

METHODS, COMPOSITIONS, AND DEVICES FOR INFORMATION STORAGE

CROSS-REFERENCE TO RELATED APPLICATIONS

[001] This application claims priority to U.S. Provisional Patent Application Nos. 62/301,538, filed February 29, 2016, and 62/415,430, filed October 31, 2016, the contents of each of which are hereby incorporated by reference in their entireties.

FIELD

[002] The invention relates to novel methods, compositions and devices useful for information storage and retrieval, using nanopore devices to synthesize and sequence polymers, e.g., nucleic acids.

BACKGROUND

[003] There is a continuing demand to store ever more data on or in physical media, with storage devices getting ever smaller as their capacity gets bigger. The amount of data stored is reportedly doubling in size every two years, and according to one study, by 2020 the amount of data we create and copy annually will reach 44 zetabytes, or 44 trillion gigabytes. Moreover, existing data storage media such as hard drives, optical media, and magnetic tapes, are relatively unstable and become corrupted after prolonged storage.

[004] There is an urgent need for alternative approaches to storing large volumes of data for extended periods, e.g. decades or centuries.

[005] Some have proposed using DNA to store data. DNA is extremely stable and could in theory encode vast amounts of data and store the data for very long periods. See, for example, Bancroft, C., et al., *Long-Term Storage of Information in DNA*, Science (2001) 293: 1763-1765. Additionally, DNA as a storage medium is not susceptible to the security risks of traditional digital storage media. But there has been no practical approach to implementing this idea.

[006] WO 2014/014991, for example, describes a method of storing data on DNA oligonucleotides, wherein information is encoded in binary format, one bit per nucleotide, with a 96 bit (96 nucleotide) data block, a 19 nucleotide address sequence, and flanking sequences for amplification and sequencing. The code is then read by amplifying the sequences using PCR and sequencing using a high speed sequencer like the Illumina HiSeq machine. The data block sequences are then arranged in the correct order using the address tags, the address and flanking

sequences are filtered out, and the sequence data is translated into binary code. Such an approach has significant limitations. For example, the 96 bit data block could encode only 12 letters (using the conventional one byte or 8 bits per letter or space). The ratio of useful information stored relative to "housekeeping" information is low – approximately 40% of the sequence information is taken up with the address and the flanking DNA. The specification describes encoding a book using 54,898 oligonucleotides. The ink-jet printed, high-fidelity DNA microchips used to synthesize the oligonucleotides limited the size of the oligos (159-mers described were at the upper limit). Furthermore, reading the oligonucleotides requires amplification and isolation, which introduces additional potential for error. See also, WO 2004/088585A2; WO 03/025123 A2; C. BANCROFT: "Long-Term Storage of Information in DNA", Science (2001) 293 (5536): 1763c-1765; COX J P L: "Long-term data storage in DNA", Trends in Biotechnology (2001)19(7): 247-250.

[007] While the potential information density and stability of DNA make it an attractive vehicle for data storage, as has been recognized for over twenty-five years, there is still no practical approach to writing and reading large amounts of data in this form.

BRIEF SUMMARY

[008] We have developed a new approach to nucleic acid storage, using nanofluidic systems to synthesize the nucleic acid sequences and nanopore readers to read the sequences. Our approach allows for the synthesis, storage and reading of DNA strands which are hundreds, thousands or even millions of bases long. Because the sequences are long, only a relatively small proportion of the sequence is taken up with identifying information, so that the information density is much higher than in the approach described above. Moreover, in some embodiments, the nucleic acid as synthesized will have a specific location on a nanochip, so the sequence can be identified even without identifying information. The sequencing carried out in nanochambers is very rapid, and reading the sequence through a nanopore can be extremely rapid, on the order of up to one million bases per second. Since only two base types are required, the sequencing can be faster and more accurate than sequencing procedures that must distinguish among four nucleotide base types (adenine, thymine, cytosine, guanine). In particular embodiments, the two bases will not pair with one another and form secondary structures and will also be of different sizes. For

example, adenine and cytosine would be better for this purpose than adenine and thymine, which tend to hybridize, or adenine and guanine, which are of similar size.

[009] In some embodiments, this system can be used to synthesize long polymers encoding data, which can be amplified and/or released, and then sequenced on a different sequencer. In other embodiments, the system can be used to provide custom DNA sequences. In still other embodiments, the system can be used to read DNA sequences.

[010] The nanochips used in one embodiment contain at least two separate reaction compartments connected by at least one nanopore, which prevents at least some of the components from mixing, but allows as few as a single molecule of DNA, or other charged polymers, e.g., RNA or peptide nucleic acid (PNA), to cross from one reaction compartment into another in a controllable manner. The transfer of the polymer (or at least the end of the polymer to which monomers are added) from one compartment to another permits sequential manipulations/ reactions to the polymer, such as addition of bases, using enzymes which are prevented from crossing through the nanopore, for example because they are too large or because they are tethered to a substrate or bulky portion. Nanopore sensors report back on the movement or location of the polymer and its state, for example its sequence and whether the attempted reaction was successful. This allows data to be written, stored, and read, for example wherein the base sequence corresponds to a machine readable code, for example a binary code, with each base or group of bases corresponding to a 1 or 0.

[011] Accordingly, the invention includes, *inter alia*, the following embodiments,

- A nanochip for synthesis of an electrically charged polymer, e.g., DNA, comprising at least two distinct monomers, the nanochip comprising two or more reaction chambers separated by one or more nanopores, wherein each reaction chamber comprises an electrolytic fluid, one or more electrodes to draw the electrically charged polymer into the chamber and one or more reagents to facilitate addition of monomers or oligomers to the polymer. The nanochip may optionally be configured with functional elements to guide, channel and/or control the DNA, it may optionally be coated or made with materials selected to allow smooth flow of DNA or to attach the DNA, and it may comprise nanocircuit elements to provide and control electrodes proximate to the nanopores. For example, the one or more nanopores may optionally each be associated with

electrodes which can control the movement of the polymer through the nanopore and/or detect changes in electric potential, current, resistance or capacitance at the interface of the nanopore and the polymer, thereby detecting the sequence of the polymer as it passes through the one or more nanopores. In particular embodiments, the oligomers are synthesized using polymerases or site specific recombinases. In some embodiments, the polymer is sequenced during the course of synthesis, to allow for the detection and optionally correction of mistakes. In some embodiments, the polymer thus obtained is stored on the nanochip and can be sequenced when it is desired to access the information encoded in the polymer sequence.

- A method of synthesizing a polymer, e.g., DNA, using a nanochip as described.
- A single stranded DNA molecule wherein the sequence consists essentially of only nonhybridizing nucleotides, for example adenine and cytosine nucleotides (As and Cs), which are arranged in sequence to correspond to a binary code, e.g., for use in a method of data storage.
- A double stranded DNA comprising a series of nucleotide sequences corresponding to a binary code, wherein the double stranded DNA further comprises
- A method of reading binary code encoded in DNA, comprising using a nanopore sequencer.
- A method of data storage and devices therefor, using the above nanochip to make an electrically charged polymer, e.g., DNA, comprising at least two distinct monomers, wherein the monomers are arranged in sequence to correspond to a binary code.

[012] Further aspects and areas of applicability of the present invention will become apparent from the detailed description provided hereinafter. It should be understood that the detailed description and specific examples, while indicating preferred embodiments of the invention, are intended for purposes of illustration only and are not intended to limit the scope of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[013] The present invention will become more fully understood from the detailed description and the accompanying drawings, wherein:

[014] Figure 1 shows a diagram of a simple two-chamber nanochip design, with a dividing membrane perforated by a nanopore, and electrodes on either side of the membrane.

[015] Figures 2 and 3 show how the charged polymer, e.g. DNA, is drawn towards the anode.

[016] Figures 4 and 5 show that the polymer can be moved back by reversing the polarity of the electrodes.

[017] Figure 6 shows a two chamber nanochip design for DNA synthesis, in which a polymerase enzyme is located in one chamber, a de-blocking enzyme is in the other chamber, and neither can pass through the nanopore.

[018] Figure 7 shows addition of an adenine nucleotide when a 3'-blocked dATP (A) flows through left chamber, and the current is set 'forward' to bring the DNA into the chamber.

[019] Figure 8 shows deprotection of the oligonucleotide so an additional nucleotide can be added. For example, deprotection occurs after moving the DNA into the chamber by setting the current to 'reverse'.

[020] Figure 9 shows addition of a 3'-blocked dCTP (C). In certain embodiments, fluid flow is used to exchange the contents of this chamber, e.g., as depicted, previously there was 'A' in this chamber.

[021] Figure 10 shows how multiple separate retaining chambers can be provided while the flow chamber becomes a single lane to provide reagents.

[022] Figure 11 shows an approach to keeping the DNA associated with its chamber, by attaching to the chamber (upper DNA fragment in figure) or by coupling to a bulky group that cannot get through the nanopore (lower DNA fragment in figure). In this system, the end of the DNA can still move into the flow chamber and receive additional nucleotides, but the other end remains in the retaining chamber.

[023] Figure 12 shows a configuration where the DNA is attached to the wall of the chamber and controlled by multiple electrodes.

[024] Figure 13 shows how the DNA can be retained in the chamber when desired, simply by controlling the polarity of the electrodes.

[025] Figure 14 shows an array with free-flowing reagents through both sides, with the DNA bound to the surface of a chamber.

- [026] Figure 15 shows an alternate design with the electrodes on the sides adjacent to the dividing membrane, which allows for less expensive manufacture.
- [027] Figure 16 shows a three-compartment arrangement, where the DNA can be moved from compartment to compartment by the electrodes. This system does not require significant flow of reagents during synthesis.
- [028] Figure 17 shows an example of how reagents could be configured in a three compartment arrangement.
- [029] Figure 18 depicts an oligonucleotide tethered adjacent to a nanopore, where the nanopore has electrode elements on either side of the membrane.
- [030] Figure 19 depicts a series of DNA molecules attached along a membrane comprising nanopores and each under control of electrodes adjacent to a nanopore, with a flow lane on either side of the membrane. For example, as depicted, the left flow lane provides a flow of buffer wash / 3'-blocked dATP (A) / buffer wash / 3'-blocked dCTP (C) / buffer wash, wherein the DNA molecules are brought into the flow chamber only when the desired nucleotide is present. The right lane provides deblocking agent(s) to deprotect the 3' end of the nucleotide and allow for addition of another nucleotide. In one embodiment, the deblocking agent(s) flow when the left lane is being washed with buffer. In another embodiment, the deprotecting agent(s) are too bulky to cross to the left lane via the nanopores.
- [031] Figures 20-22 depict the schematically the proof of concept experiments wherein the bits used to encode the data are short oligomers attached using topoisomerase.
- [032] Figure 23 depicts a format for a nanopore sequencer wherein the polymer sequence is read using capacitive variance. In this capacitive readout scheme, electrodes form the top and bottom plates of a capacitor, separated by a membrane comprising a nanopore. The capacitor is embedded in a resonant circuit, wherein a pulsating direct current can draw the charged polymer through the nanopore. The change in capacitance is measured as the polymer, e.g. DNA, passes through the nanopore, using high frequency impedance spectroscopy. A major advantage of this approach, particularly with DNA, is that the measurement frequency can be very high (effectively a measurement for every cycle, so a 100MHz frequency corresponds to 100 million measurements per second), and much greater than the rate of transfer of monomers through the nanopore (DNA, for example, unless somehow constrained, will pass through the nanopore in response to electrical current at a speed on the order of 1 million nucleotides per second).

[033] Figure 24 depicts a dual addition chamber layout, suitable for adding two different types of monomers or oligomers, e.g., for 2-bit or binary encoding. The upper part of the figure shows a top view. The lower part shows a side view cross-section. The full device in this embodiment can be assembled from up to 3 independently fabricated layers and joined by wafer bonding, or may be formed by etching a single substrate. The chip comprises an electrical control layer (1), a fluidics layer (2) which contains the two addition chambers atop a reserve chamber, with the charged polymer (e.g., DNA) anchored between nanopore entrances to the first and second addition chambers, and an electrical ground layer (3).

[034] Figure 25 depicts the operation of the dual addition chamber layout of Figure 24. It will be observed that at the base of each addition chamber, there is a nanopore (4). The nanopore is made, for example, by drilling with FIB, TEM, wet or dry etching, or via dielectric breakdown. The membrane (5) comprising the nanopores is, e.g., from 1 atomic layer to 10's of nm thick. It is made from, e.g., SiN, BN, SiOx, Graphene, transition metal dichalcogenides e.g. WS₂ or MoS₂. Underneath the nanopore membrane (5) there is a reserve or deblocker chamber (6), which contains reagents for deprotection of the polymer following addition of a monomer or oligomer in one of addition chambers (it will be recalled that the monomers or oligomers are added in end-protected form, so that only a single monomer or oligomer is added at a time). The polymer (7) can be drawn into or out of the addition chambers by changing the polarity of the electrodes in the electrical control layer (1).

[035] Figure 26 depicts a top view of similar layout to figures 24 and 25, but here there are four addition chambers which share a common reserve or deblocker chamber and the polymer is tethered at a position (9) with access to each of the four chambers. The cross section of this layout would be as depicted in Figures 24 and 25, and the charged polymer can be moved into each of the four addition chambers by operation of the electrodes in the electrical control layer (1 in Figure 24).

[036] Figure 27 depicts a top view of a nanopore chip having multiple sets of dual addition chambers as depicted in Figure 24 and 25, allowing multiple polymers to be synthesized in parallel. The monomers are (here dATP and dGTP nucleotides represented as A and G) are loaded into each chamber via serial flow paths. One or more common deblocker flow cells allows for the polymers to be deprotected after addition of a monomer or oligomer in one of the addition chambers. This also allow the polymers to be detached on demand (for example using

a restriction enzyme in the case of DNA, or a chemical detachment from the surface adjacent to the nanopore, and collected externally. In this particular embodiment, the deblocker flow cells are perpendicular to the fluidics loading channels used to fill the addition chambers.

[037] Figure 28 depicts further details of the wiring for the dual addition chamber layouts. The electrical control layer (1) includes wiring made from metal or polysilicon. The wiring density is increased by 3D stacking, with electrical isolation provided by dielectric deposition (e.g., via PECVD, sputtering, ALD etc). The contact (11) to the top electrode by in the addition chamber in in one embodiment is made using Through Silicon Via (TSV) by Deep Reactive Ion Etch (DRIE) (cryo or BOSCH process). Individual voltage control (12) allows for each addition chamber to be addressed individually, allowing fine control of the sequence of multiple polymers in parallel. The right side of the figure depicts a top view illustrating wiring to multiple addition cells. The electrical ground layer (3) may be common (as shown) or split to reduce cross coupling between the cells.

[038] Figure 29 depicts an alternative configuration where the control electrodes (13) for the addition chambers may be deposited on the side of the chamber in a wrap around fashion instead of at the top of the chamber.

[039] Figure 30 depicts a SDS-PAGE gel confirming that topoisomerase addition protocol as described in Example 3 works, with bands corresponding to the expected A5 and B5 products being clearly visible. .

[040] Figure 31 depicts an agarose gel confirming that the PCR product of Example 5 is the correct size. Lane 0 is a 25 base pair ladder; lane 1 is product of experiment, line corresponding to expected molecular weight; lane 2 is negative control #1; lane 3 is negative control #2; lane 4 is negative control #4.

[041] Figure 32 depicts an agarose gel confirming that the restriction enzyme as described in Example 5 produces the expected product. The ladder on the left is a 100 base pair ladder. Lane 1 is undigested NAT1/NAT9c, Lane 2 is digested NAT1/NAT9c. Lane 3 is undigested NAT1/NAT9cI, Lane 4 is digested NAT1/NAT9cI.

