether within the porin channel.

[0033] Figure 30 illustrates an applied-voltage cycle that is used, in one implementation, to prevent the occurrence of disruptive noise in the output voltage signal.

DETAILED DESCRIPTION

[0034] The current document discusses sequence detectors that generate a signal from which the sequence of types of entities in a target can be determined. A sequence detector described herein may be, for example, an electromechanical device. An electromechanical device is a device that includes both electrical and mechanical components, and that may include additional optical, fluid, and other components. Examples of a sequence detector detect a sequence by generating a signal from which the sequence of entity types in a target can be determined. The sequence detector may include a component, such as a microprocessor-controlled signal-analysis

component, that analyzes the signal to determine the sequence of entity types. A target contains a sequence of entities, each entity having a type. A sequence detector physically interacts with a target to generate a signal that varies as a mechanical-change sensor within the sequence detector specifically interacts with different types of entities in the target. A specific interaction is an interaction between the mechanical-change sensor and an entity that deterministically produces a mechanical change in the mechanical-change sensor characteristic of the entity type to which the entity belongs that is then transduced into a corresponding signal by the mechanical-change-to-signal transducer that is also characteristic of the entity type to which the entity belongs. A mechanical change may include a change in shape and/or size of the mechanical-change-to-signal transducer, change in the position of the mechanical-change-to-signal transducer relative to another component, change in the orientation of the mechanical-change-to-signal transducer relative to another component, and other such mechanical changes. In the current document, two different examples of targets and corresponding sequence detectors are discussed in two subsections, below. A first example target is linear sequence of macroscale objects connected together by a string or wire. The objects each have one of four different shapes. The type of an object corresponds to the object's shape. The first corresponding sequence detector produces a time-varying electrical signal as the target passes through a mechanical-change sensor and the mechanical-change sensor specifically interacts with each object, producing a mechanical change in the mechanical-change sensor that is transduced into an electrical signal characteristic of the object's shape by a variable resistor. A second example target is a biopolymer containing a sequence of monomers linked together by covalent bonds. There are four commonly occurring different types of monomers that differ from one another in chemical composition and structure, with additional types of monomers less frequently encountered, in the DNA and RNA polymers used as examples in the following discussion. The second corresponding sequence detector produces a time-varying electrical signal as the target biopolymer passes through a mechanical-change sensor and the mechanical-change sensor specifically associates with monomers to produce a mechanical change in the mechanical-change sensor that is, in turn, communicated to a variable-resistor component by a coupler. These two types of detectors are examples, and other configurations may also exist and be implemented.

A First Type of Mechanical-Change-Based Electromechanical Sequence Detector

[0035] Figure 1 illustrates, in one example, a target input to, and a determined sequence representation output by, a first (type of) mechanical-change-based sequence detector. The target 102 is a linear sequence of objects 104-116. In this example, there are four different types of objects included in the target: (1) a cylindrical-object type 120, referred to as type *a*; (2) a cubic-object type 122, referred to as type *b*; (3) a spherical-object type 124, referred to as type *c*; and (4) a four-sided-pyramid-object type 126, referred to as type *d*. The target 102 includes a linear spacing member 130, such as a wire or cord, to which the objects 104-116 are securely attached. The target is mechanically input to the first type of sequence detector which outputs a symbolic representation of the sequence of object types 132 within the target.

[0036] Figure 2 illustrates one implementation of a sequence-detection system based on the first type of mechanical-change-based sequence detector. The sequence-detection system 202 includes a mechanical-change-based sequence detector 204 into which a target 206 is mechanically input. The mechanical-change-based sequence detector mechanically outputs the target 208 and outputs an electrical signal 210 to an analysis subsystem, implemented as a computer program running on a computer system 212 in certain sequence detectors, which processes the electrical signal to determine the sequence of object types within the target and to output the determined sequence of object types 214 on a computer-display device 216. In alternative implementations, the analysis subsystem is implemented by processor-controlled subsystems other than general-purpose computer systems. The computer system may additionally store an encoded representation of the sequence in one or more memories and/or one or more mass-storage devices. The encoded representation of the sequence may be transmitted to remote computer systems and may be subsequently retrieved for display to a user and for further analysis.

[0037] Figures 3A-3F illustrates, in one example, the mechanical-change sensor component of the sequence-detection system shown in Figure 2. As shown in Figure 3A, the mechanical-change sensor component 302 is a funnel-shaped device comprising a rigid circular ring 304 to which a large number of flexible, spring-like tines, including tine 306, are attached. The tines are arranged as if lying on the surface of a conical section with inward, radial orientations, as shown in a top-down projection view in Figure 3B. Figure 3C shows a logical representation of the mechanical-change sensor component 310. In the series of figures comprising Figures 3D-3F, operation of the mechanical-change sensor component is illustrated.

An object of type *b* 312 (122 in Figure 1) is shown positioned behind the mechanical-change sensor component 302 in Figure 3D. In Figures 3E-3F, the object is mechanically translated through the mechanical-change sensor component 302. As it moves through the mechanical-change sensor component, the object pushes the flexible tines outward, as shown in Figure 3F, distorting of the funnel shape of the mechanical-change component.

[0038] Figure 4 illustrates, in one example, generation of a voltage signal from changes in the shape of the mechanical-change sensor component by a mechanical-change-detection subsystem within the sequence-detection system shown in Figure 2. On the left-hand side of Figure 4, the mechanical-change-detection subsystem 402 is shown when the funnel-like shape of the mechanical-change sensor component is undistorted. The right-hand side of Figure 4 shows the mechanical-change-detection subsystem 404 when the funnel-like shape of the mechanical-change sensor component is distorted by the presence of an object within the mechanical-change-sensor component. The mechanical-change-detection subsystem 402 includes the mechanical-change sensor component 302, a potentiometer 406, and a voltmeter 408. The mechanical-change sensor component 302 is mechanically connected to the potentiometer 406 by a cord or wire 410 attached, at one end, to one of the tines 412 of the mechanical-change sensor component and attached to a slidable potentiometer arm 414, at the other end. The cord or wire 410 passes over three freely rotating pulleys 416-418. The slidable potentiometer arm 414 is held in a first position by the tine 412 against the force of a weak spring 420 within a potentiometer-arm cylinder 422. When the mechanical-change sensor component is distorted by the presence of an object, as shown in the right-hand side of Figure 4, the tine 412 is forced downward, as a result of which the slidable potentiometer arm 414 is pulled upward by spring 420 within the potentiometer-arm cylinder 422. In this example, the potentiometer 406 act as a variable resistor. A variable resistor is a circuit element with a resistance to current flow that can be changed, and a potentiometer is one example of a variable resistor. In the position shown in the left-hand side of Figure 4, the potentiometer arm is connected to the potentiometer circuit 424 below resistor 426, as a result of which there is little or no voltage drop across the voltmeter 408. However, when the potentiometer arm is in the position shown in the right-hand side of Figure 4, the potentiometer arm is connected to the potentiometer circuit 424 at a point part way up the resistor 428, as a result of which there is a significant voltage drop across the voltmeter. Thus, the mechanical-change-detection subsystem generates a varying voltage signal in correspondence with a degree of distortion in the shape of

the mechanical-change sensor component. The magnitude of the output voltage signal corresponds to the degree of distortion of the mechanical-change sensor component.

[0039] Figure 5 illustrates, in one example, the internal components of the mechanical-change-based sequence detector 204 shown in Figure 2. The mechanical-change-based sequence detector includes the mechanical-change-detection subsystem 402 shown in Figure 4 and two electric-motor-driven pairs of counter-rotating geared drums 502 and 504 that feed the objects of a target 506 through the mechanical-change sensor component 302. As the target is pulled through the mechanical-change sensor component, the voltmeter outputs a voltage signal 210 that is input to an analysis subsystem (212 in Figure 2).