[042] Figure 33 depicts Immobilization of DNA near nanopore. Panel (1) shows DNA with an origami structure on one end in the left chamber (in the actual nanochip, there initially are many such origami structures in the left chamber). Panel (2) illustrates the system with anode on the right, which drives the DNA to the nanopore. While the DNA strand is able to transit the

nanopore, the origami structure is too large to pass through, so the DNA is 'stuck'. Turning the current off (panel 3) allows the DNA to diffuse. With suitable chemistry, the end of the DNA strand is able to bind when it comes in contact with the surface near the nanopore. In panel (4) a restriction enzyme is added, which cuts the origami structure from the DNA. The chamber is washed to remove enzyme and residual DNA. The final result is a single DNA molecule attached near a nanopore, able to be moved back and forth through the nanopore.

[043] Figure 34 depicts a basic functioning nanopore. In each panel, the y-axis is current (nA) and the x-axis is time (s). The left panel "Screening of RF Noise" illustrates the utility of the Faraday cage. A chip with no nanopore is placed in the flow cell and 300mV applied. When the lid of the Faraday cage is closed (first arrow) the noise reduction can be seen. A small spike occurs when the latch is closed (second arrow). Notice the current is ~0 nA. After pore manufacture (middle panel), application of 300mV (arrow) results in a current of ~3.5 nA. When DNA is applied to the ground chamber and +300mV is applied DNA translocations (right panel) can be observed as transient decreases in the current. (Note, in this case the TS buffer is used: 50mM Tris, pH 8, 1M NaCl). Lambda DNA is used for this DNA translocation experiment.

[044] Figure 35 depicts a simplified picture illustrating the main features of the DNA origami structure: a large single stranded region, the cubic origami structure, and the presence of 2 restriction sites (SwaI and AlwN1) near the origami structure.

[045] Figure 36 depicts an electron microscope image of the manufactured DNA origami structure, and demonstrates the expected topology. Origami is made in 5mM Tris base, 1 mM EDTA, 5 mM NaCl, 5mM MgCl₂. In order to maintain the origami structure, it is preferable to have Mg⁺⁺ concentrations of ~5mM or Na⁺/K⁺ concentrations around 1M. The origami structure is stored at 4°C at 500nM.

[046] Figure 37 depicts a restriction digestion of the DNA origami to confirm correct assembly and function. The lane on the far left provides MW standards. The restriction sites are tested by digesting the origami with AlwN1 and SwaI. The four test lanes contain reagents as follows (units are microliters):

	1	2	3	4
origami	10	10	10	10
SwaI	-	1	-	1
AlwNI	-	-	1	1
NEB 3.1 10x	2	2	2	2
water	8	7	7	6

[047] Test lane (1) is a negative control; (2) is digestion with SwaI; (3) is digestion with AlwNI; (4) is double digestion with SwaI/AlwNI. Digestion is performed at room temperature for 60 minutes, followed by 37° C for 90 minutes. Agarose gel 1/2x TBE-Mg (1/2x TBE with 5mM MgCl₂), visualized with ethidium bromide staining. Individual digestion with either enzyme shows no mobility effect in a gel, but digestion with both enzymes together (lane 4) results in two fragments of different lengths, as expected.

[048] Figure 38 depicts binding of biotin-labeled oligonucleotides to streptavidin-coated beads vs. binding to control BSA coated beads. The y-axis is fluorescence units, 'pre-binding' is oligo fluorescence from test solution prior to binding beads, (-) controls are fluorescence seen after binding to two different batches of BSA-conjugated beads, SA-1 and SA-2 are fluorescence seen after binding to 2 different batches of streptavidin-conjugated beads. A small apparent amount of binding is observed with BSA-conjugated beads, but much larger binding is seen with the streptavidin-conjugated beads.

[049] Figure 39 depicts binding of biotin-labeled oligonucleotides to streptavidin-coated beads vs. binding to control BSA coated beads in different buffer systems, MPBS and HK buffer. The left bar 'Neg Ctrl' is the oligo fluorescence from test solution prior to binding the beads. Middle column shows fluorescence of 'BSA beads' and right column of 'SA beads' after binding to BSA or streptavidin beads respectively. In both buffer systems, the fluorescence is reduced by the streptavidin beads relative to controls, indicating that the biotin-labeled oligonucleotides are binding well to streptavidin-coated beads in different buffer systems.

[050] Figure 40 depicts a functioning conjugated SiO₂ nanopore, wherein the surface is streptavidin coated on one side and BSA coated on the other. The x-axis is time and the y-axis is current. The dot shows the point where the current is reversed. There is a brief overshoot when

the current is reversed, then the current settles to approximately the same absolute value. The nanopore shows a current of $\sim +3\text{nA}$ at 200mV and $\sim -3\text{nA}$ at -200mV .

[051] Figure 41 shows a representation of an origami DNA structure inserted into a nanopore.

[052] Figure 42 shows a representation of attachment of the single stranded DNA to the streptavidin-coated surface adjacent to the nanopore.

[053] Figure 43 shows experimental results of an origami DNA attached to the surface near a nanopore. Current is $+ \text{ or } - \sim 2.5\text{nA}$ in both directions, which is less than the original current of $\pm \sim 3\text{nA}$, reflecting partial obstruction by the origami structure. The x-axis is time (s), y-axis is current (nA), circles represent voltage switch points.

[054] Figure 44 shows the insertion of origami DNA, resulting in a slight drop in current. The origami immediately exits the nanopore when the current is released. The x-axis is time (s), y-axis is current (nA), circles represent voltage switch points.

[055] Figure 45 shows a representation of controlled movement of a DNA strand back and forth through a nanopore by application of current. On the left side the DNA is in the pore, so the observed current will be lower than if there was no DNA in the pore. When the current is reversed (right side) there is no DNA in the pore so the current will be unchanged.

[056] Figure 46 shows experimental results confirming this representation. When a positive voltage is applied the current is $\sim 3\text{nA}$, comparable to the current typically observed when the pore is open. When the voltage is reversed the current is $\sim -2.5\text{nA}$. This is lower than the current typically seen when the pore is open, and corresponds to the current typically observed when the pore is blocked by a strand of DNA. Several sequential voltage switches show consistent results, suggesting that the DNA is alternating in configuration as depicted in Figure 45.

[057] Figure 47 shows different conjugation chemistries to link the DNA to the surface adjacent to the nanopore.

DETAILED DESCRIPTION

[058] The following description of the preferred embodiment(s) is merely exemplary in nature and is in no way intended to limit the invention, its application, or uses.

[059] As used throughout, ranges are used as shorthand for describing each and every value that is within the range. Any value within the range can be selected as the terminus of the range. In addition, all references cited herein are hereby incorporated by referenced in their entireties.

In the event of a conflict in a definition in the present disclosure and that of a cited reference, the present disclosure controls.

[060] Unless otherwise specified, all percentages and amounts expressed herein and elsewhere in the specification should be understood to refer to percentages by weight. The amounts given are based on the active weight of the material.

[061] “Nanochip” as used herein refers to a nanofluidic device, comprising multiple chambers containing fluid and optionally channels allowing for fluid flow, wherein the critical dimensions of the features of the nanochip, for example the width of the elements dividing the chambers from one another, are from one atom to 10 microns in thickness, e.g., smaller than one micron, e.g. 0.01-1 micron. The flow of materials in the nanochip may be regulated by electrodes. For example, as DNA and RNA are negatively charged, they will be drawn to a positively charged electrode. See, e.g., Gershow, M, et al., *Recapturing and Trapping Single Molecules with a Solid State Nanopore*, Nat Nanotechnol. (2007) 2(12): 775–779, incorporated herein by reference. The flow of fluids may in some cases also be regulated by gate elements, and by flushing, injecting, and/or suctioning fluids into or out of the nanochip. The system is capable of precise multiplexed analysis of nucleic acids (DNA/RNA). In certain embodiments, the nanochip can be made of a silicon material, for example silicon dioxide or silicon nitride. Silicon nitride (e.g., Si_3N_4) is especially desirable for this purpose because it is chemically relatively inert and provides an effective barrier against diffusion of water and ions even when only a few nm thick. Silicon dioxide (as used in the examples herein) is also useful, because it is a good surface to chemically modify. Alternatively, in certain embodiments, the nanochip, may be made in whole or in part out of materials which can form sheets as thin as a single molecule (sometimes referred to as single layered materials), for example graphene, e.g., as described in Heerema, SJ, et al, *Graphene nanodevices for DNA sequencing*, Nature Nanotechnology (2016) 11: 127–136; Garaj S et al., *Graphene as a subnanometre trans-electrode membrane*, Nature (2010) **467** (7312), 190–193, the contents of each of which are incorporated herein by reference, or a transition metal dichalcogenide, e.g., molybdenum disulfide (MoS_2) as described in Feng, et al., *Identification of single nucleotides in MoS_2 nanopores*, Nat Nanotechnol. (2015) 10(12):1070–1076, the contents of which are incorporated herein by reference, or boron nitride, as described in Gilbert, et al. *Fabrication of Atomically Precise Nanopores in Hexagonal Boron Nitride*, eprint arXiv:1702.01220 (2017).

[062] In some embodiments, the nanochip comprises such a single layered material which is relatively stiff and inert, e.g., at least as inert and stiff as graphene, such as MoS₂. Single layered materials may, for example be used as all or part of the membrane comprising the nanopore. The nanochip may be lined in parts with metal, for example the walls may be layered (e.g. metal – silicon nitride – metal), and the metal can then be configured to provide a controllable pair of electrodes near the nanopore, so that the nucleic acid can be moved back and forth through the nanopore by electromotive force, and also can be sequenced by measuring the change in electric potential as the nucleic acid passes through the nanopore.

[063] Nanochip nanofluidic devices for sequencing DNA are generally known, for example as described in Li, J., et al, *Solid-state nanopore for detecting individual biopolymers*, Methods Mol Biol. (2009)544:81-93; Smeets RM, et al. *Noise in solid-state nanopores*, PNAS (2008)105(2):417-21; Venta K, et al., *Differentiation of short, single-stranded DNA homopolymers in solid-state nanopores*, ACS Nano. (2013)7(5):4629-36; Briggs K, et al. *Automated fabrication of 2-nm solid-state nanopores for nucleic acid analysis*, Small (2014)10(10):2077-86; and Chen Z, *DNA translocation through an array of kinked nanopores*, Nat Mater. (2010)9(8):667-75; the entire contents of each of which are incorporated herein by reference, e.g. for their teachings on the design and manufacture of nanochips comprising nanopores.

[064] “Nanopore” as used herein is pore having a diameter of less than 1 micron, e.g., 2-20 nm diameter, for example on the order of 2-5 nm. Single stranded DNA can pass through a 2 nm nanopore; single or double stranded DNA can pass through a 4 nm nanopore. Having a very small nanopore, e.g., 2-5 nm, allows the DNA to pass through, but not the larger protein enzymes, thereby allowing for controlled synthesis of the DNA (or other charged polymer). Where larger nanopores (or smaller protein enzymes) are used, the protein enzyme may be conjugated to a substrate that will prevent it from passing through the nanopore, e.g. to a larger molecule, such as a larger protein, to a bead, or to a surface in the chamber. Different types of nanopores are known. For example, biological nanopores are formed by assembly of a pore-forming protein in a membrane such as a lipid bilayer. For example, α -hemolysin and similar protein pores are found naturally in cell membranes, where they act as channels for ions or molecules to be transported in and out of cells, and such proteins can be repurposed as nanochannels. Solid-state nanopores are formed in synthetic materials such as silicon nitride or

graphene e.g., by configuring holes in the synthetic membrane, e.g. using feedback controlled low energy ion beam sculpting (IBS) or high energy electron beam illumination. Hybrid nanopores can be made by embedding a pore-forming protein in synthetic material. Where there is a metal surface or electrode at either end or either side of the nanopore, a current flow across the nanopore may be established through the nanopore via an electrolyte media. Electrodes may be made of any conductive material, for example silver, gold, platinum, copper, titanium dioxide, for example silver coated with silver chloride.

[065] Methods for configuring a nanopore in a solid state, e.g., silicon nitride, membrane, are known. In one approach, a silicon substrate is coated with the membrane material, e.g., silicon nitride, and the overall configuration of the membrane is created using photolithography and wet chemical etching, to provide silicon nitride membranes of the desired size for incorporation into a nanochip, e.g., about 25x25 microns. Initial 0.1 micron diameter holes or cavities are punched in the silicon nitride membrane using a focused ion beam (FIB). Ion beam sculpting can configure the nanopore either by shrinking a larger pore, e.g., by ion beam induced lateral mass transport on the membrane surface, or by removing membrane material by ion beam sputtering layer by layer from the flat side of the membrane containing a cavity from opposing sides, so that when the cavity is ultimately reached, there is a sharp-edged nanopore. The ion beam exposure is extinguished then the ion current transmitted through the pore is appropriate for the desired pore size. See, e.g., Li, J., et al., *Solid-state nanopore for detecting individual biopolymers*, Methods Mol Biol. (2009)544:81-93. Alternatively, the nanopores can be configured using high energy (200-300 keV) electron beam illumination in a TEM. Using semiconductor processing techniques, e-beam lithography, reactive-ion etching of SiO₂ mask layers, and anisotropic KOH etching of Si, pyramidal 20 × 20 nm and larger pores are made in a 40 nm thick membrane. The electron beam in a TEM is used to shrink the larger 20 nm pores to smaller ones. The TEM allows the shrinking process to be observed in real-time. Using a thinner membrane (e.g., < 10 nm thick) nanopores can be drilled with a high energy focused electron beam in a TEM. See, generally, Storm AJ, et al. *Fabrication of solid-state nanopores with single-nanometre precision*. Nature Materials (2003) 2:537–540; Storm AJ, et al. *Translocation of double-stranded DNA through a silicon oxide nanopore*. Phys. Rev. E (2005)71:051903; Heng JB, et al. *Sizing DNA Using a Nanometer-Diameter Pore*. Biophys. J (2004) 87(4):2905–11; the contents of each of which are incorporated herein by reference.

[066] In other embodiments, the nanopores are made using dielectric breakdown, using a relatively high voltage potential across the membrane, wherein the voltage is raised until current is detected, e.g., as described in Kwok, et al., “Nanopore Fabrication by Controlled Dielectric Breakdown,” PLOS ONE (2014) 9(3): e92880, the contents of which are incorporated herein by reference.

[067] Using these techniques, and depending of course on the exact technique used and the thickness and exact composition of the membrane, the overall shape of the nanopore in a solid material such a silicon nitride may roughly resemble two funnels with their apexes coming together at the narrowest point, i.e., the actual nanopore. Such a double cone shape is conducive to steering the polymer through the nanopore and back. Imaging techniques, for example atomic force microscopy (AFM) or transmission electron microscopy (TEM), particularly TEM, can be used to verify and measure the size, location and configuration of the nanomembranes, the FIB holes or cavities, and the final nanopores.

[068] In some embodiments, one end of the polymer, e.g., DNA, is tethered near the nanopore or on the inner wall of the funnel leading to the nanopore. Since the polymer approaches the nanopore initially by diffusion, then is driven by the electrical gradient, the gradient-driven motion is maximized and the diffusive motion minimized, and speed and efficiency thereby enhanced, if one end of the polymer is tethered close to the nanopore. See, e.g. Wanunu M, *Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient*, Nat Nanotechnol. (2010) 5(2):160-5; Gershow M., *Recapturing and trapping single molecules with a solid-state nanopore*. Nat Nanotechnol. (2007) 2(12):775-9; Gershow, M., *Recapturing and Trapping Single Molecules with a Solid State Nanopore*. Nat Nanotechnol. (2007) 2(12): 775–779.

[069] In one embodiment, one end of the polymer, e.g., DNA, is attached to a bead and the polymer is driven through the pore. Attachment to the bead will stop the polymer from moving all the way through the nanopore on the opposite side of the dividing membrane in an adjacent chamber. The current is then turned off, and the polymer, e.g., DNA, attaches to the surface adjacent to the nanopore in a chamber on the other side of the dividing membrane. For example, in one embodiment, one end of ssDNA is covalently attached to a 50nm bead, and the other end is biotinylated. Streptavidin is bound to the area at the desired point of attachment in the chamber on the other side of the dividing membrane. The DNA is pulled through the nanopore

by an electrical potential, and the biotin attaches to the streptavidin. The attachments to the bead and/or the surface adjacent to the nanopore can be either covalent bonds or strong noncovalent bonds (like the biotin-streptavidin bond). The bead is then cut off with an enzyme and flushed away. In some embodiments, the single stranded DNA is cleaved with a restriction enzyme which cleaves single stranded DNA, e.g., as described in K. Nishigaki, *Type II restriction endonucleases cleave single-stranded DNAs in general*. Nucleic Acids Res. (1985) 13(16): 5747–5760, incorporated herein by reference. In other embodiments, a complementary oligonucleotide is provided to make a double-stranded restriction site, which can then be cleaved with the corresponding restriction enzyme.

[070] As the polymer passes through the nanopore, the change in electric potential or current across the nanopore caused by the partial blockage of the nanopore as the polymer passes through can be detected and used to identify the sequence of monomers in the polymer, as the different monomers can be distinguished by their different sizes and electrostatic potentials.

[071] The use of nanochips comprising nanopores in a method of DNA fabrication, as described herein, is not disclosed in the art, but such chips are well known and commercially available for rapid sequencing of DNA. For example, the MinION (Oxford Nanopore Technologies, Oxford, UK) is small and can be attached to a laptop computer. As a single strand of DNA passes through a protein nanopore at 30 bases per second, the MinION measures the electrical current. The DNA strands in the pore disrupts the ionic flow, resulting in changes in current corresponding to the nucleotides in the sequence. Mikheyev, AS, et al. *A first look at the Oxford Nanopore MinION sequencer*, Mol. Ecol. Resour. (2014)14, 1097–1102. While the accuracy of the MinION is poor, requiring repeated resequencing, the speed and accuracy of the sequencing using the nanochips of the present invention can be greatly improved if the DNA being read contains only two easily distinguishable bases, e.g. A and C.

[072] The membrane comprising the nanopores may, in some embodiments, have a trilayer configuration, with a metal surface on either side of an insulating core material, e.g., a silicon nitride membrane. In this embodiment, the metal surfaces are configured, e.g., by lithographic means, to provide a microcircuit with paired electrodes, one at each end of each nanopore, e.g., such that a current flows across the nanopore may be established between the electrodes and through the nanopore via an electrolyte media, which current can draw the polymer through the nanopore and by reversing the polarity, can draw it back. As the polymer passes through the

nanopore, the electrodes can measure the change in electric potential across the nanopore so as to identify the sequence of monomers in the polymer.

[073] In some embodiments, the sequence of the polymer is designed to store data. In some embodiments, the data is stored in a binary code (1's and 0's). In some embodiments each base corresponded to a 1 or 0. In other embodiments, an easily recognized sequence of two or more bases corresponds to a 1 and another easily recognized sequence of two or more bases corresponds to a 0. In other embodiments, the data is can be stored in a ternary, quaternary or other code. In a particular embodiment, the polymer is DNA, for example single stranded DNA, wherein the DNA contains only two base types and does not contain any bases capable of self-hybridizing, e.g., wherein the DNA comprises adenines and guanines, adenines and cytosines, thymidines and guanines, or thymidines and cytosines. In some embodiments, the two bases may be interspersed with one or more additional bases, for example A and C may contain a T to "punctuate" the sequence, e.g., by indicating a break in a coding sequence, at a frequency that does not result in significant self-hybridization. In other embodiments, e.g., where the nucleic acid is double stranded, some or all available bases may be employed.

[074] The nucleotide bases may be natural or may in some embodiments consist of or include nonnatural bases, e.g. as described in Malyshev, D. et al. "A semi-synthetic organism with an expanded genetic alphabet", *Nature* (2014) 509: 385–388, incorporated herein by reference.

[075] In one embodiment, the data is stored by addition of single monomers, e.g., single nucleotides in the case of DNA, to the polymer. In one embodiment, the polymer is DNA and the monomers are adenine (A) and cytosine (C) residues. A and C residues have an advantage because (i) A and C have a large size difference, so differentiation through the nanopore should be facilitated, (ii) A and C do not pair with one another so do not form significant secondary structure which could complicate interpretation of the nanopore signal, and (iii) for the same reason, G's are less preferred as they are know to form guanine tetrads. Nucleotides are added by terminal transferase (or polynucleotide phosphorylase), but the nucleotides are 3'-blocked so that only a single nucleotide is added at a time. The block is removed prior to addition of the next nucleotide.

[076] In some embodiments, the DNA is left in the nanochip. In other embodiments, it is removed, and optionally converted to double stranded DNA and/or optionally converted to crystalline form, e.g. to enhance long term stability. In still other embodiments, DNA can be

amplified and the amplified DNA removed for long term storage, while the original template DNA, for example DNA bound to the wall of a chamber in the nanochip, can be left in the nanochip, where it can be read and/or used as a template to make additional DNA.

[077] In some embodiments, the DNA or other polymer is anchored to a surface proximate to the nanopore during synthesis. For example, in one embodiment, single stranded DNA molecules are each attached at the 5' end to a surface proximate to a nanopore, wherein the current at each nanopore can be independently regulated by electrodes for that nanopore, so that the 3' end of the DNA molecule can be pulled forward through the nanopore from a retaining chamber into a flow chamber containing a flow of 3'-protected dNTPs together with a polymerase or terminal transferase enzyme to add a 3'-protected dNTPs, or retained in the retaining chamber where the nanopore excludes the enzyme, so that the dNTP is not added. See, e.g., depictions at Figures 12-16 and also Figures 18 and 19. In other embodiments, single stranded DNA is built by addition to the 5' end (with the 3' end attached), using topoisomerase, as described more fully below. By controlling whether or not each DNA molecule participates in each cycle, the sequence of each DNA molecule can be precisely controlled, e.g., as follows:

Step	Flow Chamber	Nanopore 1	Nanopore 2
0		Retain	Retain
1	flow 'A'		
2		Forward into flow chamber 'A' gets added	retain
3		Reverse back into resting chamber; oligo is deprotected	retain
4	Flush with buffer	retain	retain
5	flow 'C'	forward 'C' gets added	forward 'C' gets added

6		reverse oligo is deprotected	reverse oligo is deprotected
7		retain	forward 'C' gets added
8		retain	reverse oligo is deprotected
9	Flush with buffer	retain	retain

Flow A = 3'-protected dATP

Flow C = 3'-protected dCTP

Nanopore 1 and Nanopore 2 in this schematic are associated with different DNA strands and the positions of which (in or out of the flow chambers) are separately controllable. The DNA can be deprotected either by a specific enzyme in the retaining chamber, or by changing the flow in the flow chamber to provide deprotection by enzymatic, chemical, light-catalyzed or other means. In one embodiment, the deblocking agent(s) flow between cycles of Flow A and Flow C, e.g., when the flow chamber is being washed with buffer, so that the deblocking agent does not deprotect the nucleotide building blocks. In other embodiments, the deprotecting agent(s) are too bulky to cross to the flow chamber via the nanopores.

[078] The end result in the foregoing example would be that an A and a C were added to the DNA at Nanopore 1, and a C and a C were added to the DNA at Nanopore 2.

[079] In another embodiment, the chamber configuration is similar, but with double stranded DNA anchored to the surface proximate to a nanopore, and oligonucleotide fragments, for example of two or more types, each corresponding to a binary code, are added sequentially, e.g., using site-specific recombinases, i.e., enzymes that spontaneously recognize and cleave at least one strand of a double strand of nucleic acids within a sequence segment known as the site-specific recombination sequence, for example using topoisomerase-charged oligonucleotides as described below.

[080] In certain embodiments, it may be desirable to keep the electrically charged polymer, e.g., DNA, in a condensed state subsequent to synthesis. There are several reasons for this:

- the polymer should be more stable in this form,
- condensing the polymer will keep down crowding and allow use of longer polymers in small volumes,
- orderly condensation can reduce potential that the polymer will form knots or tangles,
- if any of the chambers are interconnected it will help keep the polymer from getting so long that it goes through a different pore than it is supposed to when current is applied,
- condensation will help keep polymer away from the electrodes, where electrochemistry could damage the polymer.

A human cell is about 10 microns but contains 8 billion base pairs of DNA. Stretched out it would be over a meter long. The DNA fits into the cell because it is wound around histone proteins. In certain embodiments, histones or similar proteins provide a similar function in the nanochips of the invention. In some embodiments, the interior surfaces of the nanochips are slightly positively charged so that electrically charged polymer, e.g., DNA tends to stick weakly to them.

[081] In certain embodiments, the charged polymer, e.g., single or double stranded DNA, bound to a surface proximate to a nanopore. This can be accomplished in various ways.

Generally, the polymer is localized to the nanopore by attaching the polymer to a relatively bulky structure (e.g. a bead, a protein, or a DNA origami structure (described below), having a diameter too large to fit through the nanopore, e.g., >10nm, e.g., about 20-50nm), pulling the charged polymer through the nanopore using current, anchoring the end of the polymer distal to the bulky structure to the surface adjacent to the nanopore, and cleaving off the bulky structure.

[082] The step of anchoring the end of the polymer distal to the bulky structure to the surface adjacent to the nanopore, can be accomplished in various ways. In one embodiment, the polymer is a single stranded DNA, and there are pre-attached DNA strands (about 50bp) which are complementary to part of the single stranded DNA, so that the single stranded DNA and the pre-attached DNA strands can join via base pairing. If the pairing is strong enough, it will be sufficient to keep the DNA anchored even while being manipulated. An advantage of this method of attachment is that it allows the DNA to be removed from the nanopore chip if desired for long term storage of the DNA. Alternatively, the strand is attached to the surface covalently, either using conjugation chemistry, e.g., streptavidin-biotin conjugation as described in Example 1 below, or 'click' chemistry (see Kolb, et al. *Angew. Chem. Int. Ed.* (2001)40: 2004-2021,

incorporated herein by reference, and/or using enzymatic attachment, for example by pre-attaching oligos covalently to the distal surface, and then using DNA ligase to connect them.

[083] Once the distal end of the strand is attached to the surface adjacent to the nanopore, the bulky structure is cleaved off, e.g., using an endonuclease which cleaves at a restriction site near the bulky structure.

[084] The bulky structure may be a bead, a bulky molecule, e.g., a protein which is reversibly bound to a DNA strand, or a DNA origami structure. DNA origami involves the use of base pairing to create three dimensional DNA structures. DNA origami techniques are generally described in Bell, et al, Nano Lett. (2012)12: 512–517, incorporated herein by reference. For example, in the current invention, DNA origami can be used to attach the single DNA molecule to a surface adjacent to the nanopore. In one embodiment, the structure is a ‘honey comb cube’, e.g., about 20nm on each side. This prevents this part of the DNA from going through the nanopore (just like in the attached paper). There is a long strand of DNA (single or double stranded) attached to the origami structure. The DNA strand goes through through the nanopore, until the origami cube meets the nanopore and blocks further progress. The current is then turned off and the strand is attached to the surface adjacent to the nanopore.

[085] In another embodiment, the electrically charged polymer, e.g., DNA, with the origami structure is in the middle chamber of a three chamber configuration. The origami will keep the DNA from completely entering the other 2 chambers (or other one chamber in the 2 chamber example). Thus, in this example the polymer doesn’t need to be anchored to the surface. This reduces the risk that the polymer will knot up and avoids the need for the step of binding one end of the polymer to the surface and cleaving off the bulky portion at the other end. The volume of the chamber with the origami should be kept as small as practical so that the polymer stays relatively close to the pore, which will help ensure that it translocates quickly when current is applied. It should be noted that while the middle chamber containing the origami portion of the polymer can’t be interconnected with other middle chambers (or else the different polymers will get mixed up), the other chambers (or sets of chambers in the 3 chamber example) can be interconnected. These other chambers can have larger volumes if desired, as the polymer will necessarily be close to the pore (some of it will be in the pore in fact) when the DNA is moved to that chamber.

[086] In some embodiments, the device comprises three in-line chambers, wherein the addition chambers are contiguous to allow for flow, and have common electrodes, while the 'deprotect' chambers are fluidically isolated except for the flow through the nanopore and have unique electrodes.

[087] In other embodiments, the DNA or other charged polymer is not anchored but can move between synthesis chamber(s) and deprotection chamber(s), under control of electrodes in the chambers, while the polymerase and the deprotecting agents are restricted from movement between chambers because they are too bulky to pass through the nanopores connecting the chambers and/or are anchored to a surface in a chamber. See, e.g., Figures 1-9 and 16-17.

[088] The current needed to move the charged polymer through the nanopore depends on, e.g., the nature of the polymer, the size of the nanopore, the material of the membrane containing the nanopore, and the salt concentrations, and so will be optimized to the particular system as required. In the case of DNA as used in the examples herein, examples of voltage and current would be, e.g., 50-500mV, typically 100-200mV, and 1-10nA, e.g., about 4nA, with salt concentrations on the order of 100mM to 1M.

[089] The movement of charged polymer, e.g., DNA, through the nanopore is normally very rapid, e.g., 1 to 5 μ s per base, so on the order of one million bases per second (1 MHz, if we adopt the nomenclature of frequency), which presents challenges for getting an accurate reading distinct from the noise in the system. Using current methods, either (i) a nucleotide needed to be repeated in a sequence, e.g., ca. 100 times successively, in order to produce a measurable characteristic change, or (ii) using protein pores, such as Alpha hemolysin (α HL) or Mycobacterium smegmatis porin A (MspA), which provide a relatively long pore with potential for multiple reads as the base moves through the polymer, and in some cases, can be adapted to provide a controlled feed of DNA through the pore one base at a time, in some cases using an exonuclease to cleave each base as it passes through. Various approaches are possible, e.g.,

- slowing down the speed of the polymer, from ca. 1MHz to ca. 100-200Hz, for example using a medium comprising an electrorheological fluid in which becomes more viscous when a voltage is applied, thereby slowing down the speed of the polymer through the nanopore, or a plasmonic fluid system, wherein the viscosity of the medium can be controlled by light; or a molecular motor or ratchet;

- providing a sequence in the polymer, e.g., in single stranded DNA, which will form a bulky secondary structure, e.g., a “hairpin”, “hammerhead”, or “dumbbell” configuration, which will have to be linearized in order to fit through the nanopore, thereby making the information less dense and providing a signal having a longer duration;
- providing many reads of the same sequence, e.g., by using rapidly alternating current, allowing for many reads of the same sequence frame, and combined with brief bursts of direct current to pull the molecule to the next sequence frame, by reading the entire sequence multiple times, or by reading multiple identical sequences in parallel, in each case collating the reads to provide a consensus read that amplifies the signal ;
- measuring an impedance change in a high frequency signal induced by a change in capacitance as monomers (e.g., nucleotides) pass through the nanopore, rather than measuring changes in current flow or resistance directly;
- enhancing the differences in current, resistance or capacitance between different bases, e.g., by using non-natural bases which have a greater difference in size or are otherwise modified to give different signals, or by forming larger secondary structures within the DNA, such as a “hairpin”, “hammerhead”, or “dumbbell” configuration, which provide an enhanced signal because of their larger size;
- using an optical reading system, for example using an integrated optical antenna adjacent to the nanopore, which acts as an optical transducer (or optical signal enhancer) to complement or replace standard ion current measurements, e.g., as described in Nam, et al., “*Graphene Nanopore with a Self-Integrated Optical Antenna*”, Nano Lett. (2014)14: 5584–5589, the contents of which are incorporated herein by reference. In some embodiments, the monomers, e.g. DNA nucleotides, are labeled with fluorescent dyes so that each different monomer fluoresces at a signature intensity as it passes through the junction of the nanopore and its optical antenna. In some embodiments, a solid-state nanopore strips off fluorescent labels, leading to a series of detectable photon bursts, as the polymer passes through the nanopore at high speed, e.g. as described in McNally et al., “*Optical recognition of converted DNA nucleotides for single molecule DNA sequencing using nanopore arrays*”, Nano Lett. (2010)10(6): 2237–2244, and Meller A., “*Towards Optical DNA Sequencing Using Nanopore Arrays*”, J. Biomol. Tech. (2011) 22(Suppl): S8–S9, the contents of each of which are incorporated herein by reference.