[0040] Figures 6A-6D illustrate, in one example, the voltage signals produced by each of the four different types of objects that occur within targets. Figure 6A shows a plot 602 of the voltage signal generated when an object of type *d* 604 passes through the mechanical-change sensor component. The voltage of the output signal is represented by a vertical axis 606 and the position of the object in a horizontal direction is represented by a horizontal axis 608 in the plot. Note that the position may be expressed either in a horizontal displacement or in time, assuming that the target moves through the sensor at a constant velocity. Because objects of type *d* are rotationally unstable with respect to an internal axis that passes through the top vertex and the center of the base, objects of type *d* tend to rotate back and forth about this axis as they pass through the mechanical-change sensor component. They also tend to rotate about four internal horizontal axes. As a result, the voltage signal 610 tends to oscillate as the object passes through the mechanical-change sensor. Objects of type *c* 612 produce a smooth and symmetrical signal 614, as shown in Figure 6B. Objects of type *b* 616 show a symmetrical signal 618 with minor oscillations due to slight rotational instability, as shown in Figure 6C. Objects of type *a* 620 also produce a symmetrical output signal 622 with slight oscillations, as shown in Figure 6D. The output voltage signals are analyzed by the computational analysis subsystem (212 in Figure 2). Data is collected from a region of each voltage-signal curve that begins when the voltage signal rises to half peak height and that ends when the signal falls back to half peak height, shown in each plot of Figures 6A-6D by a horizontal double-headed arrow, such as arrow 624, and vertical dashed lines, such as vertical dashed lines 626 and 628. The analysis subsystem computes, from the voltage-signal-magnitude data collected from each object-indicating voltage-signal curve, a mean

voltage magnitude μ , such as mean voltage magnitude 630, a variance σ^2 , such as variance 632, and an area A under the voltage-signal curve, such as area 634.

[0041] The computed values are obtained by collecting n sample voltage magnitudes v_i from different timepoints or displacements within the central portion of voltage-signal curve. In various different implementations, current magnitudes or other values may be instead sampled. A sampling rate of 1 KHz, for example, would provide 1000 sample voltage magnitudes. The area A is computed by discrete integration:

$$A = \sum_{i=1}^n v_i.$$

The mean voltage magnitude is computed as:

$$\mu = \frac{\sum_{i=1}^n v_i}{n}.$$

The standard deviation is computed as:

$$\sigma^2 = \frac{\sum_{i=1}^n (v_i - \mu)^2}{n}.$$

Finally, the standard deviation is computed as:

$$\sigma = \sqrt{\sigma^2}.$$

[0042] Figure 7 shows hypothetical analytical results produced by the sequence-detection system from a test target that includes 10 objects of each of the four object types a , b , c , and d . The analytical results for each of the four types of objects are shown in tables 702-705. Each table includes three columns corresponding to the computed voltage-magnitude mean, variance, and area for each voltage-signal curve output by the mechanical-change-based sequence detector. As commonly occurs in experimental data, the three computed values vary across the 10 instances of the four different types of objects. The analysis subsystem attempts to use the output voltage-signal curves to uniquely identify the type of each object passing through the mechanical-change sensor component.

[0043] A common approach for using output voltage signals is to choose a single computed value, such as the mean voltage magnitude, to differentiate each type of object from the remaining types of objects. Figure 8 shows a plot of the mean-voltage-magnitude data contained in the data

tables shown in Figure 7. A key 802 is shown in the upper left-hand portion of Figure 8. The key describes the different symbols used for plotting mean-voltage-magnitude values for each of the different types of objects. The lower portion of Figure 8 shows a plot 804 of the mean-voltage-magnitude data with respect to a horizontal axis 806 representing voltage magnitude. The mean-voltage-magnitude values for each different type of object cluster within subregions of the horizontal axis, as indicated in Figure 8 by the dashed ellipses 808-811. The mean voltage magnitudes for objects of type *c* fall within the voltage-magnitude range indicated by ellipse 808, for example. From the data plot shown in Figure 8, objects of type *c* are uniquely distinguishable from the remaining object types based on mean voltage magnitude, alone, since in this example there is no overlap between the range of mean voltage magnitudes for objects of type *c* and the ranges of mean voltage magnitude for objects of type *a*, *b*, and *d*. Similarly, objects of type *b*, the mean-voltage-magnitude values of which fall within the range represented by ellipse 811, are uniquely distinguishable from the remaining types of objects based on mean voltage magnitude, alone. However, the mean-voltage-magnitude ranges for objects of type *d* and *a*, represented by ellipses 809-810, almost completely overlap with one another, as a result of which it is not possible to distinguish between objects of type *d* and *a* using mean-voltage-magnitude values, alone.

[0044] In some examples, when a particular output signal is insufficient for distinguishing the different types of objects in a sequence, a sequence-detector designer would seek to incorporate an additional type of sensor into the sequence detector to produce an additional output signal, so that the combination of multiple output signals provides sufficient information for distinguishing the objects from one another. For example, change counters use separate size-detection sensors, weight detectors, and magnetic-susceptibility detectors to produce separate output signals that together provide an unambiguous output-signal-derived fingerprint for each type of coin. By contrast the currently described example sequence-detection systems compute multiple derived values from a single output voltage signal in order to differentiate each type of object in a target. As discussed above with reference to Figures 6A-6D and Figure 7, the analysis subsystem computes not only the mean voltage magnitude from the voltage-signal curve corresponding to an object, but also the variance. Figure 9 shows a two-dimensional plot of the data contained in the tables shown in Figure 7. The horizontal axis 902 represents the mean voltage magnitude obtained from voltage-signal curve, as in Figure 8, and the vertical axis 904 represents the variance obtained from the voltage-signal curve. As in Figure 8, dashed ellipses 906-909 surround clusters of data

points plotted for each of the different object types. As can be easily seen in Figure 9, the two-dimensional areas contained within these ellipses do not overlap. Thus, a pair of mean-voltage-magnitude and variance values computed from the single output voltage signal for a particular object contains sufficient information to unambiguously assign a type to the object. In other words, in this example the single output voltage signal produced by the mechanical-change-detection subsystem contains sufficient information for assigning a type to each object, but the information within each voltage-signal curve is, in a sense, two-dimensional.

[0045] As discussed above with reference to Figure 9, use of the two derived values, including mean voltage magnitude and variance, by the analysis subsystem of the sequence-detection system is sufficient, for the target described with reference to Figure 1, to identify each object or entity within the target. The analysis system uses the two derived values as coordinates to map the two derived values to a range area corresponding to a particular object or entity type. Were the range areas overlapping, then an additional derived value, such as the computed area below the voltage-signal curve, might be used to uniquely differentiate object types within targets. Figure 10 illustrates, in one example, use of three derived values for determination of object types. In Figure 10, each of three different derived values are represented by the three axes 1002, 1004, and 1006. Plotted data points for the four different object types fall into the four discrete and nonoverlapping elliptical range volumes 1010-1013. Generally, as the number of derived values is increased, the probability of overlap in the ranges of the derived values for the different object types decreases when the derived signals are reasonably orthogonal and sensitive to differences in object type.

[0046] Figure 11 summarizes the sequence detection system discussed above with reference to Figures 1-10. A target comprising a sequence of objects or entities 1102 is input to the sequence-detection system 1104 and is mechanically translated through a mechanical-change component 1106 by a mechanical-translator component 1108. A power source 1110 provides power for the mechanical translation. The mechanical-change component 1106 is mechanically coupled, by a coupler 1112, to a variable-resistance component 1114. A mechanical coupler joins two or more entities by a physical coupler, such as a string or cord, in a macroscale device, or a linear molecule, such as a DNA polymer, in nanoscale and microscale devices. The variable-resistance component 1114 provides a variable resistance to a current flow 1116 in response to motion of the coupler 1112, in turn induced by changes in the shape of the mechanical-change

component 1106. A power source 1118 drives the current flow 1116. The measurement component 1120 measures the potential in the current-flow channel or the current flow, itself, to produce an output electrical signal 1122 that varies with variation of the shape of the mechanical-change component 1106. The output signal is computationally processed by an analysis subsystem, not shown in Figure 11, to generate a representation of the sequence of object types in the target 1102. The sequence-detection system, discussed above with reference to Figures 1-11, is a macroscale device that determines the sequence of macroscale-object types within a target.

A Second Type of Mechanical-Change-Based Electromechanical Sequence Detector

[0047] The second (type of) sequence-detection system discussed in the current document is a mixed-scale device that includes macroscale, microscale, and nanoscale components. The second sequence-detection system determines the sequence of deoxynucleotide monomers within nucleic-acid polymers.