[090] In one embodiment, the charged polymer is a nucleic acid, e.g., single stranded DNA, wherein the sequences provide a secondary structure. Bell, et al., *Nat Nanotechnol.*(2016)11(7):645-51, incorporated herein by reference, describes using a relatively short sequence of dumbbell configurations detectible in a solid state nanopore format, to label antigens in an immunoassay. The nanopores used in Bell, et al. were relatively large, so the entire dumbbell structure could pass through the pore, but using nanopores smaller than the diameter of the dumbbell configuration, the DNA will “unzip” and become linearized. More complex configurations can be used, e.g. wherein each bit corresponds to a sequence similar to a tRNA (see, e.g., Henley, et al. *Nano Lett.* (2016)16: 138–144, incorporated herein by reference). Thus the invention provides charged polymers, e.g. single stranded DNA, having at least two types secondary structure, wherein the secondary structure encode data (e.g. binary data, wherein one secondary structure type is a 1 and a second is a 0). In other embodiments, secondary structures are used to slow down the passage of the DNA through the nanopore or to provide breaks in the sequence, to facilitate reading of the sequence.

[091] In another embodiment, the invention utilizes a DNA molecule comprising a series of at least two different DNA motifs, wherein each motif specifically binds to a particular ligand, for example a gene regulatory protein for double stranded DNA or a tRNA for single stranded DNA, wherein the at least two different DNA motifs encode information, e.g. in a binary code, wherein one motif is a 1 and a second is a 0, e.g., wherein the ligand enhances the signal difference (e.g. change in current or capacitance) across the nanopore as the DNA passes through the nanopore.

[092] As discussed above, when different monomers pass through the nanopore, they affect the current flow through the nanopore, primarily by physically blocking the nanopore and changing the conductance across the nanopore. In existing nanopore systems, this change in current is measured directly. The problem with current reading systems is that there is considerable noise in the system, and in the case of DNA, for example, when measuring current fluctuations as different nucleotide units pass through the nanopore, a relatively long integration time, on the order of one hundredth of a second, is needed to accurately detect differences between different monomers, e.g., between different bases. Recently, it has been shown that changes in impedance and capacitance can be useful to study cells and biological systems, despite the potential for complex interactions with salts and biological molecules. For example, Laborde, et al. *Nat*

Nano. (2015)10(9):791-5 (incorporated herein by reference), demonstrates that high-frequency impedance spectroscopy can be used to detect small changes in capacitance under physiological salt conditions and image microparticles and living cells beyond the Debye limit.

[093] In one embodiment of the invention, therefore, we measure capacitive variance rather than measuring current variance directly, for example, wherein the sequence of the charged polymer is identified by measuring the phase change in a radiofrequency signal induced by change in capacitance as the monomers (e.g., nucleotides) pass through the nanopore.

[094] Simply stated, capacitance exists in any circuit where there is a gap between one electrical conductor and another. While current varies directly with capacitance, it does not vary simultaneously with capacitance. For example, if we were to plot the current and voltage over time in a capacitive circuit with an alternating electrical current, we would see that while both current and voltage each form a sine wave, the waves are out of phase. When there is a change in current, there is a change in capacitance, which is reflected in a change in the phase of the signal. A radiofrequency alternating current provides a signal with fixed frequency and amplitude, while the phase of the signal will vary with the capacitance of the circuit. In our system, we use a pulsating direct current rather than an alternating current (i.e., the voltage alternates between two values, but the voltage does not cross the “zero” line, such that polarity is maintained and one electrode remains positive and the other negative), so that the charged polymer can be drawn through the nanopore (towards the positive electrode in the case of DNA). When there is nothing in the nanopore, the capacitance has one value, which changes as the different monomers of the polymer pass through the nanopore. Suitable frequency ranges are in the radiofrequency range, e.g. 1 MHz to 1GHz, e.g. 50-200MHz, for example about 100MHz, e.g. below higher microwave frequencies that could cause significant dielectric heating of the medium. To reduce the potential for interference, different frequencies can be applied at different nanopores so that multiple nanopores can be measured simultaneously with a single radiofrequency input line.

[095] Measuring impedance changes (due to, e.g. changes in capacitance) at high frequencies increases the signal to noise available within a certain time span, as it reduces the effects of 1/f noise, or ‘pink’ noise that is inherent in electronic measurement circuits. Using a high frequency signal enhances the signal-to-noise ratio, as many measurements are made within a given time span, providing a more stable signal which is readily distinguished from impedance changes due to environmental or device variation and fluctuation.

[096] Applying these principles to the instant invention, the invention provides in one embodiment, a method of measuring an impedance change in a high frequency signal induced by a change in capacitance as monomers (e.g., nucleotides) pass through a nanopore, for example, a method of reading a monomer sequence of a charged polymer comprising at least two different types of monomers, for example a DNA molecule, comprising applying a radiofrequency pulsating direct current, e.g. at a frequency of 1 MHz to 1GHz, e.g. 50-200MHz, for example about 100MHz, across a nanopore, wherein the pulsating direct current draws the charged polymer through the nanopore and the monomer sequence is read by measuring the capacitive variance across the nanopore as the charged polymer goes through the nanopore.

[097] In some embodiments, the invention provides a nanochip for sequencing an electrically charged polymer, e.g., DNA, comprising at least two distinct monomers, the nanochip comprising at least a first and second reaction chambers, each comprising electrolytic medium, and separated by a membrane comprising one or more nanopores, wherein a pair of electrodes (for example in the form of opposing plates), connected in circuit, is disposed on either side of the membrane comprising one or more nanopores, the electrodes being separated by a distance of 1-30 microns, e.g., about 10 microns, such that the gap between the electrodes has a capacitance when a radiofrequency pulsating direct current, e.g. 1 MHz to 1GHz, is applied to the electrodes so as to draw the electrically charged polymer through the nanopore, e.g., from one chamber to the next, and such that the phase of the pulsating direct radiofrequency current changes with changes in capacitance as the electrically charged polymer passes through the nanopore, thereby allowing detection of the monomer sequence of the electrically charged polymer. In certain embodiments, the nanochip comprises multiple sets of reaction chambers wherein the reaction chambers within a set are separated by membrane having one or more nanopores, and the sets of reaction chambers are separated by a screening layer to minimize electrical interference between the sets of reaction chambers and/or to separate multiple linear polymers and allow them to be sequenced in parallel.

[098] For example, in one embodiment the electrodes form the top and bottom plates of a capacitor embedded in a resonant circuit, and the change in capacitance is measured as the DNA passes through the pore between the plates.

[099] In certain embodiments, the nanochip further comprises reagents for synthesizing the polymer, e.g. DNA, e.g., according to any of Nanochip 1, et seq., below.

[0100] In one embodiment, therefore the invention provides a method (Method 1) for synthesizing a charged polymer [e.g., a nucleic acid (e.g., DNA or RNA)] comprising at least two distinct monomers in a nanochip, the nanochip comprising

one or more addition chambers containing reagents for addition of one or more monomers [e.g. nucleotides] or oligomers [e.g., oligonucleotides] to the charged polymer in a buffer solution in terminal protected form, such that only a single monomer or oligomer can be added in one reaction cycle; and

one or more reserve chambers containing buffer solution but not all reagents necessary for addition of the one or more monomers or oligomers,

wherein the chambers are separated by one or more membranes comprising one or more nanopores and

wherein the charged polymer can pass through the nanopore but the least one of the reagents for addition of one or more monomers or oligomers cannot,

the method comprising

- a) moving the first end of a charged polymer having a first end and a second end, by electrical attraction into an addition chamber, whereby monomers or oligomers are added to said first end in blocked form,
- b) moving the first end of the charged polymer with the added monomer or oligomer in blocked form into a reserve chamber,
- c) deblocking the added monomer or oligomer, and
- d) repeating steps a-c, wherein the monomers or oligomers added in step a) are the same or different, until the desired polymer sequence is obtained.

[0101] For example, the invention provides

1.1. Method 1, wherein the polymer is nucleic acid, e.g., wherein the polymer is DNA or RNA, e.g., wherein it is DNA, e.g dsDNA or ssDNA.

1.2. Any foregoing method wherein the second end of the polymer, e.g. the nucleic acid, is either protected or bound to a substrate adjacent to the nanopore.

1.3. Any foregoing method wherein the electrical attraction is provided by applying an electric potential between the electrodes in each chamber, wherein the polarity and

current flow between the electrodes can be controlled, e.g., such that the nucleic acid is attracted to a positive electrode.

1.4. Any foregoing method wherein the polymer is a nucleic acid and

(i) the said first end of the nucleic acid is the 3'-end, the addition of nucleotides is in the 5' to 3' direction and is catalyzed by a polymerase, e.g., wherein the polymerase is hindered (e.g. due to its size or due to being tethered to a substrate in the first chamber) from passing through the nanopore, the nucleotides are 3'-protected when added, and following addition of the 3'-protected nucleotide to the 3'-end of the nucleic acid, the 3'-protecting group on the nucleic acid is removed, e.g., in the reserve chamber; or

(ii) the said first end of the nucleic acid is the 5' end, the addition of nucleotides is in the 3' to 5' direction, the nucleotides are 5'-protected when added, and following addition of the 5'-protected nucleotide to the 5'-end of the nucleic acid, the 5'protecting group is removed, e.g., in the second chamber; (for example wherein the phosphate on the 5'-protected nucleotide is a nucleoside phosphoramidite coupled via the 5'-protecting group to a bulky group which cannot pass through the nanopore, so that following coupling to the nucleic acid, the unreacted nucleotides are flushed away, the bulky 5'-protecting group is cleaved from the nucleic acid, and flushed away, and the 5'-end of the nucleic acid can be moved into the reserve chamber);

wherein the addition of nucleotides to the nucleic acid is controlled by movement of the first end of the nucleic acid into and out of the one or more addition chambers, and the cycle is continued until the desired sequence is obtained.

1.5. Any foregoing method wherein the sequence of monomers or oligomers in the polymer [e.g., the sequence of nucleotides in the nucleic acid] thus synthesized corresponds to a binary code.

1.6. Any foregoing method wherein the polymer thus synthesized is single stranded DNA.

1.7. Any foregoing method wherein the sequence of the polymer [e.g. the nucleic acid] is checked during the process or synthesis by sequencing the monomers or oligomers

[e.g., nucleotide bases] as they pass through the nanopore to identify errors in sequencing.

1.8. Any foregoing method wherein the polymer thus synthesized is single stranded DNA, wherein at least 95%, e.g. at least 99%, e.g., substantially all of the bases in the sequence are selected from two bases that do not hybridize with other bases in the strand, e.g. bases selected from adenine and cytosine.

1.9. Any foregoing method wherein a multiplicity of polymers [e.g. oligonucleotides] are synthesized independently in parallel, such that polymers [oligonucleotides] having different sequences are obtained by separately controlling whether they are present in one or more addition chambers or one or more reserve chambers.

1.10. Any foregoing method wherein there are at least two addition chambers that contain reagents suitable for adding different monomers or oligomers, e.g. different nucleotides, e.g., wherein there are one or more addition chambers containing reagents suitable for adding a first monomer or oligomer and one or more addition chambers containing reagents suitable for adding a second different monomer or oligomer, for example wherein there are one or more addition chambers containing reagents suitable for adding adenine nucleotides and one or more addition chambers containing reagents suitable for adding cytosine nucleotides.

1.11. Any foregoing method wherein at least one addition chamber is a flow chamber, providing a flow cycle comprising (i) providing to the flow chamber reagents suitable for adding a first monomer or oligomer, (ii) flushing, (iii) providing to the flow chamber reagents suitable for adding a second different monomer or oligomer, and (iv) flushing, and repeating the cycle, until the synthesis is complete, wherein the sequence of monomers or oligomers in the polymer is controlled by introducing or excluding the first end of the polymer from the flow chamber during step (i) or (iii) in each cycle;

1.12. Any foregoing method wherein the polymer is DNA and at least one addition chamber is a flow chamber, providing a flow cycle comprising (i) providing to the flow chamber reagents suitable for adding a first type of nucleotide, (ii) flushing, (iii) providing to the flow chamber reagents suitable for adding a second type of nucleotide, and (iv) flushing, and repeating the cycle until the synthesis is complete, wherein the

sequence is controlled by controlling the presence or absence of the first end of the DNA (e.g. the 3'-end) in the flow chamber.

1.13. Any foregoing method wherein the polymer is DNA and at least one addition chamber is a flow chamber, providing a flow cycle comprising (i) providing to the flow chamber reagents suitable for adding a first type of nucleotide, (ii) flushing, (iii) providing to the flow chamber reagents suitable for adding a second type of nucleotide, and (iv) flushing, (i) providing to the flow chamber reagents suitable for adding a third type of nucleotide, (ii) flushing, (iii) providing to the flow chamber reagents suitable for adding a fourth type of nucleotide, and (iv) flushing, and repeating the cycle until the synthesis is complete, wherein the sequence is controlled by controlling the presence or absence of the first end of the DNA (e.g. the 3'-end) in the flow chamber when reagents suitable for adding the different types of nucleotides are present.

1.14. Any foregoing method wherein the polymer is DNA and the nanochip comprises two addition chambers which are flow chambers, (a) the first flow chamber providing a flow cycle comprising (i) providing to the first flow chamber reagents suitable for adding a first type of nucleotide, (ii) flushing, (iii) providing to the first flow chamber reagents suitable for adding a second different type of nucleotide, and (iv) flushing, and repeating the cycle until the synthesis is complete, and (b) the second flow chamber providing a flow cycle comprising (i) providing to the second flow chamber reagents suitable for adding a third type of nucleotide, (ii) flushing, (iii) providing to the second flow chamber reagents suitable for adding a fourth different type of nucleotide, and (iv) flushing, and repeating the cycle until the synthesis is complete, wherein the nucleotides are selected from dATP, dTTP, dCTP, and dGTP and wherein the sequence is controlled by directing the first end of the DNA (e.g. the 3'-end) into the flow chamber where the next desired nucleotide is provided.

1.15. Any foregoing method wherein the polymer is DNA and the nanopore chip comprises one or more addition chambers for adding dATP, one or more addition chambers for adding dTTP, one or more addition chambers for adding dCTP, and one or more addition chambers for adding dGTP.

1.16. Any foregoing method wherein the polymers [e.g. nucleic acids] synthesized are each bound via their second end to a surface proximate to a nanopore.