[0048] Figure 12 provides a table that compares the first sequence-detection system, discussed above with reference to Figures 1-11, and the second sequence-detection system, discussed in the current subsection of the current document. A first column 1202 in the table lists sequence-detection-system components, discussed above with reference to Figure 11. A second column 1204 of the table further describes each of the components listed in the first column with respect to the first sequence-detection system. A third column 1206 further describes each of the components listed in the first column with respect to the second sequence-detection system, further described below. In the first sequence-detection system, the power source for mechanical translation of the target is an electrical current obtained from a battery or from line current while, in the second sequence-detection system, the power source for mechanical translation of the target is chemical energy produced by hydrolysis of a phosphoanhydride bond and hydrolysis of inorganic pyrophosphate, as indicated in the first row 1208 of the table. In the first sequence-detection system, the mechanical translator that translates the target is two pairs of counterrotating electric-motor-driven geared drums while, in the second sequence-detection system, the mechanical translator is a Klenow fragment of E. coli DNA polymerase I, as indicated in the second row 1210 of the table. In the first sequence-detection system, the mechanical-change component is a funnel-shaped set of spring-like bristles, or tines, while, in the second sequence-

detection system, the mechanical-change component is a Klenow fragment of DNA polymerase I, as indicated in the third row 1212 of the table. In the first sequence-detection system, the coupler is a cord or wire while, in the second sequence-detection system, the coupler is a DNA polymer, as indicated in the fourth row 1214 of the table. In the first sequence-detection system, the variable resistor is a potentiometer while, in the second sequence-detection system, the variable resistor is a portion of the DNA-polymer coupler lying within an MspA-porin channel, as indicated in the fifth row 1216 of the table. In both sequence-detection systems, the current is an electrical current. In the first sequence-detection system, the charge carriers are conduction-band electron flowing through a metal wire while, in the second sequence-detection system, the charge carriers are positively and negatively charged ions, as indicated in the sixth row 1218 of the table, although, of course the electrodes are connected by current-carrying wires. In the first sequence-detection system, the power source for driving the current is obtained from a battery or from line current while, in the second sequence-detection system, the power source for driving the current is obtained from line current, as indicated in the seventh row 1220 of the table. In the first sequence-detection system, the current-flow or potential measurement device is a voltmeter while, in the second sequence-detection system, the current-flow or potential measurement device is a current-to-voltage converter, as indicated in the eighth row 1222 of the table. The first and second sequence-detection systems are thus similar to one another in configuration and operation, but include different specific types of components.

[0049] Figure 13 illustrates, in one example, a mechanical-change-based sequence detector that is included in the second sequence-detection system. In the following discussion, the mechanical-change-based sequence detector 1302 is referred to as a "cell." The illustration in Figure 13 does not reflect the relative sizes and volumes of the various components. In many implementations, the cell 1302 is a macroscale or microscale device while the mechanical-change sensor component 1304 is a nanoscale component. Figure 13 is intended to illustrate the overall configuration and relative positions and orientations of the various components of the mechanical-change-based sequence detector, rather than to accurately portray the relative scales of the components.

[0050] The cell 1302 includes a two-part vessel 1306, with a first solution-containing chamber 1308 separated from a second solution-containing chamber 1310 by a Teflon barrier 1312 and a lipid bilayer 1314. In one implementation, the lipid bilayer comprises 1, 2-diphytanoyl-sn-

glycerol-3-phosphocholine. The Teflon barrier includes an aperture 1316 that is covered by the lipid bilayer, so that the first solution-containing chamber 1308 and the second solution-containing chamber 1310 are separated only by the lipid bilayer within the aperture 1316. A narrow channel 1316 through the lipid bilayer is provided by a Mycobacterial porin ("MspA porin"), an octameric protein aggregate with eightfold rotational symmetry. The narrow channel is sufficiently wide to allow for passive diffusion of ions between the two solution-containing chambers. The first and second solution-containing chambers 1308 and 1310 contain a buffer solution at pH 8.0. In one implementation, the buffer solution includes 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid ("HEPES"), 150 mM KCl, 1 mM dithiothreitol ("DDT"), and 3 mM MgCl_2 . HEPES is a zwitterionic buffering compound. KCl provides ions that carry ionic current. Dithiothreitol is a reducing agent that promotes free sulfhydryl groups in proteins. MgCl_2 contributes Mg^{2+} ions that aids the catalytic activity of the mechanical-change sensor component 1304, discussed below. The first solution-containing chamber 1308 additionally contains deoxynucleotide triphosphates and primer-associated deoxyribonucleic-acid templates. The primer-associated deoxyribonucleic-acid templates 1318-1319 are the targets for which sequences of deoxynucleotide monomers are determined by the second sequence-detection system. In certain implementations, the second solution-containing chamber 1310 additionally contains locking components, discussed below. The solution within the second solution-containing chamber 1310 is in fluid contact with a positive electrode, or reference electrode 1320, and the solution within the first solution-containing chamber 1308 is in fluid contact with a negative electrode 1322. In one implementation, silver/silver-chloride electrodes are used. When a voltage is applied across the cell through the electrodes, negative ions flow through the porin channel towards the positive anode and positive ions flow through the porin channel towards the negative electrode. In one example, the downward flow of negative ions may be inhibited when a DNA or RNA polymer is resident within the pore. In other words, an electrical current is established within the porin channel by application of a voltage across the two electrodes 1320 and 1322. As further discussed below, the polarity of the applied voltage may be temporarily reversed, at various times during operation of the cell, by reversing the polarities of the electrodes.

[0051] In alternative implementations, rather than using a Teflon barrier and lipid bilayer, the aperture is produced in a silicon substrate or other type of substrate using a photolithographic process and a synthetic-polymer membrane is employed to prevent fluid communication between

the two chambers except through the porin channel. As one example, a triblock copolymer may be used for the membrane. In alternative implementations, MspA-porin variants may be employed, including a single-chain version or a version with fewer or greater than 8 subunits. Certain variants may comprise multiple subunits that differ in sequence. Other types of pore-containing biopolymers and synthetic polymers may alternatively be used in alternative implementations. Additional types of divalent metal ions may also be used, in alternative implementations, including Mn^{2+} . Non-catalytic metal ions, including Ca^{2+} and Sr^{2+} may also be used, in certain circumstances.

[0052] The mechanical-change sensor component 1304 is, in one implementation, a Klenow fragment of *E. coli* DNA polymerase I. The Klenow fragment may be obtained by removing the 5'→3' exonuclease structural domain from *E. coli* DNA polymerase I by treatment with a protease or by expressing the desired fragment from a genetically modified bacterial strain. The Klenow fragment retains the 5'→3' polymerization functionality. As discussed further, below, when supplied with a primer-associated DNA template and deoxynucleotide triphosphates, the Klenow fragment of *E. coli* DNA polymerase I catalyzes sequential polymerization of the deoxynucleotide triphosphates to form a copy DNA strand complementary in sequence to the template DNA strand. The coupling component 1324 is a DNA-polymer tether. It is attached to the mechanical-change sensor component 1304 and pulled into the porin channel by the voltage applied to the electrodes, since DNA polymers are negatively charged and migrate towards the positive electrode under an applied voltage. A small region of the DNA-polymer tether spanning a narrow constriction within the porin channel, along with the narrow constriction, act together as a variable resistor that regulates the flow of ions between the two solution-containing chambers to different extents depending on the position of the small region of the DNA-polymer tether relative to the narrow constriction, as further discussed below. Because the system comprising the Klenow fragment of *E. coli* DNA polymerase I and the porin exhibits differences for each different type of deoxynucleotide triphosphate that occupies the active site within the Klenow fragment of *E. coli* DNA polymerase I, the DNA-polymer tether has a different dynamic position relative to the narrow constriction within the porin channel when different deoxynucleotide triphosphates are specifically associated with the active site, which is reflected in a different dynamic current flow through the porin channel for each different type of deoxynucleotide triphosphate sequentially incorporated within the growing DNA copy strand. Specific association between a molecule and