- 1.17. Any foregoing method wherein the sequence of the polymer [e.g. nucleic acid] is determined following each cycle by detecting the change in electric potential, current, resistance, capacitance, and/or impedance as the polymer passes through the nanopore.
- 1.18. Any foregoing method wherein the polymer is a nucleic acid and synthesis of the nucleic acid takes place in a buffer solution, e.g., a solution comprising a buffer for pH 7-8.5, e.g. ca. pH 8, e.g. a buffer comprising tris(hydroxymethyl)aminomethane (Tris), a suitable acid, and optionally a chelator, e.g., ethylenediaminetetraacetic acid (EDTA), for example TAE buffer containing a mixture of Tris base, acetic acid and EDTA or TBE buffer comprising a mixture of Tris base, boric acid and EDTA; for example a solution comprising 10mM Tris pH 8, 1 mM EDTA, 150 mM KCl, or for example, 50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, pH 7.9 @ 25°C.
- 1.19. Any foregoing method wherein the polymer is single stranded DNA further comprising converting the synthesized single stranded DNA into double stranded DNA.
- 1.20. Any foregoing method further comprising removing the polymer [e.g. the nucleic acid] from the nanochip after the polymer synthesis is complete.
- 1.21. Any foregoing method wherein the polymer is a nucleic acid, further comprising amplifying and retrieving copies of synthesized nucleic acid using an appropriate primer and a polymerase (e.g. Phi29).
- 1.22. Any foregoing method wherein the polymer is a nucleic acid, further comprising cleaving the synthesized nucleic acid with a restriction enzyme and removing the the nucleic acid from the nanochip.
- 1.23. Any foregoing method wherein the polymer is a nucleic acid, further comprising amplifying the nucleic acid thus synthesized.
- 1.24. Any foregoing method further comprising removing the polymer [e.g., the nucleic acid] from the nanochip and crystallizing the polymer.
- 1.25. Any foregoing method wherein the polymer is a nucleic acid, further comprising stabilizing the nucleic acid, e.g., by drying a solution comprising the nucleic acid together with one or more of a buffer (e.g., a borate buffer), an antioxidant, a humectant, e.g. a polyol, and optionally a chelator, for example as described in US 8283165 B2, incorporated herein by reference; or by forming a matrix between the nucleic acid and a polymer, such as poly(ethylene glycol)–poly(l-lysine) (PEG–PLL) AB type block

copolymer; or by addition of a complementary nucleic acid strand or a protein that binds the DNA.

1.26. Any foregoing method comprising:

- (i) reacting a nucleic acid with a 3'-protected nucleotide in an addition chamber, in the presence of a polymerase which catalyzes the addition of the 3'-protected nucleotide to the 3' end of the nucleic acid, ;
- (ii) drawing at least the 3' end of the 3'-protected nucleic acid thus obtained out of the addition chamber, through the at least one nanopore, into a reserve chamber, wherein the polymerase is hindered (e.g. due to its size or due to being tethered to a substrate in the first chamber) from passing through the nanopore;
- (iii) deprotecting the 3'-protected nucleic acid, e.g., chemically or enzymatically; and
- (iv) if it is desired that an additional 3'-protected dNTP be added to the oligonucleotide, drawing the 3' end of the oligonucleotide into the same or different addition chamber, so that steps (i) – (iii) are repeated, or if it is not so desired, allowing the 3' end of the nucleic acid to remain in the reserve chamber until a further cycle wherein the desired 3'-protected dNTP is provided to the addition chamber; and
- (v) repeating the cycle of steps (i) – (iv) until the desired nucleic acid sequence is obtained.

1.27. Any foregoing method wherein the polymer is nucleic acid single-stranded DNA (ssDNA) and the one or more nanopores have a diameter allowing ssDNA to pass but not double stranded DNA (dsDNA), e.g., a diameter of about 2nm.

1.28. Any foregoing method wherein the monomer is a 3'-protected nucleotide, e.g., deoxynucleotide triphosphate (dNTP), e.g. selected from deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), deoxycytidine triphosphate (dCTP), deoxythymidine triphosphate (dTTP), for example dATP or dCTP.

1.29. Any foregoing method wherein the polymer is a nucleic acid and the addition of the nucleotide to the nucleic acid is catalyzed by a polymerase, e.g., a template independent polymerase, e.g., terminal deoxynucleotidyl transferase (TdT), or

polynucleotide phosphorylase, e.g., wherein the polymerase catalyzes the incorporation of a deoxynucleotide at the 3'-hydroxyl terminus of DNA.

1.30. Any foregoing method wherein the membrane contains a multiplicity of nanopores and a multiplicity of polymers each bound to a surface proximate to a nanopore, e.g., a multiplicity of nucleic acids each bound via their 5' end to a surface proximate to a nanopore.

1.31. Any foregoing method wherein a multiplicity of polymers each bound to a surface proximate to a nanopore, e.g., a multiplicity of nucleic acids each bound at the 5' end to a surface proximate to a nanopore, are synthesized independently, wherein each nanopore has an associated pair of electrodes, wherein one electrode in the pair is located proximate to one end of the nanopore and the other electrode located proximate to the other end of the nanopore, such that each polymer can be independently moved between the first and second chamber by current provided by the pair of electrodes.

1.32. Any foregoing method wherein the polymer is a 3'-protected nucleic acid bound at the 5' end to a surface proximate to a nanopore and the 3' end of the 3'-protected nucleic acid is drawn through the nanopore by using an electrical force, e.g., by using an electrical force applied from an electrode in an adjacent chamber.

1.33. Method 1.20 wherein the new 3'-protected dNTP is the same or different from the first 3'-protected dNTP.

1.34. Method 1.20 wherein the 3'-protected dNTP used in step (i) of the cycle alternates with each cycle between 3'-protected dATP and 3'-protected dCTP.

1.35. Any foregoing method wherein the polymer is a nucleic acid and deprotection of the nucleic acid is carried out by an enzyme that removes a 3'-protecting group on ssDNA but not on a 3'-protected dNTP.

[0101] For example, the invention provides a method for synthesizing a nucleic acid in a nanochip, comprising at least a first chamber and a second chamber separated by a membrane comprising at least one nanopore, the synthesis being carried out in a buffer solution by a cycle of nucleotide addition to a first end of a nucleic acid having a first end and a second end, wherein the first end of the nucleic acid is moved by electrical attraction between one or more addition chambers (which contains reagents capable of adding nucleotides) and one or more reserve chambers (which do not contain reagents necessary to add nucleotides), the chambers being

separated by one or more membranes each comprising one or more nanopores, wherein the nanopore is large enough to permit passage of the nucleic acid but is too small to permit passage of at least one reagent essential for adding a nucleotide, e.g., wherein the method corresponds to any of Method 1, et seq.

[0102] In certain embodiments, the sequence of the polymer corresponds to a binary code, for example where the polymer is a nucleic acid and the sequence corresponds to a binary code, where each bit (0 or 1) is represented by a base, e.g. A or C.

[0103] In certain embodiments, the polymer is DNA

[0104] In certain other embodiments, each bit is represented by a short sequence of monomers rather than by a single monomer. For example, in one such embodiment, blocks of DNA are synthesized, where each block generates a unique signal via the nanopore and corresponds to a zero or a one. This embodiment has certain advantages in that single nucleotides are more difficult to detect in nanopores, especially solid-state nanopores, so using blocks is less prone to reading errors, although the information density in the polymer is correspondingly reduced.

[0105] For example, blocks of (double stranded) nucleotides can be added, using site-specific recombinases, i.e., enzymes that spontaneously recognize and cleave at least one strand of a double strand of nucleic acids within a sequence segment known as the site-specific recombination sequence. In one such embodiment, the site specific recombinase is a topoisomerase used to ligate a topo-conjugated dsDNA oligonucleotide block to the sequence. These oligonucleotides themselves will not have a structure compatible with further ligation until they are cleaved with a restriction enzyme. Vaccinia virus topoisomerase I specifically recognises DNA sequence 5'-(C/T)CCTT-3'. The topoisomerase binds to double-stranded DNA and cleaves it at the 5'-(C/T)CCTT-3' cleavage site. Note that the cleavage is not complete, as the topoisomerase only cleaves the DNA on one strand (although having a nearby nick on the other strand does cause a double-strand break of sorts), and when it cleaves, the topoisomerase attaches covalently to the 3' phosphate of the 3' nucleotide. The enzyme then remains covalently bound to the 3' end of the DNA, and can either religate the covalently held strand at the same bond as originally cleaved (as occurs during DNA relaxation), or it can religate to a heterologous acceptor DNA having compatible overhangs, creating a recombinant molecule. In this embodiment, we create dsDNA donor oligonucleotides (e.g., comprising one of at least two different sequences, one for '0' and the other for '1') flanked by a topoisomerase recombination

site and a restriction site that generates a topoisomerase ligation site. The cassettes are Topo-charged; that is, they are covalently bound to a topoisomerase, which will bind them to a topoisomerase ligation site on the receiver oligonucleotide. When the growing DNA chain of the receiver is cleaved with a restriction enzyme it becomes capable of ligation to a Topo-charged cassette. So, one just needs to cycle the growing DNA from restriction enzyme to Topo-charged cassette successively, with each cycle adding another donor oligonucleotide. A related approach has been described for cloning, see, e.g., Shuman S., *Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase*. J Biol Chem. (1994); 269(51):32678-84, the contents of which are incorporated by reference.

[0106] Single bases can be added using a similar strategy. In the presence of a suitable single stranded 'deprotected' 'acceptor' DNA, the topo-charged DNA is enzymatically and covalently ligated ('added') to the acceptor by the topoisomerase, which in the process becomes removed from the DNA. A type IIS restriction enzyme can then cleave all of the added DNA with the exception of a single base (the base which is being 'added'). This process of deprotect-add can be repeated to add additional bases (bits). As demonstrated in the examples herein, it is feasible to use a Topo / TypeIIS restriction enzyme combination to add a single nucleotide to the 5' end of a target single stranded DNA. The use of a TypeIIS restriction enzyme enables cleavage of DNA in a location different from that of the recognition sequence (other TypeIIS restriction enzymes can be found at <https://www.neb.com/tools-and-resources/selectioncharts/type-iis-restriction-enzymes>). The use of inosines (which act as 'universal bases' and pair with any other base) in this system allows this reaction to occur without any specific sequence requirements in the target DNA. The identity of the nucleotide added to the single strand target DNA is the 3' nucleotide to which vaccinia topoisomerase conjugates via the 3' phosphate. Since the recognition sequence of vaccinia topoisomerase is (C/T)CCTT, we have used this system to add a 'T' to the target DNA. There is a related topoisomerase, SVF, that can use the recognition sequence CCCTG (<https://www.ncbi.nlm.nih.gov/pubmed/8661446>). Thus SVF can be used to add a 'G' instead of a 'T'. Paired with vaccinia topo, binary data can be encoded in T's and G's.

[0107] In another approach to single base addition, a 5'phosphate provides a blocking group to provide single base addition in the 3' to 5' direction. The charging reaction charges the topoisomerase with a single T (or G, or other nucleotide as desired), having a 5' phosphate group. When the charged topoisomerase 'sees' a free 5' unblocked (unphosphorylated) single

stranded DNA chain it will add the T to that chain, providing a DNA with a T added to the 5'. This addition is facilitated by the presence of an adapter DNA having sequences to which the topoisomerase and the single stranded acceptor DNA can bind. (Note that the adapter DNA is catalytic – it can be reused as a template in repeated reactions.) The added nucleotide has a 5' phosphate on it, so it won't be a substrate for further addition until it is exposed to a phosphatase, which removes the 5' phosphate. The process is repeated, using vaccinia topoisomerase to add a single "T" to the 5' end of a target single stranded DNA and SVF topoisomerase to add a single 'G', thus allowing construction of a sequence encoding binary information with T and G. Other topoisomerases can be used to add A's or C's, although this reaction is less efficient.

[0108] One advantage of using a topoisomerase-mediated strategy is that the monomer is covalently attached to the topoisomerase, and therefore cannot "escape" to interfere with other reactions. When polymerase is used, the monomers can diffuse so the polymerases and/or the deblocking agents should be specific (e.g. selective for A vs C, for example) or alternatively, the monomers are provided by a flow so they don't have a chance to mix.

[0109] In one aspect, the invention provides a topoisomerase charged with a single nucleotide, i.e., a topoisomerase conjugated to a single nucleotide, e.g., wherein the topoisomerase is conjugated via the 3'-phosphate of the nucleotide, and the nucleotide is protected, e.g., phosphorylated, at the 5'-position.

[0110] In another aspect the invention provides a method (Method A) of synthesizing a DNA molecule using topoisomerase-mediated ligation, by adding single nucleotides or oligomers to a DNA strand in the 3' to 5' direction, comprising (i) reacting a DNA molecule with a topoisomerase charged with the desired nucleotide or oligomer wherein the nucleotide or oligomer is blocked from further addition at the 5' end, then (ii) deblocking the 5' end of the DNA thus formed, and repeating steps (i) and (ii) until the desired nucleotide sequence is obtained, e.g.,

- A1.1. Method A which is a method of synthesizing a DNA molecule by adding single nucleotides in the 3' to 5' direction comprising (i) reacting a DNA molecule with a topoisomerase charged with the desired nucleotide in 5'protected form, e.g., 5'phosphorylated form, such that the desired nucleotide in 5'protected form is added to the 5' end of the DNA, then (ii) deprotecting the 5' end of the DNA thus formed through

the use of a phosphatase enzyme, and repeating steps (i) and (ii) until the desired nucleotide sequence is obtained; or

- A1.2. Method A which is a method of synthesizing a DNA molecule by adding oligomers in the 3' to 5' direction comprising (i) reacting a DNA molecule with a topoisomerase charged with the desired oligomer, thereby ligating the oligomer to the DNA molecule, then (ii) using a restriction enzyme to provide a 5' site for a topoisomerase-mediated ligation for another oligomer, and repeating steps (i) and (ii) until the desired oligomer sequence is obtained.
- A1.3. Any foregoing method comprising providing ligase and ATP to seal nicks in the DNA [NB: the topoisomerase ligation only ligates one strand].
- A1.4. Any foregoing method wherein the topoisomerase-charged donor oligonucleotide comprises a 5' overhang on the strand complementary to the strand bearing the topoisomerase, comprising a polyinosine sequence [NB: inosines act as 'universal bases' and pair with any other base].
- A1.5. Any foregoing method wherein the restriction enzyme is a type IIS restriction enzyme which can cleave all of the added DNA with the exception of a single base (the base which is being 'added').
- A1.6. Any foregoing method wherein the topoisomerase is selected from vaccinia topoisomerase and SVF topoisomerase I.
- A1.7. Any foregoing method wherein vaccinia topoisomerase (which recognizes (C/T)CCTT) is used to add dTTP nucleotides and SVF topoisomerase I (which recognizes CCCTG) is used to add dGTP nucleotides, e.g., to provide binary code
- A1.8. Any foregoing method wherein the DNA is double stranded and the reserve chamber further comprises a ligase and ATP, to repair the DNA strand not joined by the topoisomerase .
- A1.9. Any foregoing method comprising use of a topoisomerase inhibitor to suppress binding and activity of free topoisomerase to the DNA oligomer, e.g., wherein the inhibitors is selected from novobiocin and coumermycin.
- A1.10. Any foregoing method wherein the DNA strand thus provided has a sequence comprising thymidine (T) nucleosides and deoxyguanosine (G) nucleosides.

- A1.11. Any foregoing method wherein the topoisomerase adds a single base, but the restriction enzyme cleaves at a position which is one nucleotide in the 5' direction from the base added by topoisomerase.
- A1.12. Any foregoing method wherein the DNA strand thus provided has a sequence comprising a sequence of 'TT' and 'TG' dinucleotides.
- A1.13. Any foregoing method wherein the DNA is single stranded,
- A1.14. Any foregoing method wherein the DNA double stranded.
- A1.15. Any foregoing method wherein the DNA is on a substrate or magnetic bead, where it can be selectively exposed to or removed from the reagents as required to provide the desired sequence.
- A1.16. Any foregoing method wherein some or all of the reagents for adding or deblocking the DNA are supplied by flow and removed by flushing.
- A1.17. Any foregoing method wherein the attachment of the single nucleotides or oligomers to a single-stranded DNA is facilitated by the presence of an adapter DNA having sequences to which the topoisomerase and the single stranded acceptor DNA can bind.
- A1.18. Any foregoing method carried out in a system where a nanopore separates a chamber comprising the topoisomerase from a chamber comprising the phosphatase or restriction enzyme, wherein the nanopore allows movement of the DNA by electrical attraction, but not the enzymes, e.g. as described in any of Method 2, et seq.

[0111] One possible concern is poly-G sequences may form G-quartet secondary structures. By moving the restriction enzyme back one base (to the 5' of the topo sequence) and following a similar Topo/IIS strategy a 'TT' or 'TG' can be added, each of which can represent a different bit. While this would require 2 bases to encode a bit, it has the advantage of avoiding poly-G sequences. In other embodiments, other bases in the 3' end of the topo recognition sequence - although less efficient than (C/T)CCTT, can allow conjugation using poxvirus topoisomerase with (C/T)CCTA, (C/T)CCTC and (C/T)CCTG (<https://www.ncbi.nlm.nih.gov/pubmed/17462694>). Protein engineering/selection techniques can be used to improve the efficiency of these reactions as well, and similar approaches can be used to add non-canonical bases.

[0112] In certain embodiments, the method of synthesizing DNA by this method includes treating the DNA with a ligase and ATP. The topoisomerase only joins together one side of the DNA (the other is essentially nicked). The ligase would repair the nick and ensure that the topoisomerase itself doesn't recut the reaction product and cleave it.

[0113] In certain embodiments, the method comprises using a topoisomerase inhibitor to suppress binding and activity of free topoisomerase to the DNA oligomer. Suitable inhibitors include novobiocin and coumermycin. Note that complete inhibition is not desirable, as a low level of topoisomerase activity can help 'relax' coiled DNA, which is useful especially when synthesizing long DNA chains.

[0114] Thus, in another embodiment, the disclosure provides a method (Method 2) for synthesizing DNA in a nanochip, comprising one or more addition chambers containing a topoisomerase-charged oligonucleotide (i.e., oligonucleotide bound at the 3' end to a topoisomerase), and one or more reserve chambers comprising a restriction enzyme or deblocker, e.g., phosphatase, said chambers also containing compatible buffer solution and being separated by a membrane comprising at least one nanopore, wherein the topoisomerase and the restriction enzyme are prevented from passing through the nanopore (e.g. because they are too large and/or because they are tethered to a substrate in the first and second chambers respectively), the synthesis being carried out by a cycle of adding single nucleotides or short oligonucleotide blocks to a first end of a nucleic acid having a first end and a second end, wherein the first end of the nucleic acid is moved by electrical attraction between an addition chambers and a reserve chamber, for example in one embodiment as follows:

- (i) moving the 5' end of a receiver DNA (e.g., a double-stranded DNA) into a first addition chamber, by means of an electrical force,
- (ii) providing in the first addition chamber a topoisomerase-charged donor oligonucleotide, wherein the donor oligonucleotide comprises a topoisomerase binding site, an informational sequence (e.g., selected from at least two different nucleotides or sequences, e.g., wherein one sequence corresponds to '0' and the other to '1' in a binary code), and a restriction site which when cleaved by a restriction enzyme will yield a topoisomerase ligation site;
- (iii) allowing sufficient time for the donor oligonucleotide to ligate to and thereby extend the receiver DNA;

(iv) moving the 5' end of the receiver DNA thus extended into the reserve chamber, by means of an electrical force, e.g., so that the restriction enzyme cleaves the receiver DNA to provide a topoisomerase ligation site, or in the case of single nucleotide addition, the deblocker, e.g., phosphatase, generates a 5' unblocked nucleotide on the single stranded DNA; and

(v) repeating the cycle of steps (i) – (iv), adding oligonucleotides having the same or different informational sequence, until the desired DNA sequence or sequences are obtained.

[0115] For example, the invention provides

2.1. Method 2 wherein the 3' end of the receiver DNA is attached proximate to a nanopore and the 5' end of the receiver oligonucleotide comprises a topoisomerase ligation site, and comprising a step after step (iv) of adding an additional oligonucleotide to the 5' end of the receiver DNA by flushing the first addition chamber and providing new topoisomerase-charged donor oligonucleotide to the first addition chamber, wherein the new donor oligonucleotide has a different informational sequence from the previous donor oligonucleotide; and if desired that the new donor oligonucleotide be added to the receiver DNA, drawing the 5' end of the receiver nucleic acid back into the first chamber, and repeating steps (i) – (iii), or if not so desired, allowing the receiver DNA to remain in the second chamber until the desired donor oligonucleotide is provided to the first chamber.

2.2. Any foregoing method wherein a multiplicity of receiver DNA molecules are synthesized independently in parallel, such that DNA molecules having different sequences are obtained by separately controlling whether they are present in the first chamber.

2.3. Any foregoing method wherein a multiplicity of receiver DNA molecules each bound at the 3' end to a surface proximate to a nanopore are synthesized independently, wherein each nanopore has an associated pair of electrodes, wherein one electrode in the pair is located proximate to one end of the nanopore and the other electrode located proximate to the other end of the nanopore, such that each receiver DNA molecule can be

independently moved between the first and second chamber by current provided by the pair of electrodes.

2.4. Any foregoing method wherein the donor oligonucleotides used in step (i) of the cycle alternate with each cycle between donor oligonucleotides comprising a first informational sequence and donor oligonucleotides comprising a second informational sequence.

2.5. Method 2 comprising the step of adding an additional oligonucleotide to the 5' end of the receiver DNA by returning the 5' end of the receiver DNA to the first addition chamber to add an oligonucleotide having the same informational sequence or moving the 5' end of the receiver DNA to a second addition chamber to having a donor oligonucleotide bound at the 3' end to a topoisomerase, wherein the donor oligonucleotide in the second addition chamber has a different informational sequence from the donor oligonucleotide in the first addition chamber.

2.6. Any foregoing method wherein the donor oligonucleotide comprises a structure as follows:

```
5' CGAAGGG <Informational sequence A or B> GTCGACNNNNN
3' GCTTCCC <-----Complement-----> CAGCTGNNNNN
```

wherein N refers to any nucleotide and the restriction enzyme is AccI, which can cut the DNA (e.g. GTCGAC in the above sequence) so as to provide an appropriate overhang.

2.7. Any foregoing method wherein the donor oligonucleotide has a hairpin structure, e.g., 2.6 wherein the NNNNN groups on the top and bottom strands are joined.

2.8. Any foregoing method wherein at least one of the topoisomerase charged oligonucleotides has a structure as follows:

```
5' CGAAGGG <Informational sequence A or B> GTCGACNNNNN
3' *TTCCC <-----Complement-----> CAGCTGNNNNN
(* = topoisomerase)
```

2.9. Any foregoing method wherein at least one of the topoisomerase charged oligonucleotides has a structure as follows:

```
5' pCACGTCAGGCGTATCCATCCCTT*
3' GTGCAGTCCGCATAGGTAGGGAAGCGC
```

- 2.10. The preceding method wherein the topoisomerase charged oligonucleotide
- 2.11. Any foregoing method wherein the sequence of DNA synthesized is determined following each cycle by detecting the change in electric potential, current, resistance, capacitance and/or impedance as the oligonucleotide passes through the nanopore.
- 2.12. Any foregoing method wherein the synthesis of the DNA takes place in a buffer solution, e.g., a solution comprising a buffer for pH 7-8.5, e.g. ca. pH 8, e.g, a buffer comprising tris(hydroxymethyl)aminomethane (Tris), a suitable acid, and optionally a chelater, e.g., ethylenediaminetetraacetic acid (EDTA), for example TAE buffer containing a mixture of Tris base, acetic acid and EDTA or TBE buffer comprising a mixture of Tris base, boric acid and EDTA; for example a solution comprising 10mM Tris pH 8, 1 mM EDTA, 150 mM KCl, or for example, 50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, pH 7.9 @ 25°C.
- 2.13. Any foregoing method further comprising removing the DNA from the nanochip.
- 2.14. Any foregoing method further comprising amplifying the DNA thus synthesized.
- 2.15. Any foregoing method further comprising removing the DNA from the nanochip and crystallizing the DNA.
- 2.16. Any foregoing method further comprising stabilizing the DNA, e.g., by drying a solution comprising the DNA together with one or more of a buffer (e.g., a borate buffer), an antioxidant, a humectant, e.g. a polyol, and optionally a chelator, for example as described in US 8283165 B2, incorporated herein by reference, or by forming a matrix between the nucleic acid and a polymer, such as poly(ethylene glycol)–poly(l-lysine) (PEG–PLL) AB type block copolymer.
- 2.17. Any foregoing method comprising providing ligase and ATP to seal nicks in the DNA [NB: the topoisomerase ligation only ligates one strand].
- 2.18. Any foregoing method wherein the topoisomerase-charged donor oligonucleotide comprises a 5' overhang on the strand complementary to the strand bearing the topoisomerase, comprising a polyinosine sequence [NB: inosines act as 'universal bases' and pair with any other base].
- 2.19. Any foregoing method wherein the restriction enzyme is a type IIS restriction enzyme which can cleave all of the added DNA with the exception of a single base (the base which is being 'added').

- 2.20. Any foregoing method wherein the topoisomerase is selected from vaccinia topoisomerase and SVF topoisomerase I.
- 2.21. Any foregoing method wherein vaccinia topoisomerase (which recognizes (C/T)CCTT) is used to add dTTP nucleotides and SVF topoisomerase I (which recognizes CCCTG) is used to add dGTP nucleotides, e.g., to provide binary code information.
- 2.22. Any foregoing method wherein the reserve chamber further comprises a ligase and ATP, to repair the DNA strand not joined by the topoisomerase .
- 2.23. Any foregoing method comprising use of a topoisomerase inhibitor to suppress binding and activity of free topoisomerase to the DNA oligomer, e.g., wherein the inhibitors is selected from novobiocin and coumermycin.
- 2.24. Any foregoing method wherein the DNA strand thus provided has a sequence comprising thymidine (T) nucleosides and deoxyguanosine (G) nucleosides.
- 2.25. Any foregoing method wherein the topoisomerase adds a single base, but the restriction enzyme cleaves at a position which is one nucleotide in the 5' direction from the base added by topoisomerase.
- 2.26. Any foregoing method wherein the DNA strand thus provided has a sequence comprising a sequence of 'TT' and 'TG' dinucleotides.
- 2.27. Any foregoing method which is a method of synthesizing a DNA molecule by adding single nucleotides in the 3' to 5' direction comprising (i) reacting a DNA molecule with a topoisomerase charged with the desired nucleotide in 5'protected form, e.g., 5'phosphorylated form, such that the desired nucleotide in 5'protected form is added to the 5' end of the DNA, then (ii) deprotecting the 5' end of the DNA thus formed through the use of a phosphatase enzyme, and repeating steps (i) and (ii) until the desired nucleotide sequence is obtained.
- 2.28. Any foregoing method which is a method of synthesizing a DNA molecule by adding oligomers in the 3' to 5' direction comprising (i) reacting a DNA molecule with a topoisomerase charged with the desired oligomer, thereby ligating the oligomer to the DNA molecule, then (ii) using a restriction enzyme to provide a 5' site for a topoisomerase-mediated ligation for another oligomer, and repeating steps (i) and (ii) until the desired nucleotide sequence is obtained.

2.29. Any foregoing method which is a method in accordance with any of Method A, *et seq.*

[0116] The product of the synthesis reactions can be detected, reviewed for quality control purposes, and read to extract the data encoded on the polymer. For example the DNA may be amplified and sequenced by conventional means to confirm that the nanopore sequencing is robust.

[0117] In another embodiment, the invention provides an oligonucleotide comprising a topoisomerase binding site, an informational sequence (e.g., selected from at least two different sequences, e.g., wherein one sequence corresponds to '0' and the other to '1' in a binary code), and a restriction site which when cleaved by a restriction enzyme will yield a topoisomerase ligation site, e.g., comprising the following sequence:

```
5' CGAAGGG <Informational sequence A or B> GTCGAC
3' GCTTCCC <-----Complement-----> CAGCTG
```

wherein the Informational Sequence A or B is a sequence of 3-12, e.g., about 8 nucleotides.

[0118] In another embodiment, the invention provides a topoisomerase charged oligonucleotide wherein the oligonucleotide comprises a topoisomerase binding site, an informational sequence (e.g., selected from at least two different sequences, e.g., wherein one sequence corresponds to '0' and the other to '1' in a binary code), and a restriction site which when cleaved by a restriction enzyme will yield a topoisomerase ligation site; for example a topoisomerase charged oligonucleotide having a structure as follows:

```
5' CGAAGGG <Informational sequence A or B> GTCGACNNNNN
3' *TTCCC <-----Complement-----> CAGCTGNNNNN
```

wherein the Informational Sequence A or B is a sequence of 3-12, e.g., about 8 nucleotides and * is topoisomerase covalently bound to the oligonucleotide; e.g., wherein the topoisomerase is Vaccinia virus topoisomerase I.

[0119] In another embodiment, the invention provides a single or double stranded DNA molecule as described above, wherein the single strand or the coding sequence consists essentially of nonhybridizing bases, for example adenines and cytosines (As and Cs), which are arranged in sequence to correspond to a binary code, e.g., for use in a method of data storage.

For example, the invention provides DNA (DNA 1), wherein the DNA is single or double stranded, at least 1000 nucleotides long, e.g., 1000 – 1,000,000 nucleotides or, for example, 5,000 to 20,000 nucleotides long, wherein the sequence of the nucleotides corresponds to a binary code; e.g.,

- 1.1. DNA 1 wherein the DNA is single stranded.
- 1.2. DNA 1 wherein the DNA is double stranded.
- 1.3. Any foregoing DNA wherein the nucleotides in a single strand or in the coding strand are selected from adenine, thymine and cytosine nucleotides, e.g. are selected from adenine and cytosine nucleotides or thymine and cytosine nucleotides
- 1.4. Any foregoing DNA consisting primarily of nonhybridizing nucleotides, so that it will not form significant secondary structures when in the form of a single strand.
- 1.5. Any foregoing DNA wherein the nucleotides are at least 95%, e.g. 99%, e.g., 100% adenine and cytosine nucleotides.
- 1.6. Any foregoing DNA comprising a nucleotide or sequence of nucleotides added to separate or punctuate the nucleotides comprising a binary code, e.g., to separate the 1's and 0's or groups of 1's and 0's, so that consecutive 1's or 0's can be more easily read.
- 1.7. Any foregoing DNA wherein (a) each bit in the binary code corresponds to a single nucleotide, e.g. each of 1 and 0 correspond to A or C; or (b) each bit in the binary code corresponds to a series of more than 1 nucleotides, e.g. 2, 3 or 4 nucleotides, e.g., AAA or CCC.
- 1.8. Any foregoing DNA which is crystallized.
- 1.9. Any foregoing DNA which is provided in a dry form together with one or more of a buffer salt (e.g., a borate buffer), an antioxidant, a humectant, e.g. a polyol, and optionally a chelator, for example as described in US 8283165 B2, incorporated herein by reference; and/or in a matrix between the nucleic acid and a polymer, such as poly(ethylene glycol)–poly(l-lysine) (PEG–PLL) AB type block copolymer; ; and/or together with a complementary nucleic acid strand or a protein that binds the DNA.
- 1.10. Any foregoing DNA made by any of Method 1 *et seq.* or Method 2 *et seq.* or Method A, *et seq.*.

[0120] The nanochips can be fabricated for example as depicted in figures 23-29. For example, in one format, each polymer strand is associated with two or four addition chambers, wherein the

two addition chamber format is useful for encoding binary code in the polymer, and the four addition chamber format is particularly useful for making custom DNA sequences. Each addition chamber contains a separately controllable electrode. The addition chambers contain reagents to add monomers to the polymer in buffer. The addition chambers are separated by a membrane comprising one or more nanopores from a reserve chamber, which may be common to multiple addition chambers, and which contains deprotection reagents and buffer, to deprotect the protected monomers or oligomers added in the addition chambers. The nanochips comprise a multiplicity of addition chamber sets, to allow parallel synthesis of many polymers.

[0121] High-bandwidth and low-noise nanopore sensor and detection electronics are important to achieving single-DNA base resolution. In certain embodiments, the nanochip is electrically linked to a Complementary Metal-Oxide Semiconductor (CMOS) chip. Solid-state nanopores can be integrated within a CMOS platform, in close proximity to the biasing electrodes and custom-designed amplifier electronics, e.g., as described in Uddin, et al., “Integration of solid-state nanopores in a 0.5 μm cmos foundry process”, *Nanotechnology* (2013) 24(15): 155501, the contents of which are incorporated herein by reference.