an active site involves a key-in-lock or induced-fit type of association in which particular electrostatic and chemical features of the molecule associate with complementary electrostatic and chemical features of the active site, leading to larger binding affinities for the molecule or a class of molecules than for molecules that do not specifically associate with the active site. The differences may result from one or more of conformational changes, movement of the Klenow fragment relative to the poring channel, and other changes. Current-detection circuitry, discussed below, produces a voltage signal that varies in correspondence with variation in the current flow through the porin channel. In alternative implementations, many different polymerases, polymerase fragments, and other types of biomolecules that interact with the biopolymer target for sequencing may be used in place of the above-discussed Klenow fragment. Different types of natural and synthetic nucleotides may be used, including nucleotides with larger phosphate esters, such as deoxynucleotide hexaphosphates, with different carbohydrate components, with different bases, and different functional groups. Many additional types of mechanical-change components may be used for sequencing a variety of different types of target biopolymers and synthetic polymers, including enzymes and other proteins and protein/nucleic-acid complexes that interact with target proteins in sequence-specific fashions. In addition, it is important to note that the phrase "mechanical-change sensor component," when applied to the second sequence-detection system, indicates signal generation is a product of one or more of changes in the shape of the polymerase fragment, changes in the relative positions the polymerase fragment with respect to the porin, and/or changes in the orientations of the polymerase fragment with respect to the porin, as one example. In certain implementations, changes in the shape of the polymerase fragment provide the mechanical changes that lead to movement of the variable resistor. Ultimately, the mechanical-change sensor component, and interactions of the mechanical-change sensor component with the target and with the pore-containing component, produce a mechanical change in the position of the variable resistor. Coupling connectors and variable resistors other than DNA polymers may be used in alternative implementations.

[0053] Although the first sequence-detection system is a macroscale system and the second sequence-detection system is a mixed-scale system that includes macroscale and nanoscale components, the second sequence-detection system is analogous to the first sequence-detection system. Both sequence-detection systems employ a mechanical-change sensor component to generate a mechanical signal that varies with the type of object or entity currently being processed

by, or associated with, the mechanical-change sensor component. Both sequence-detection systems employ mechanical coupling to couple the mechanical-change sensor component to variable-resistance component. Both sequence-detection systems generate an output voltage signal by transduction of the mechanical signal produced by the mechanical-change sensor component into an electrical signal. Both sequence-detection systems employ computational analysis of the output signal to generate multiple derived values that are used together to identify the sequence of types of objects or entities in a target sequence. In the second sequence-detection system, the sequence of deoxynucleotide-monomer types detected is complementary to, and has reverse polarity with respect to, the sequence of deoxynucleotide-monomer types within the template-strand target.

[0054] Figures 14A-14B illustrate two different current-to-voltage converter circuits that are used separately or together in various implementations of the second sequence-detection system. In Figure 14A, a voltage source 1402 applies a voltage across the cell 1404. An inverting amplifier, or op amp, 1406 with a feedback loop 1408 containing a feedback resistor 1410 outputs a voltage signal proportional to the current flowing through the cell. A second op amp 1412 amplifies the voltage differential of its inputs to generate an amplified voltage signal that is passed through a frequency-correction circuit 1414 to produce a final output voltage signal 1416 proportional to current flow through the cell 1404. Figure 14B shows a current-to-voltage converter that uses a feedback capacitor rather than a feedback resistor.

[0055] Figure 15 illustrates, in one example, an array of cells that provides for parallel sequence determination. Parallel sequence determination may be used to increase the sequence-determination throughput of the system. The array of cells includes multiple cells, such as cell 1502, and an analysis subsystem 1504 that processes and analyzes the voltage-signal outputs from the multiple cells in parallel. In general, the sequences output by a sequence-detection system may contain errors due to a variety of different operational error sources often present in the sequence-detection-system components, including the mechanical-change sensor component, the mechanically coupled variable-resistance component, and the current or potential measuring circuitry. Therefore, depending on the level of accuracy desired, multiple identical targets may be sequenced and a consensus sequence may be computationally generated from the multiple sequences determined for the multiple identical targets. It is often the case that sequences for multiple targets are desired. In the array of cells illustrated in Figure 15, each cell may be loaded

with multiple copies of each of multiple different types of targets. The cells continuously produce sequence information from the multiple types of targets, with the targets processed in a nondeterministic order based on random association of primer-associated templates with the Klenow fragment of *E. coli* DNA polymerase I. The analysis subsystem 1502 continuously collects sequence information from multiple cells, such as cell 1504, assigns each sequence to a group of sequences generated by a particular target, and then compiles consensus sequences for each of the different types of target from the group of sequences obtained for each target type. Use of parallelism allows for rapid and efficient consensus-sequence determination for multiple targets.

[0056] Figures 16-19 illustrate deoxyribonucleic acids and peptides. Figure 16 illustrates a short DNA polymer. Deoxyribonucleic acid ("DNA") is a linear polymer, synthesized from four different types of deoxy nucleotide triphosphates that, when incorporated within the polymer, are referred to as deoxynucleotide monomers. The deoxynucleotide monomers include: (1) deoxyadenylate, abbreviated "A," a purine-containing deoxynucleotide; (2) deoxythymidylate, abbreviated "T," a pyrimidine-containing deoxynucleotide; (3) deoxycytidylate, abbreviated "C," a pyrimidine-containing deoxynucleotide; and (4) deoxyguanidylate, abbreviated "G," a purine-containing deoxynucleotide. The corresponding nucleosides, which lack phosphate groups attached through phosphodiester bonds to ribose hydroxyl oxygens, are referred to as deoxyadenosine, deoxythymidine, deoxycytidine, and deoxyguanosine. Figure 16 illustrates a short DNA polymer 1600, called an "oligomer" or "oligonucleotide," composed of the following subunits: (1) deoxyadenylate 1602; (2) deoxythymidylate 1604; (3) deoxycytidylate 1606; and (4) deoxyguanidylate 1608. The deoxynucleotide subunits are linked together through phosphodiester bonds 1610-1615 to form the DNA polymer. A linear DNA molecule, such as the oligomer shown in Figure 16, has a 5' end 1618 and a 3' end 1620. Often, the 5' end 1618 includes a phosphate group linked to the 5' hydroxyl oxygen through a phosphoester bond. A DNA polymer can be chemically characterized by writing, in sequence from the 5' end to the 3' end, the single letter abbreviations for the deoxynucleotide subunits that together compose the DNA polymer. For example, the oligomer 1600 shown in Figure 16 can be symbolically represented as "ATCG." A deoxynucleotide comprises a purine or pyrimidine base (e.g. adenine 1622 of the deoxyadenylate 1602), a deoxyribose sugar (e.g. deoxyribose 1624 of the deoxyadenylate 1602), and a phosphate group (e.g. phosphate 1626) that links one deoxynucleotide to another deoxynucleotide in the DNA polymer. Many non-natural nucleotides may be incorporated into

DNA-like and RNA-like polynucleotides. Example modified nucleobases that can be included in a polynucleotide, whether having a native backbone or analogue structure, include, inosine, xanthine, hypoxanthine, isocytosine, isoguanine, 2-aminopurine, 5-methylcytosine, 5-hydroxymethyl cytosine, 2-aminoadenine, 6-methyl adenine, 6-methyl guanine, 2-propyl guanine, 2-propyl adenine, 2-thioLiracil, 2-thiothymine, 2-thiocytosine, 5-halouracil, 5-halocytosine, 5-propynyl uracil, 5-propynyl cytosine, 6-azo uracil, 6-azo cytosine, 6-azo thymine, 5-uracil, 4-thiouracil, 8-halo adenine or guanine, 8-amino adenine or guanine, 8-thiol adenine or guanine, 8-thioalkyl adenine or guanine, 8-hydroxyl adenine or guanine, 5-halo substituted uracil or cytosine, 7-methylguanine, 7-methyladenine, 8-azaguanine, 8-azaadenine, 7-deazaguanine, 7-deazaadenine, 3-deazaguanine, 3-deazaadenine or the like. Certain nucleotide analogues cannot become incorporated into a polynucleotide, for example, nucleotide analogues such as adenosine 5'-phosphosulfate.

[0057] The DNA polymers that contain the organization information for living organisms occur in the nuclei of cells in pairs, forming double-stranded DNA helices. One polymer of the pair is laid out in a 5' to 3' direction, and is paired with a complementary polymer laid out in a 3' to 5' direction. The two DNA polymers in a double-stranded DNA helix are therefore described as being anti-parallel. The two DNA polymers, or strands, within a double-stranded DNA helix are bound to each other through attractive forces including hydrophobic interactions between stacked purine and pyrimidine bases and hydrogen bonding between purine and pyrimidine bases, the attractive forces emphasized by conformational constraints of DNA polymers. Because of a number of chemical and topographic constraints, double-stranded DNA helices are most stable when deoxyadenylate subunits of one strand hydrogen bond to deoxythymidylate subunits of the other strand and when deoxyguanylate subunits of one strand hydrogen bond to corresponding deoxycytidylate subunits of the other strand.

[0058] Figures 17A-17B illustrate the hydrogen bonding between the purine and pyrimidine bases of two anti-parallel DNA strands. Figure 17A shows hydrogen bonding between adenine and thymine bases of corresponding deoxyadenylate and deoxythymidylate subunits and Figure 17B shows hydrogen bonding between guanine and cytosine bases of corresponding deoxyguanylate and deoxycytidylate subunits. Note that there are two hydrogen bonds 1002 and 1003 in the adenine/thymine base pair, and three hydrogen bonds 1004-1006 in the guanosine/cytosine base pair, as a result of which GC base pairs contribute greater thermodynamic

stability to DNA duplexes than AT base pairs. AT and GC base pairs, illustrated in Figures 10A-B, are known as Watson-Crick ("WC") base pairs.

[0059] Figures 18A-18B illustrate double-stranded DNA. As shown in Figure 18A, two strands of DNA polymer 1802 and 1804 with complementary sequences form an anti-parallel double-stranded complex through hydrogen bonds between complementary bases of the two strands. The double-strand complexes are antiparallel because the two strands have opposite 5'-3' orientations or polarities. An adenine base on one strand 1806 is paired with a thymine base 1808 of the other strand and a guanine base on one strand 1810 is paired with a cytosine base 1812 on the other strand. The sequence of deoxynucleotides in the 5'-3' direction along one strand is complementary to the sequence of deoxynucleotides in the 3'-5' direction along the other strand. The complementarity of the two strands within an anti-parallel double-stranded DNA polymer is produced when a DNA polymerase catalyzes the polymerization of a copy strand onto a template strand. Figure 18B shows the familiar double-helix conformation of double-stranded DNA that occurs under physiological temperatures, pressures, pHs, and ion concentrations.

[0060] Figure 19 shows the names and chemical structures of the 20 common amino acids. Amino acids are polymerized in a ribosome-mediated translation process to form proteins. Amino acids are polymerized in a ribosome-mediated translation process to form proteins. Figure 19 shows a short four-amino-acid polymer 1902, referred to as a "peptide," that includes alanine, glutamic acid, glycine, and lysine monomers. Protein polymers commonly have hundreds to thousands of amino-acid monomers. Many proteins, such as the MspA porin, include multiple protein polymers. Under physiological conditions, proteins generally have complex three-dimensional conformations, such as the goblet-like conformation of the MspA porin octamer.

[0061] Figure 20 illustrates the polymerization reaction catalyzed by the Klenow fragment of E. coli DNA polymerase I, used as the mechanical-change sensor component in the second sequence-detection system. This polymerization reaction adds a nucleotide triphosphate 2002 to the 3' end of a growing copy strand 2004. In Figure 20, the remaining deoxynucleotide monomers in the copy strand are indicated by the arrow 2006 and the 5' label 2008. The deprotonated 3' hydroxyl 2010 of the 3'-terminal deoxynucleotide monomer in the copy strand carries out a nucleophilic attack on the α phosphate 2012 of the deoxynucleotide triphosphate 2002, forming a phosphodiester bond 2014 and displacing inorganic pyrophosphate 2016. This reaction has a relatively small change in free energy, under standard physiological conditions, but is driven by

subsequent hydrolysis of the pyrophosphate 2018, which is accompanied by a large free-energy change. The chemical energy released by hydrolysis of the pyrophosphate not only drives the polymerization reaction, but also drives translation of the DNA polymerase relative to the primer-associated template strand and may contribute to the different dynamic conformations exhibited by the DNA polymerase when different deoxynucleotide triphosphates are specifically associated with the active site and hydrogen bond with complementary deoxynucleotide monomers in the template strand.

[0062] Figures 21A-21E illustrate copy-strand extension catalyzed by the Klenow fragment of *E. coli* DNA polymerase I. Figures 21A-E all use the same illustration conventions, next described with reference to Figure 21A. The DNA polymerase is represented by a sphere 2102, which appears as a circle in cross-section. The active site within the DNA polymerase is represented by a vertically oriented, shaded rectangle 2104. The template DNA strand 2106 and the copy DNA strand 2108 are represented by a series of rectangles and discs. The purine and pyrimidine bases are represented by long, vertically oriented, labeled rectangles, such as rectangle 2110, which represents a guanine base. The ribose moiety within each deoxynucleotide monomer is represented by a small square, such as small square 2112. The phosphodiester bonds joining to deoxynucleotide monomers within a strand are represented by circles, such as circle 2114. Curved arrows, such as curved arrow 2116, indicate that the strands continue in the indicated directions.

[0063] Figure 21A shows the DNA polymerase without a deoxynucleotide triphosphate occupying the active site. The DNA polymerase is ready to receive a next deoxynucleotide triphosphate for addition to the 3' end of the copy strand. The active site of the DNA polymerase is a complex chemical environment that includes several bound magnesium ions and numerous functional groups of amino-acid-monomer sidechains that all contribute to specific binding of the template and copy strands, to specific binding of a deoxynucleotide triphosphate for addition to the 3' end of the copy strand via the reaction illustrated in Figure 20, and to catalysis of the polymerization reaction, including stabilization of one or more transition states. In Figure 21B, a deoxynucleotide triphosphate diffuses through channels in the DNA polymerase towards the active site. In at least one example, any of the four different types of deoxynucleotide triphosphate may approach the active site, but only a deoxynucleotide triphosphate that is complementary to the unpaired deoxynucleotide monomer of the template strand within the active site is stably associated with the active site for hydrogen bonding with the unpaired active-site-resident

template-strand deoxynucleotide monomer, as shown in Figure 21C. Stable association of the deoxynucleotide triphosphate with the active site is associated with a conformational change in the DNA polymerase, represented by an ellipsoid shape in Figure 21C. The actual conformational changes are complex, affecting multiple different domains within the DNA polymerase. The DNA-polymerase conformation is dynamic, and is generally associated with various types of subtle oscillation modes and relative motions of various structural domains. The specific association of each different type of deoxynucleotide triphosphate with the active site induces a different dynamical DNA-polymerase conformation, a change in the relative positions or orientations of the DNA-polymerase and porin, and/or other changes which are thought to be the source of the mechanical-change mechanical signal generated by the DNA polymerase acting as the mechanical-change sensor component of the second sequence-detection system.

[0064] In the sequence of figures that includes Figures 21C-21D, the polymerization reaction illustrated in Figure 20 occurs, forming the phosphodiester-bond bridge 2120 that incorporates the new deoxynucleotide monomer into the copy strand. The pyrophosphate 2122 is released from the active site. In the sequence of figures that includes Figures 21C-21E, the DNA polymerase translates relative to the template and copy strands to again form an active site 2126 without a nucleotide triphosphate, ready for specific incorporation of a subsequent deoxynucleotide triphosphate. Note that the conformation of the DNA polymerase has reverted to the original conformation, represented in Figure 21E by a spherical shape 2128. It should be noted that, in certain implementations, specific association of deoxynucleotide monomers with the active site, alone, without incorporation, can still lead to mechanical changes of the polymerase that can be transduced into a signal from which the target sequence can be determined. Incorporation of nucleotides into a copy strand is not necessary in these implementations.

[0065] Figures 22A-22B illustrate, in one example, the variable-resistance component of the second sequence-detection system. As shown in Figure 22A, the DNA-polymer tether 2202 is attached to the DNA polymerase 2204 and extends through the porin channel into the second solution-containing chamber 1310 of the cell below the lipid bilayer 2206. As the deoxynucleotide triphosphates are specifically associated with the active site of the DNA polymerase during copy-strand extension, changes in the dynamical conformation of the DNA polymerase result in translation of the DNA-polymer tether relative to a narrow constriction 2208 in the porin channel.