[0122] In another embodiment, the disclosure provides a nanochip (Nanochip 1) for synthesis of and/or sequencing an electrically charged polymer, e.g., DNA, comprising at least two distinct monomers, the nanochip comprising at least a first and second reaction chambers, separated by a membrane comprising one or more nanopores, wherein each reaction chamber comprises one or more electrodes to draw the electrically charged polymer into the chamber and further comprises an electrolytic media and optionally reagents for addition of monomers to the polymer, for example,

- 1.1. Nanochip 1 wherein the nanopore has a diameter of 2-20 nm, e.g. 2-10 nm, for example 2-5 nm.
- 1.2. Any foregoing nanochip wherein the some or all of the walls of the reaction chambers of the nanochip comprise a silicon material, e.g., silicon, silicon dioxide, silicon nitride, or combinations thereof, for example silicon nitride.
- 1.3. Any foregoing nanochip wherein the some or all of the walls of the reaction chambers of the nanochip comprise a silicon material, e.g., silicon, silicon dioxide, silicon nitride, or combinations thereof, for example silicon nitride, and some or all of the nanopores are made by ion bombardment.

- 1.4. Any foregoing nanochip wherein some or all of the nanopores are comprised of a pore-forming protein, α -hemolysin, in a membrane, e.g. a lipid bilayer.
- 1.5. Any foregoing nanochip wherein some or all of the walls of the reaction chambers are coated to minimize interactions with the reagents, e.g., coated with a polymer such as polyethylene glycol, or with a protein, such as a bovine serum albumin.
- 1.6. Any foregoing nanochip comprising an electrolyte media.
- 1.7. Any foregoing nanochip comprising an electrolyte media comprising a buffer, e.g., a buffer for pH 7-8.5, e.g. ca. pH 8, e.g, a buffer comprising tris(hydroxymethyl)aminomethane (Tris), a suitable acid, and optionally a chelater, e.g., ethylenediaminetetraacetic acid (EDTA), for example TAE buffer containing a mixture of Tris base, acetic acid and EDTA or TBE buffer comprising a mixture of Tris base, boric acid and EDTA; for example a solution comprising 10mM Tris pH 8, 1 mM EDTA, 150 mM KCl, or for example, 50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, pH 7.9 @ 25°C.
- 1.8. Any foregoing nanochip comprising reagents for addition of monomers to the polymer.
- 1.9. Any foregoing nanochip capable of both synthesizing (“writing”, e.g., by adding monomers or groups of monomers sequentially to the polymer) and sequencing (“reading”, e.g., by measuring changes in current and/or inductance as the monomers pass through the nanopore) the polymer.
- 1.10. Any foregoing nanochip wherein the membrane comprising one or more nanopores comprises a metal surface on both sides, the metal surface being separated by an insulator, e.g. a silicon nitride membrane, the metal surfaces being configured, e.g., by lithographic means, to provide electrodes at either end of each nanopore, e.g., such that a current flow across the nanopore may be established through the nanopore via an electrolyte media, e.g., such that the current can draw the polymer through the nanopore and as the polymer passes through the nanopore, the change in electric potential across the nanopore can be measured and used to identify the sequence of monomers in the polymer.
- 1.11. Any foregoing nanochip comprising an electrically charged polymer which is DNA.

- 1.12. Any foregoing nanochip comprising an electrically charged polymer which is single stranded DNA (ssDNA).
- 1.13. Any foregoing nanochip comprising an electrically charged polymer which is DNA comprising a predetermined restriction site.
- 1.14. Any foregoing nanochip comprising an electrically charged polymer which is DNA wherein the DNA is a DNA as described in any of DNA 1, *et seq.*, above.
- 1.15. Any foregoing nanochip comprising an electrically charged polymer which is DNA, wherein the DNA comprises at least 95%, e.g. 99%, e.g., 100% adenines and cytosines.
- 1.16. Any foregoing nanochip comprising an electrically charged polymer which is DNA, wherein the DNA comprises only adenines and cytosines.
- 1.17. Any foregoing nanochip comprising one or more ports to permit introduction of and flushing out of buffer and reagents.
- 1.18. Any foregoing nanochip comprising a buffer solution, e.g., a solution comprising a buffer for pH 7-8.5, e.g. ca. pH 8, e.g, a buffer comprising tris(hydroxymethyl)aminomethane (Tris), a suitable acid, and optionally a chelater, e.g., ethylenediaminetetraacetic acid (EDTA), for example TAE buffer containing a mixture of Tris base, acetic acid and EDTA or TBE buffer comprising a mixture of Tris base, boric acid and EDTA; for example a solution comprising 10mM Tris pH 8, 1 mM EDTA, 150 mM KCl, or for example, 50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, pH 7.9 @ 25°C..
- 1.19. Any foregoing nanochip which is or is capable of being lyophilized for storage and subsequently rehydrated, e.g., wherein the structure of the nanochip comprises a hydratable or water permeable polymer.
- 1.20. Any foregoing nanochip which is synthesized in a dry form, e.g., wherein the structure of the nanochip comprises a hydratable or water permeable polymer, followed by hydration prior to use, optionally followed by lyophilization for long term storage once the write process is complete.
- 1.21. Any foregoing nanochip wherein the electrically charged polymer, e.g., DNA, is stabilized with histone.
- 1.22. Any foregoing nanochip wherein the interior surface is positively charged.

- 1.23. Any foregoing nanochip wherein the electrodes are operably connected in a capacitive circuit capable of providing a radiofrequency pulsating direct current, e.g. at a frequency of 1 MHz to 1GHz, e.g. 50-200MHz, for example about 100MHz, across the nanopore, e.g., wherein the pulsating direct current can draw the charged polymer through the nanopore and the monomer sequence can be determined by measuring the capacitive variance across the nanopore as the charged polymer goes through the nanopore.
- 1.24. Any foregoing nanochip comprising a reserve or deblocker chamber, which contains reagents for deprotection of the polymer following addition of a monomer or oligomer in one of addition chambers.
- 1.25. Any foregoing nanochip comprising a multiplicity of pairs of addition chambers.
- 1.26. Any foregoing nanochip comprising an electrical control layer, a fluidics layer and an electrical ground layer, e.g., as depicted in figure 24, joined by wafer bonding.
- 1.27. Any foregoing nanochip wherein the nanopore is made by drilling with FIB, TEM, wet or dry etching.
- 1.28. Any foregoing nanochip wherein the membrane comprising the nanopores is from 1 atomic layer to 30 nm thick.
- 1.29. Any foregoing nanochip wherein the membrane comprising the nanopores is made of SiN, BN, SiOx, Graphene, transition metal dichalcogenides e.g. WS₂ or MoS₂.
- 1.30. Any foregoing nanochip comprising wiring made from metal or polysilicon.
- 1.31. Any foregoing nanochip wherein the wiring density is increased by 3D stacking, with electrical isolation provided by dielectric deposition (e.g., via PECVD, sputtering, ALD etc).
- 1.32. Any foregoing nanochip wherein the contact to the electrode in the addition chamber is made using Through Silicon Via (TSV) by Deep Reactive Ion Etch (DRIE), e.g. using cryo or BOSCH process, or via wet silicon etching.
- 1.33. Any foregoing nanochip wherein individual voltage control for the electrode in each addition chamber allows the electrode in each addition chamber to be controlled and monitored individually.
- 1.34. Any foregoing nanochip wherein each polymer is associated with a first addition chamber, a second addition chamber, and a deblocking chamber.

- 1.35. Any foregoing nanochip wherein one or more chambers have fluid flow.
- 1.36. Any foregoing nanochip wherein one or more chambers are fluidically isolated.
- 1.37. Any foregoing nanochip wherein the deblocking chamber has fluid flow..
- 1.38. Any foregoing nanochip wherein addition chambers have common fluid flow.
- 1.39. Any foregoing nanochip wherein wiring between chambers is common among chambers of similar type (e.g. among first addition chambers, among second addition chambers, and among deblocking chambers.)
- 1.40. Any foregoing nanochip wherein the addition chambers have individual voltage control and the deblocking chambers have a common electrical ground.
- 1.41. Any foregoing nanochip wherein the deblocking chambers have individual voltage control, the first addition chambers have a common electrical ground and the second addition chambers have a common electrical ground.
- 1.42. Any foregoing nanochip wherein the nanochip is fabricated by wafer bonding, and the chambers are prefilled with desired reagents prior to bonding.
- 1.43. Any foregoing nanochip wherein one or more internal surfaces are silanized.
- 1.44. Any foregoing nanochip which has one or more ports for introduction or removal of fluid.
- 1.45. Any foregoing nanochip wherein the electrodes in the chambers are restricted from direct contact with the charged polymer, e.g., wherein the electrode is placed too far from the nanopore to be reached by a charged polymer bound to a surface adjacent to the nanopore, or wherein the electrode is protected by a material which will permit the passage of water and single atom ions (e.g., Na⁺, K⁺ and Cl⁻ ions) but not the passage of the polymer or monomer or oligomer reagents to be joined to the polymer.
- 1.46. Any foregoing nanochip which is electrically linked to a Complementary Metal-Oxide Semiconductor (CMOS) chip.

[0123] For example, in one embodiment, the invention provides a nanochip, e.g., according to any of Nanochip 1, et seq., for sequencing an electrically charged polymer, e.g., DNA, comprising at least two distinct monomers, the nanochip comprising at least a first and second reaction chambers comprising an electrolyte media and separated by a membrane comprising one or more nanopores, wherein each reaction chamber comprises at least one pair of

electrodes disposed on opposite sides of the membrane, wherein the electrodes are operably connected in a capacitive circuit capable of providing a radiofrequency pulsating direct current, e.g. at a frequency of 1 MHz to 1GHz, e.g. 50-200MHz, for example about 100MHz, across the nanopore, e.g., wherein the pulsating direct current can draw the charged polymer through the nanopore and the monomer sequence can be determined by measuring the capacitive variance across the nanopore as the charged polymer goes through the nanopore.

[0124] In another embodiment, the invention provides a method of reading a monomer sequence of a charged polymer comprising at least two different types of monomers, for example a DNA molecule, comprising applying a radiofrequency pulsating direct current, e.g. at a frequency of 1 MHz to 1GHz, e.g. 50-200MHz, for example about 100MHz, across a nanopore, wherein the pulsating direct current draws the charged polymer through the nanopore and the monomer sequence is read by measuring the capacitive variance across the nanopore as the charged polymer goes through the nanopore.

[0125] In another embodiment, the invention provides the use of any of DNA 1, *et seq.* in a method for storing information.

[0126] In another embodiment, the invention provides the use of a single stranded DNA in a method for storing information, e.g., wherein the sequence is substantially non-self-hybridizing.

[0127] In another embodiment, the invention provides a method of data storage and device, using a nanochip, e.g., any of Nanochip 1 *et seq.* to make an electrically charged polymer, e.g., DNA, comprising at least two distinct monomers or oligomers, wherein the monomers or oligomers are arranged in sequence to correspond to a binary code, e.g., in accordance with any of the foregoing Methods 1 and/or 2 *et seq.*

[0128] For example, in one embodiment, the nanochip comprising the polymer thus synthesized provides a data storage device, as the nanochip can be activated and the sequence of the polymer detected by passing it through a nanopore at any time. In other embodiments the polymer is removed from the nanochip, or amplified and the amplified polymer removed from the nanochip, stored until required, and then read using a conventional sequencer, e.g., a conventional nanopore sequencing device,

[0129] In another embodiment, the invention provides a method of storing information comprising synthesizing any of DNA 1, et seq., e.g., in accordance with any of Methods 1, et seq. or Methods 2, et seq.

[0130] In another embodiment, the invention provides a method of reading a binary code, e.g., as encoded on any of DNA 1, et seq., using a nanopore sequencer, for example using Nanochip 1 et seq., as described herein.

[0131] Any foregoing method wherein the nanochip is erased using an enzyme which lyses the charged polymer, e.g., a deoxyribonuclease (DNase) to hydrolyze DNA.

EXAMPLES

Example 1 – Immobilizing one end of DNA adjacent to nanopore and controlled back and forth movement of DNA via electrical current

[0132] Experimental procedures are developed to demonstrate that DNA is moved back and forth between two chambers separated by a nanopore, via an electrical current, under conditions that a relevant protein does not move between chambers.

[0133] A nanochip comprising two chambers is fabricated from silicon nitride. Nanopores of <4 nm (for dsDNA or ssDNA) and 2 nm (for ssDNA only) are prepared as described in Briggs K, et al. *Automated fabrication of 2-nm solid-state nanopores for nucleic acid analysis*, Small (2014)10(10):2077-86. The two chambers are referred to as a 'near' and 'far' chamber, the far chamber being the chamber where 3' end of DNA is conjugated.

[0134] It is shown that ssDNA (2nm pore) and ssDNA+dsDNA (4nm pore) but not protein pass through the nanopore. Passing through the nanopores is detected by electrical current disruption.

[0135] *Conjugation of DNA to pore surface:* Attach 5' amino modified DNA to carboxy-coated polystyrene beads (Fluoresbrite® BB Carboxylate Microspheres 0.05µm, from Polysciences, Inc.) via carbodiimide mediated attachment. 3' of DNA is labeled with biotin. DNA is of a pre-specified length.

[0136] *Streptavidin conjugation:* Conjugation was performed on the 'far' side of a silicon nitride nanopore conjugate streptavidin to the surface, as described in Arafat, A. *Covalent Biofunctionalization of Silicon Nitride Surfaces*. Langmuir (2007) 23 (11): 6233-6244.

[0137] *Immobilization of DNA near the nanopore:* DNA conjugated polystyrene beads in buffer is added to 'near' chamber and buffer is added to 'far' chamber (standard buffer: 10mM Tris pH

8, 1 mM EDTA, 150 mM KCl). Voltage (~100 mV) is applied until current disruption is observed (use an Axon Nanopatch200B patch-clamp amplifier). 50 nm beads cannot pass through the nanopore, so when a DNA strand has gone through and a bead is pressed against an end of the nanopore the current is highly disrupted. Current is maintained 1-2 mins until DNA is irreversibly bound to immobilized streptavidin on the far side via binding of biotin. To confirm that the DNA has been immobilized, the current is reversed. Different currents are observed if DNA is in or out of the pore. If it appears that DNA is not immobilized, then the process is repeated.

[0138] Release the bead via endonuclease: Restriction enzyme in restriction enzyme buffer is added to the chamber where the DNA is attached. In one embodiment, the DNA is single stranded and contains a restriction site cleavable by an enzyme that will cleave single stranded DNA. See, e.g., Nishigaki, K., *Type II restriction endonucleases cleave single-stranded DNAs in general*. Nucleic Acids Res. (1985) 13(16): 5747–5760. In an alternative embodiment, a complementary oligonucleotide is added to the chamber where the DNA is attached and allowed to hybridize for 30 minutes to create dsDNA, then the restriction enzyme is added. Once the bead is released, it is washed away. Current is switched between forward and reverse to confirm that the DNA goes into/through and out of the pore.

[0139] Demonstrate controlled back and forth movement: Using standard buffer, current is applied in forward direction until signal disruption is observed and then reverted to ‘normal’ after the DNA passes. Reverse current is applied until signal disruption is observed. It is observed that the signal does not go back to normal as the DNA remains in the pore. Application of current in forward and reverse direction is repeated for several cycles to confirm that DNA moves back and forth through the nanopore.

Example 1a: Immobilizing DNA strand adjacent to nanopore in a silicon dioxide chip

[0140] A nanochip interior wall is fabricated from silicon dioxide. Both sides are silanized, but the oligonucleotide is conjugated to just one side of the chip wall, then a nanopore is created.