[0066] Figure 22B illustrates how translation of the DNA-polymer tether with respect to the narrow constriction in the porin channel leads to varying resistance to ion flow through the porin channel. Figure 22B shows four different positions of a DNA-polymer tether within the porin channel 2210-2213. The DNA-polymer tether is represented as a series of circles with different diameters. Large-diameter circles, such as circle 2216, represent one or more deoxynucleotide monomers that impart high resistance to ion flow through the porin channel when positioned within the narrow constriction 2218 of the born channel. Circles with increasingly smaller diameters represent one or more deoxynucleotide monomers that impart increasingly less resistance to ion flow through the porin channel when positioned within the narrow constriction. In the first position 2210, one or more a low-resistance deoxynucleotide monomers are positioned within the narrow constriction 2218, as a result of which there is relatively high rate of ion-current flow through the porin channel, as represented by the large number of positive 2220 and negative 2222 ion symbols shown entering the porin channel. As the DNA-polymer tether moves upward relative to the narrow constriction, in positions 2211-2213, one or more deoxynucleotide monomers that impart increasingly greater resistance to ion flow move into the narrow constriction, resulting in increasingly smaller rate of ion-current flow through the porin channel. Thus, positioning of the DNA-polymer tether within the porin channel varies the resistance to ion flow through the porin channel and transduces the mechanical mechanical-change signal generated by specific incorporation of deoxynucleotide triphosphates into the active side of the DNA polymerase into an electrical signal that is transduced, by the current-to-voltage-converter circuitry discussed above with reference to Figures 14A-14B, into an output voltage signal.

Specific Details of the Second Type of Mechanical-Change-Based Electromechanical Sequence

Detector

[0067] Figures 23A-23B illustrates one method for attaching a DNA-polymer tether to the DNA- polymerase mechanical-change sensor component 1304. In the wild-type E. coli DNA polymerase I, a leucine monomer 2302 occupies amino-acid-monomer position 790. A cysteine monomer 2304 is substituted for the leucine monomer at position 790 by genetic-engineering and biotechnology techniques to produce a mutated E. coli DNA polymerase I with a single free sulfhydryl moiety 2305, referred to a "P" in Figures 23A-23B. This is the attachment site for the

DNA-polymer tether. Any of various linker molecules, including the linker molecule 2306, can be incorporated at one end of the DNA-polymer tether. In many implementations, one or a few deoxynucleotide monomers constitute the 5' portion of the DNA-polymer tether 2308 and several tens of deoxynucleotide monomers constitute the 3' portion of the DNA-polymer tether 2310. A primary-amine functional group 2312 provides an attachment point for the DNA-polymer tether. The primary amine is attached to a crosslinking molecule 2314 that includes a sulfhydryl-reactive maleimide moiety 2316 to produce an activated tether 2318, referred to a "L" in Figures 23A-23B. As shown in Figure 23B, the mutated *E. coli* DNA polymerase I with a single free sulfhydryl moiety 2305, P, and the activated tether 2318, L, then react to covalently link the tether to the DNA polymerase 2320. The maleimide moiety 2316 of the activated tether reacts with the free sulfhydryl moiety 2304 of the mutated DNA polymerase to attach the DNA-polymer tether 2318 to the DNA polymerase 2320. This is but one example of many different possible methods for linking a DNA-polymer tether to a DNA polymerase.

[0068] Figure 24 illustrates several features of the DNA-polymer tether that mechanically couples the DNA polymerase to the variable-resistance component and that additionally forms a portion of the variable-resistance component. The DNA-polymer tether 2402 includes, in certain implementations, a linker 2404 through which the DNA-polymer tether is attached to the DNA polymerase, a pre-reporter region 2406 that includes a sufficient number of deoxynucleotide monomers to span the distance from the attachment point on the polymerase fragment and the reporter region when the polymerase fragment is seated within the porin, a reporter region 2408 that generally includes at least four deoxynucleotide monomers, and a post-reporter region that may include from several to many tens of deoxynucleotide monomers 2410. The reporter region 2408 is the region that lies within the constriction within the porin channel when the DNA polymerase exhibits various different conformations, exhibits different positions relative to the porin, and/or exhibits different orientations relative to the porin that together comprise the mechanical changes exhibited by the mechanical-change sensor component of the second sequence-detection system. The lengths of the various DNA-polymer-tether regions may vary with different implementations that use different porins and/or different DNA polymerases. The regions are defined by the distance between the attachment point of the DNA-polymer tether to the DNA polymerase and the narrow constriction in the porin channel as well as the range of

displacements in the position of the DNA-polymer tether induced by conformational changes of the DNA-polymerase.

[0069] As shown in the lower portion of Figure 24, the DNA-polymer tether may have a certain amount of dimensional flexibility, similar to an elastic member or spring member in a mechanical system. When a relatively low voltage is applied to the cell 2412, the portion of the DNA-polymer tether within the porin-channel constriction may be different than the portion of the tether within the porin-channel constriction when voltages of higher magnitudes are applied to the cell 2416, when the DNA-polymer tether inhabits a stretched or taut state 2418. A given deoxynucleotide monomer 2420 may lie at the boundary of the pre-reporter and reporter regions, in the relaxed state, but, in the stretched or taut state, the same deoxynucleotide monomer may lie well within the reporter region. By varying the applied voltage to the cell, relatively fine-grain adjustments are made to the resting or baseline position of the DNA-tether and, specifically, to the resting or baseline position of the reporter region relative to the narrow constriction in the porin channel. This provides a mechanism to calibrate the cell with respect to the output-voltage magnitude and with respect to the mechanical-to-electrical-signal-transduction responsiveness of the cell.

[0070] Different sequences of deoxynucleotide monomers within the reporter region provide different resistances to ion-current flow through the porin channel. Experiments with DNA-polymer tethers having different reporter-region deoxynucleotide-monomer sequences have led to the identification of a number of low-current, high-resistance reporter-region sequences and a number of high-current, low-resistance reporter-region sequences. Figure 25 shows numerous examples of low-current, high-resistance four-deoxynucleotide-monomer sequences and high-current, low-resistance four-deoxynucleotide-monomer sequences. The low-current sequences are shown in a first column 2502 and the high-current sequences are shown in a second column 2504. The letter "X" represents a baseless, depurinated deoxynucleotide 2506. Thus, the varying-resistance profile of the variable-resistance component with respect to mechanical translation of the DNA-polymer tether within the porin channel can be precisely designed by varying the deoxynucleotide-monomer sequence of the reporter region. It should be noted that the high-resistance deoxynucleotide-monomer sequences exhibit resistances to current flow through the porin channel greater than an average resistance exhibited by the various different possible deoxynucleotide-monomer sequences while low-resistance deoxynucleotide-monomer sequences

exhibit resistances to current flow through the porin channel less than the average resistance exhibited by the various different possible deoxynucleotide-monomer sequences.

[0071] Figures 26A-26B illustrate, in one example, tailoring the responsiveness of the mechanical-to-electrical signal transduction through DNA-polymer-tether-sequence design. In both Figures 26A and 26B, a reporter-region 2602 and 2632 is shown on the left-hand side of the figure and a corresponding output signal is plotted of the right-hand portion of the figure 2604 and 2605. Pairs of arrows, such as the pair of arrows 2606-2607, indicate the position of the narrow constriction within the porin channel relative to the reporter region in a reference position. In this example, the reporter region may move up or down within the porin channel by various distances, or displacements, as indicated by double-headed arrow 2608. The current-signal plot 2604 has a horizontal axis 2610 representing the displacement of the reporter region relative to the reference position and a vertical axis 2612 representing the magnitude of current flow through the porin channel. The reporter region 2602 is represented by a series of labeled squares, such as square 2614. Each square represents a single deoxynucleotide monomer or a short sequence of deoxynucleotide monomers. The letter "L" stands for low-current, the letter "M" stands for medium-current, and the letter "H" stands for high-current. Reporter region 2602 in Figure 26A varies slowly and symmetrically in resistance from a low-current, high-resistance portion 2616 to a high-current, low-resistance portion 2618 and then back to a low-current, high-resistance portion 2620. As a result, the current signal increases relatively slowly from a small negative displacement, or upward displacement 2620, to a larger positive or downward displacement 2624. Assuming that, during operation of the cell, the reporter displacement varies from $d_{\min}$ 2626 to the $d_{\max}$ 2628, the magnitude of the current flow falls within a response range indicated by double-headed arrow 2630. By contrast, the reporter region 2632 in Figure 26B features a sharp fall and rise in resistance and therefore produces a much steeper current increase 2634 over a similar displacement range 2636 and 2638 and a correspondingly larger current-flow response range 2640. Thus, the current-flow response to reporter-region displacement due to conformational changes in the DNA polymerase can be precisely tailored through careful design of the deoxynucleotide sequence within the DNA-polymer-tether reporter region. In general, steeper and non-linear response curves provide greater sensitivity to DNA-polymerase conformational changes, allowing detection of dynamic oscillations and relative inter-domain movement within the DNA-

polymerase, which may be additionally modified by interactions between the DNA polymerase and the MspA porin.

[0072] Figure 27 illustrates, in one example, a DNA-polymer tether that features a repetitive deoxynucleotide sequence. Figure 27 uses the same illustration conventions used in Figures 26A-26B. The DNA-polymer-tether sequence 2702 includes a repeating sequence of low-current, medium-current, and high-current deoxynucleotide tether portions. Arrows 2704-2706 show boundaries between the repeating sequence. Arrow 2708 represents a reference position of the DNA-polymer tether relative to the porin-channel constriction. A plot 2710 of the current signal generated by displacing the DNA-polymer tether upward and downward from the reference displacement 2708 shows a corresponding repeating signal form 2712. The advantage of a repeating-resistivity-sequence tether is that, by adjusting the voltage applied to the cell, as discussed above with reference to Figure 24, a current flow corresponding to a desired position on the output signal curve can be obtained regardless of the gross position of the DNA-polymer tether within the porin channel. In other words, the adjustment of the applied voltage can be considered to be a fine-grain tuning of the cell response, and, because of the repeating-sequence nature of the DNA-polymer tether, the fine-grain tuning is sufficient to select one of many identical optimal baseline positions for the DNA-polymer tether, even though the absolute position of the DNA-polymer tether within the porin channel may be difficult or impossible to determine.

[0073] As discussed above with reference to Figures 6A-10, various derived values are generated by the analysis subsystem of the sequence-detection system for use in identifying the objects or entities within a target. For certain implementations of the second sequence-detection system, the average current magnitude and the standard deviation within the central portion of the voltage signal produced by the current-to-voltage-converter circuitry, discussed above with reference to Figures 14A-14B, are generated by the analysis subsystem for voltage-signal curves generated for each deoxynucleotide added to the copy strand. However, an additional technique is needed to identify each added deoxynucleotide.

[0074] Figures 28A-28C illustrate, in one example, an approach used to distinguish the different deoxynucleotides added to the copy strand by the DNA polymerase in certain implementations of the second sequence-detection system. Figure 28A shows a plot of the mean-voltage-magnitude ranges for the four types of deoxynucleotide triphosphates within the active side of the DNA polymerase. All four ranges overlap, to some extent, with the average voltage-

signal-magnitude ranges for deoxythymidine triphosphate, deoxyadenosine triphosphate, and deoxycytosine triphosphate extensively overlapping one another. Figure 28B illustrates a two-dimensional plot using both the mean-voltage-magnitude and the standard deviation, as in Figure 9, discussed above. The two-dimensional plot effectively differentiates deoxyguanosine triphosphate 2802 from deoxycytosine triphosphate 2804, but the area ranges of deoxythymidine triphosphate and deoxyadenosine triphosphate are essentially coextensive 2806. In order to fully differentiate the four deoxynucleotide triphosphates, a modified adenine base 2810, 7-deaza adenine, is used in place of adenine 2812. When the modified deoxyadenosine-like triphosphate is present within the active site of the DNA polymerase, the DNA polymerase occupies a significantly different conformation than that occupied by the DNA polymerase when deoxyadenosine triphosphate is present within the active site. As a result, the area range of the modified deoxyadenosine-like triphosphate 2814 in a two-dimensional plot 2816 is no longer coextensive with the area range of deoxythymidine triphosphate 2818, and all four different deoxynucleotide triphosphates can be unambiguously identified based on the average voltage-signal magnitude and standard-deviation derived values generated from the output voltage signals generated when they are present in the active site.

[0075] Figures 29A-29C illustrate, in one example, use of a locking oligonucleotide as a locking component to securely hold the DNA-polymer tether within the porin channel. Figure 29A illustrates a DNA-the polymer tether with a locking oligonucleotide using the same illustration conventions used previously in Figure 24. As shown in Figure 29A, the locking oligonucleotide 2902 has a deoxynucleotide-monomer sequence complementary to a portion of the deoxynucleotide-monomer sequence of the DNA-polymer tether 2904, and hybridizes with the portion of the DNA-polymer tether via hydrogen bonding and base stacking. Figures 29B-29C illustrate use of the locking oligonucleotide during cell operation. Initially, the DNA polymerase 2906 and the DNA-polymer tether 2908 are unassociated with the porin 2910. Application of a voltage across the cell 2912 results in threading of the DNA-polymer tether into and through the porin channel 2914 and seating of the DNA polymerase within the pore 2916. The locking oligonucleotide then associates with the portion of the DNA-polymer tether 2918 extending into the second solution-containing chamber of the cell. As shown in Figure 29C, the locking oligonucleotide prevents the DNA-polymer tether from being pulled out of the porin channel 2920 when a reversed-polarity voltage 2922 is applied to the cell. As a result, restoring the original

polarity of the applied voltage 2924 reseats the DNA polymerase within the porin 2926. However, application of a much larger-magnitude voltage 2928 generates sufficient force to strip the locking oligonucleotide 2930 from the DNA-polymer tether 2932 and allow the DNA polymerase and DNA-polymer tether to fully dissociate from the port. In certain implementations, locking components other than locking oligonucleotides may be used. For example, a biotin moiety can be attached to tether and streptavidin can be used as the locking component. Other types of biopolymer locking components may also be used in alternative implementations.

[0076] There are several significant advantages obtained by using a locking oligonucleotide to secure the DNA-polymer tether within the porin channel. For certain types of DNA-polymerases, the voltage signal produced by the DNA polymerase and DNA-polymer-tether-based variable-resistance component exhibits extended periods of increased noisiness which masks the current-signal variations used to differentiate the different types of deoxynucleotide triphosphates specifically bound to the active site of the DNA polymerase. In order to prevent the occurrence of this type of noise, the polarity of the applied voltage is, in certain implementations, periodically reversed in order to unseat the DNA polymerase from the porin, after which the polarity of the applied voltage is again reversed to initiate a next interval of sequence detection. Figure 30 illustrates an applied-voltage cycle that is used, in one implementation, to prevent the occurrence of disruptive noise in the output voltage signal. In the plot shown in Figure 30, the horizontal axis 3002 represents time and the vertical axis 3004 represents the applied voltage. The applied voltage alternates between relatively long periods of a relatively large-magnitude applied voltage with normal first polarity, as shown in Figure 13, such as the period 3006, and relatively short periods during which the polarity of the applied voltage is reversed, such as period 3008.

[0077] Although the system, such as the sequence detector, disclosed in this document has been described in terms of particular implementations, it is not intended that the invention be limited to these implementations. Modification within the spirit of the invention will be apparent to those skilled in the art. For example, any particular sequence-detection system may use a variety of different types of mechanical-change sensor components, mechanical couplers, variable-resistance components, and electrical-signal-generation components and circuitry. For example, the above-described second sequence-detection system may use different DNA polymerases and DNA-polymerase fragments, may use different porins, may use different types of mechanical couplers, including different types of polymers, and may employ cells of many different shapes,

sizes, and configurations. Example polypeptide pores that may be used in place of MspA-porin include α -hemolysin, Mycobacterium smegmatis porin A, gramicidin A, maltoporin, OmpF, OmpC, PhoE, Tsx, F-pilus, SP1, mitochondrial porin (VDAC), Tom40, outer membrane phospholipase A, and Neisseria autotransporter lipoprotein (NaIP). Example polymerases, or fragments thereof, that may be used in place of the above-discussed Klenow fragment include T7 DNA polymerase, eukaryotic mitochondrial DNA polymerase γ , *E. coli* DNA Pol I, *Thermus aquaticus* Pol I, *Bacillus stearothermophilus* Pol I, eukaryotic DNA polymerases α , δ , and ϵ , DNA polymerase ζ , T4 DNA polymerase, Phi29 DNA polymerase, RB69 bacteriophage DNA polymerase, *E. coli* DNA polymerase III alpha subunit, polymerases derived from the Euryarchaeota subdomain of Archaea, eukaryotic polymerases Pol β , Pol σ , Pol λ , and Pol μ , *S. cerevisiae* Pol4, Pol η , Pol iota, Pol kappa, *E. coli* Pol IV (DINB), *E. coli* Pol V (UmuD'2C), retrovirus reverse transcriptases and eukaryotic telomerases, viral RNA polymerases such as T7 RNA polymerase, Eukaryotic RNA polymerases such as RNA polymerase I, RNA polymerase II, RNA polymerase III, RNA polymerase IV, RNA polymerase V, and Archaea RNA polymerase. Example alternative tethers include synthetic polymers such as PEG (polyethylene glycol), PPG (polypropylene glycol), PVA (polyvinyl alcohol), and PE (polyethylene). Additionally, DNA, RNA, PNA or LNA labels, morpholinos, or enantiomeric DNA, for example. Additionally, in one illustrative embodiment, isoG and isoC bases can be on the tether. Various different values derived from the output voltage signal can be employed as different dimensions for uniquely identifying the types of objects or entities in target sequences. The average voltage-signal magnitude, the variance, and the standard deviation computed from the voltage signals are only a few examples of the many different types of derived values that can be used for object or entity type identification. Many different types of modified deoxynucleotide triphosphates can be used for copy-strand extension in addition to 7-deaza deoxyadenosine triphosphate. Copy-strand elongation is only used to determine the template-strand sequence, and thus the chemical composition of the copy strand produced during sequence determination is irrelevant to successful operation of the second sequence-detection system. All combinations of the discussed concepts, components, and methods have been considered and comprehended.

[0078] It should be appreciated that all combinations of the foregoing concepts and additional concepts (provided such concepts are not mutually inconsistent) are contemplated as being part of the inventive subject matter disclosed herein. In particular, all combinations of

claimed subject matter appearing at the end of this disclosure are contemplated as being part of the inventive subject matter disclosed herein.

CLAIMS

1. A sequence-detection system comprising:
 - a mechanical-change sensor that exhibits one or more mechanical changes when specifically interacting with entities within a target, each entity having a type;
 - a mechanical-change-to-signal transducer that transduces the one or more mechanical changes into an electrical signal;
 - a mechanical coupler that couples the mechanical-change component to the mechanical-change-to-signal transducer; and
 - an analysis subsystem that determines a type of an entity within the target using the electrical signal by mapping two or more values derived from the electrical signal to a range volume, corresponding to the entity type, that does not substantially overlap a range volume corresponding to any other entity type.
2. The sequence-detection system of claim 1 wherein the analysis subsystem determines the types of entities within the target using the electrical signal by:
 - for each portion of the electrical signal corresponding to a particular entity within the target, computing n derived values from the portion of the electrical signal, wherein n is an integer greater than or equal to 2,
 - mapping the two or more derived values to an n -dimensional range volume corresponding to a particular entity type, that does not substantially overlap an n -dimensional range volume corresponding to any other entity type, and
 - assigning the particular entity type to the entity,optionally:
 - wherein the n derived values include one or more of:
 - an average of multiple signal-magnitude samples;
 - a variance of multiple signal-magnitude samples; and
 - a standard deviation of multiple signal-magnitude samples.
3. The sequence-detection system of claim 2 wherein the n derived values include one or more of:

an average of multiple signal-magnitude samples;
a variance of multiple signal-magnitude samples; and
a standard deviation of multiple signal-magnitude samples.

4. The sequence-detection system of claim 2 wherein the analysis subsystem stores a representation of a sequence of entity types assigned to a corresponding sequence of entities within the target in a memory for at least one of subsequent retrieval, transmission, display, and further analysis.

5. The sequence-detection system of claim 1 wherein the mechanical-change-to-signal transducer further comprises:

a variable resistor that regulates flow of an electrical current;
a portion of the mechanical coupler that couples the mechanical-change component to the variable resistor; and

a current-to-voltage converter that outputs a voltage signal that varies in correspondence with the rate of flow of the electrical current.

6. The sequence-detection system of claim 5

wherein the target is a primer-associated DNA template polymer;
wherein the mechanical-change component is a DNA polymerase;
wherein the mechanical coupler is a DNA-polymer tether attached, at one end, to the DNA polymerase; and

wherein the variable resistor is a reporter-region of the DNA-polymer tether that lies within a constriction within a porin channel.

7. A sequence-detection system comprising:

a nucleic-acid-polymerase mechanical-change component that exhibits mechanical changes when specifically associating with nucleotide polyphosphates within an active site;

a mechanical-change-to-signal transducer that transduces mechanical changes in the nucleic-acid-polymerase mechanical-change component into an output signal;

a mechanical coupler that couples the mechanical-change component to the mechanical-change-to-signal transducer; and

an analysis subsystem that determines a type of an entity within the target using the output signal by mapping two or more values derived from the output signal to a range volume, corresponding to the entity type, that does not substantially overlap a range volume corresponding to any other entity type.

8. The sequence-detection system of claim 7 wherein the mechanical coupler and the mechanical-change-to-signal transducer further comprise:

a tether attached, at one end, to the nucleic-acid-polymerase mechanical-change component and containing a reporter region that moves through a constriction in a porin channel as the nucleic-acid polymerase mechanical-change component undergoes mechanical changes, variably regulating flow of ions through the porin channel that constitutes an electrical current; and

a current-to-voltage converter that outputs a voltage signal that varies in correspondence with the rate of electrical current flow through the porin channel.

9. The sequence-detection system of claim 8 wherein the reporter region is an oligonucleotide and wherein different nucleotide subsequences within the reporter region provide different resistances to electrical current flow when positioned within the constriction in a porin channel.

10. The sequence-detection system of claim 9 wherein the reporter region is an oligodeoxynucleotide and wherein the reporter region exhibits a resistance to electrical current flow greater than an average resistance for oligodeoxynucleotides of similar lengths.

11. The sequence-detection system of claim 9 wherein the reporter region is an oligodeoxynucleotide and wherein the reporter region exhibits a resistance to electrical current flow less than an average resistance for oligodeoxynucleotides of similar lengths.

12. The sequence-detection system of claim 9

wherein a sensitivity of the output signal to changes in conformation of the nucleic-acid-polymerase mechanical-change component increases with an increase in a spatial rate of change in the resistance to electrical current flow provided by reporter region; and

wherein the reported region has a spatial rate of change in resistance that produces a specified output-voltage-signal signal response.

13. The sequence-detection system of claim 9

wherein the tether exhibits dimensional flexibility; and

wherein adjustment of a voltage applied across the porin channel moves the reporter region within the porin channel.

14. The sequence-detection system of claim 13

wherein the resistance of reporter region to electrical current flow has a spatial periodicity; and

wherein the sequence-detection system is calibrated to have a particular output-voltage-signal signal response by adjusting the voltage applied across the porin channel.

15. The sequence-detection system of claim 13 wherein a locking component is associated with a portion of the polymer tether that extends past the porin channel on the opposite side of the porin channel from the DNA-polymerase mechanical-change component to secure the polymer tether within the porin channel.

16. The sequence-detection system of claim 15 wherein a polarity of a voltage applied across the porin channel is periodically reversed to reduce noise in the output signal.

17. The sequence-detection system of claim 7 further comprising:

a two-chamber cell that includes

a first chamber containing a buffered solution of ionized salts, deoxynucleotide polyphosphates, and primer-associated DNA template polymers,

a second chamber containing a buffered solution of ionized salts, and

a barrier that isolates the first chamber from the second chamber within which the porin channel provides fluid communication between the first and second chambers;

a first electrode in fluid contact with the buffered solution in the first chamber and connected to the current-to-voltage converter; and

a second electrode in fluid contact with the buffered solution in the second chamber and connected to the current-to-voltage converter.