[0141] Silanization: The surface of the chip wall is cleaned with piranha solution (various brands commercially available, generally comprising a mixture of sulfuric acid (H₂SO₄) and hydrogen peroxide (H₂O₂), which remove organic residues from the surface) at 30°C, and washed with double-distilled water. A stock solution of (3-aminopropyl)triethoxysilane (APTES) is prepared,

with 50% methanol (MeOH), 47.5% APTES, 2.5% nanopure H₂O, and aged > 1 hour at 4°C. The APTES stock is then diluted 1:500 in MeOH and applied to and incubated with the chip wall at room temperature. The chip wall is then rinsed with MeOH and dried at 110°C for 30 minutes.

[0142] *Conjugation:* The chip wall is then incubated for 5 hours at room temperature in a 0.5% w/v solution of 1,4-phenylene diisothiocyanate (PDC) in dimethyl sulfoxide (DMSO). It is washed briefly twice with DMSO and the briefly twice with double distilled water. The chip wall is then incubated with 100 nM amine-modified single stranded DNA oligomers (ca. 50-mers) in double distilled water (pH 8) overnight at 37°C. Then the chip wall is washed twice with 28% ammonia solution to deactivate any unreacted material, and washed twice with double distilled water. One or more nanopores are then created in the wall.

[0143] Once the fabrication of the nanochip is complete, the interior wall is coated with DNA oligomers ca. 50bp long. This permits a single stranded DNA having an end-terminal sequence complementary to the surface bound DNA to be localized to a nanopore by attaching the ssDNA to a relatively bulky structure (e.g. a bead, a protein, or a DNA origami structure having a diameter too large to fit through the nanopore,), wherein the sequence complementary to the surface-bound DNA is distal to the bulky structure, pulling the charged polymer through the nanopore using current, allowing the ssDNA to bind to a complementary surface bound DNA oligomer adjacent to the nanopore, and cleaving off the bulky structure.

Example 2: DNA synthesis – single nucleotide addition

[0144] DNA is moved to ‘reserve’ chamber by applying appropriate current and detecting DNA movement.

[0145] Terminal transferase enzyme (TdT, New England Biolabs) in appropriate buffer (50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, pH 7.9 @ 25°C), plus reversibly blocked-dATP* is added to the ‘addition’ chamber. The buffer is also added to the ‘reserve’ chamber.

[0146] dNTPs that have reversible blocks on the 3’ –OH are used to add nucleotides to the DNA. When added to the DNA chain, the next dNTP cannot be added until the blocked dNTP is unblocked.

[0147] Deblocking can be chemical or enzymatic. Different approaches are utilized:

- [0148] a. 3' O-allyl: Allyl is removed by Pd-catalyzed deallylation in aqueous buffer solution as described in Ju J, *Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators*. Proc Natl Acad Sci U S A. (2006);103(52):19635-40; or by using using iodine (10 mol%) in polyethylene glycol-400, as described in Shankaraiah G., et al., *Rapid and selective deallylation of allyl ethers and esters using iodine in polyethylene glycol-400*. Green Chem. (2011)13: 2354-2358
- [0149] b. 3' O-NH₂: Amine is removed in buffered NaNO₂, as described in US 8034923.
- [0150] c. 3'-phosphate. Phosphate is hydrolyzed with Endonuclease IV (New England Biolabs). Other possible 3' modifications which can also be removed with Endonuclease IV include phosphoglycoaldehyde and deoxyribose-5-phosphate.
- [0151] d. 3'-O-Ac: Acetate is removed by enzymatic hydrolysis as described in Ud-Dean, A *theoretical model for template-free synthesis of long DNA sequence*. Syst Synth Biol (2008) 2:67-73,
- [0152] The DNA is then moved to the 'far' chamber by applying appropriate current and detecting DNA movement. DNA is deprotected by switching out buffers and adding deblocking buffer/solution as described in a. – d. above.
- [0153] Process is repeated as desired to make sequence of interest.

Example 3: DNA synthesis: block oligonucleotide addition

- [0154] The 3' end of double stranded DNA is attached adjacent to a nanopore with 4nm aperture. The 5' end of the DNA has an overhang of CG (reading from 5' to 3').
- [0155] Oligo cassettes A and B are made as follows:
- 5' CGAAGGG <CODEA OR B> GTCGACNNNNN
3' GCTTCCC <COMPLEMENT> CAGCTGNNNNN
- [0156] CodeA and CodeB each represent an informational sequence. Ns refer to any nucleotide. The 5' sequence comprises a topoisomerase recognition site and the 3' sequence comprises an Acc1 restriction site. The oligo is exposed to topoisomerase and the topoisomerase binds to 3' thymidine:
- 5' CGAAGGG <CODEA OR B> GTCGACNNNNN
3' *TTCCC <COMPLEMENT> CAGCTGNNNNN (* = topoisomerase)

[0157] DNA is moved to 'near' chamber by applying appropriate current and detecting DNA movement. The topoisomerase-charged 'codeA' oligo is provided in the 'addition' chamber. The DNA is moved into the addition chamber by applying appropriate current and detecting DNA movement, whereupon the code A oligo is bound to the DNA, Acc1 is added to the 'reserve' chamber, where it cleaves at the restriction site to provide a topoisomerase ligation site.

[0158] The process is repeated until the desired sequence is reached, adding other 'code A' or 'code B.' Note that it is not required to continually add new Acc1 to the 'reserve' chamber; it is just needed to flush out codeA or codeB oligos in 'addition' chamber when switching from codeA or codeB.

[0159] For sequencing a pore that only allows ssDNA to pass, some modifications to the protocol above are required. It is already known that when dsDNA encounters a small pore (2nm) only ssDNA will go through and the complement will be 'stripped' off. Thus, if doing this synthesis with a 2nm pore one must ensure that the proper dsDNA is able to 'reform' on the other side. To do this one would add "CGAAGGG <CODEA OR B> GTCGACNNNNN" to the near chamber (to ensure a restriction site is created) and "CGAAGGG <CODEA OR B> GT" to the far chamber (to ensure a topo-compatible site is generated).

[0160] Elaborating on the foregoing method, we demonstrate the sequential 'addition' of DNA-encoded information into a growing DNA chain with ≥ 2 sequential additions (representing 2 bits of data), each of which comprise an 'add' and a 'deprotect' step. Initial experiments for optimization and proof of concept are performed in microtubes.

[0161] In the approach described in this example, one bit of information is encoded in a string of nucleotides. The DNA bit to be 'added' is a short dsDNA sequence conjugated to vaccinia topoisomerase I (topo). In the presence of a suitable 'deprotected' 'acceptor' DNA, the topo-charged DNA 'bit' is enzymatically and covalently linked ('added') to the acceptor by the topoisomerase, which in the process becomes removed from the DNA. A restriction enzyme can then cleave the added bit to 'deprotect' it and create of suitable 'acceptor' sequence for addition of the next bit.

[0162] *Topo Charging:* A generic charging scheme is as follows, depicted schematically in figure 22 and below, where N indicates any nucleotide, and A, T, G, and C represent nucleotides with adenine, thymine, guanine and cytosine bases respectively. N's on top of one another are

complementary. While this example uses the restriction enzyme HpyCH4III, the basic strategy will work with other restriction enzymes, e.g., as demonstrated in Example 4.

```
N-N-N-N-N-N-N-N-N      A-G-G-G-N-N-N-N-N-N-N-N-N-N
N-N-N-N-N-N-N-N-N-T-T-C-C-C-N-N-N-N-N-N-N-N-N-N
```

+

Topoisomerase *

=

```
N-N-N-N-N-N-N-N-N
N-N-N-N-N-N-N-N-N
```

+

```
      A-G-G-G-N-N-N-N-N-N-N-N-N-N
*-T-T-C-C-C-N-N-N-N-N-N-N-N-N-N
(topo charged)
(N's on top of one another are complementary)
```

Addition

Generic 'add' reaction:

```
N-N-N-N-N-N-N-N-N-N-A
N-N-N-N-N-N-N-N-N-N..
(acceptor)
```

+

```
      A-G-G-G-N-N-N-N-N-N-N-N-N-N
*-T-T-C-C-C-N-N-N-N-N-N-N-N-N-N
```

(topo charged)

=

```
N-N-N-N-N-N-N-N-N-N-A A-G-G-G-N-N-N-N-N-N-N-N-N-N
N-N-N-N-N-N-N-N-N-N-T-T-C-C-C-N-N-N-N-N-N-N-N-N-N
```

+

*

(free topo)

Deprotection

Generic 'deprotection' reaction:

...N-N-**A-C-A-G-T**-N-N-N-N-N-N-N-N-N-N
 ...N-N-**T-G-T-C-A**-N-N-N-N-N-N-N-N-N-N

+

HpyCH4III
 (restriction enzyme)

=

...N-N-**A-C-A**
 ...N-N-**T-G**. .
 (deprotected)

+

. .**G-T**-N-N-N-N-N-N-N-N-N-N
T-C-A-N-N-N-N-N-N-N-N-N-N
 (side product)

[0163] The following oligonucleotides are ordered from Integrated DNA Technologies (IDT).

The “b” at the end of some of the oligonucleotides indicates biotin):

BAB:

CGATAGTCTAGGCACTGTTTGCTGCGCCCTTGTCGGTGTCGCCCTTATCTACTT
 AAGAGATCATACAGCATTGCGAGTACG

B1: b-CACGTACTCGCAATGCTGTATGATCTCTTAAGTAGATA

B2: ATCTACTTAAGAGATCATACAGCATTGCGAGTACG

TA1: b-CACACTCATGCCGCTGTAGTCACTATCGGAAT

TA2: AGGGCGACACGGACAGTTTGAATCATACCG

TA3b:

AACTTAGTATGACGGTATGATTCAAAGTGTCCGTGTCGCCCTTATTCCG
 ATAGTGACTACAGCGGCATGAG

TB1: b-CACACTCATGCCGCTGTAGTCACTATCGGAAT

TB2: AGGGCGCAGCAAACAGTGCCTAGACTATCG

TB3b:

AACTTAGTATGACGATAGTCTAGGCACTGTTTGCTGCGCCCTTATTCCGATAG
 TGACTACAGCGGCATGAG

FP1: CACGTACTCGCAATGCT

FP2: CGGTATGATTCAAACGTGCCG

FP3: GCCCTTGTCCGTGTC

[0164] Oligonucleotides are solubilized to 100uM in TE buffer and stored at -20C.

[0165] Hybridized oligonucleotides are made by mixing oligonucleotides as described below, heating to 95°C for 5 minutes, and then dropping the temperature by 5°C every 3 minutes until the temperature reaches 20°C. Hybridized oligonucleotides are stored at 4°C or -20°C. The combinations of oligonucleotide are as follows:

B1/2

48 uL B1

48 uL B2

4 uL 5M NaCl

A5

20 uL TA1

20 uL TA2

5 uL TA3b

4 uL 5M NaCl

51 uL TE

B5

20 uL TB1

20 uL TB2

5 uL TB3b

4 uL 5M NaCl

51 uL TE

[0166] The following buffers and enzymes are used in this example:

TE: 10M Tris pH 8.0, 1mM EDTA, pH 8.0

WB: 1M NaCl, 10mM Tris pH8.0, 1mM EDTA pH8.0

1x Topo: 20mM Tris pH7.5, 100mM NaCl, 2mM DTT, 5mM MgCl₂

1x RE: 50mM K-acetate, 20mM Tris-acetate, 10mM Mg-acetate, 100ug/ml BSA pH 7.9 @ 25C.

Vaccinia DNA Topoisomerase I (**topo**) is purchased from Monserate Biotech (10,000 U/mL)

HypCH4III is purchased from NEB

Streptavidin-coated magnetic beads (**s-MagBeads**) are purchased from ThermoFisher.

[0167] Acceptor is prepared as follows: 5uL of s-magbeads are washed one time in 200uL WB (binding time 1 minute). 5uL B1/2 + 195uL WB is added to beads and incubated 15 minutes at room temperature, then washed one time with 200uL WB, then washed one time with 200uL 1x Topo, and resuspended in 150uL of 1x Topo

[0168] Topo-charged A5 (see figure 20) is prepared as follows: 4uL 10x topo buffer + 23 uL water + 8uL A5 + 5 uL topo are incubated at 37°C for 30 minutes, added to 5uL s-magbeads (washed 1x with 200uL WB, 1x with 200uL 1x Topo, resuspended in 150uL 1x topo), and allowed to bind for 15 minutes at room temperature.

[0169] 'Add' charged A5 to Acceptor: s-magbeads are removed from from Topo-charged-A5, added to Acceptor, and incubated at 37°C for 60 minutes. The aliquot is removed, diluted 1/200 in TE, and stored at -20C

[0170] Deprotection: The material is washed one time with 200uL of WB, when washed one time with 200uL of 1x RE, and resuspended in 15uL 10x RE and 120uL water). 15uL HypCH4III is added. The mixture is incubated at 37°C for 60 minutes, then washed one time with 200uL WB, washed one time with 200uL 1x topo, to produce a product which we term 'Acceptor-A5'.

[0171] Topo charged B5 (see figure 21) is prepared as follows: 4uL 10x topo buffer. 23 uL water and 8uL B5 + 5 uL topo are combined and incubated at 37°C for 30 min. the product is added to 5uL s-magbeads (washed one time with 200uL WB, one time with 200uL 1x Topo, and resuspended in 150uL 1x topo) and allowed to bind for 15 minutes at room temperature.

[0172] 'Add' charged B5 to Acceptor-A5: s-magbeads are removed from Topo-charged-B5, added to Acceptor-A5 and incubated at 37°C for 60 minutes. The aliquot is then removed, diluted 1/200 in TE, and stored at -20°C

[0173] Deprotection: The material is washed one time with 200uL of WB, then washed one time with 200uL of 1x RE, and resuspended in 15uL 10x RE and 120uL water. 15uL HypCH4III is added, and the mixture is incubated at 37°C for 60 minutes.

[0174] Confirmation that the above reactions worked is provided by PCR amplification of aliquots from A5 (Acceptor with A5 added: step iii, 'A5 Added' in schematic) and B5 (Acceptor-A5 with B5 added: step vi, 'B5 Added' in schematic). 'No template' is used as negative control for A5, A5 is used as negative control for B5, oligo BAB is used as positive control for B5. The expected product size for A5 PCR is 68bp, the expected product size for B5 PCR is 57bp. (B1/2 is also run on the gel, expected size is ~47bp, but this may be approximate as there are overhangs and it is biotinylated). PCR reactions (30 cycles of 95/55/68 (1 minutes each) are carried out as follows:

	A5 (-) ctrl	A5	B5 (-) ctrl	B5
Template		1 uL	1 uL	1 uL
FP1	1 uL	1 uL	1 uL	1 uL
FP2	1 uL	1 uL		
FP3			1 uL	1 uL
Water	8 uL	7 uL	7 uL	7 uL
Maxima MM	10 uL	10 uL	10 uL	10 uL

[0175] SDS-PAGE using 4-20% Tris-glycine gels is used to confirm that expected size oligonucleotides are produced. Charging is performed as described above, but directly after charging (37°C incubation step), loading buffer is mixed in and samples are heated to 70°C for 2 minutes and allowed to cool prior to running the gel. Gel is stained with Coomassie. For the negative control, water is added to the reaction instead of topo. Figure 30 depicts the results, clearly showing bands corresponding to the expected product sizes for A5 PCR and for B5 PCR.

[0176] DNA bit addition via topoisomerase-charged DNA cassettes and deprotection performed via restriction enzyme are thus shown to be feasible. In these proof of concept experiments the

DNA is immobilized via streptavidin-conjugated magnetic beads, and moved sequentially into different reaction mixes, but in the nanopore chip format, we create separate reaction chambers and use electrical current to move the DNA into those different reaction chambers.

[0177] Finally, PCR demonstrates that the expected DNA sequences are created when performing sequential additions of DNA sequences corresponding to ‘bits’ of information. These reactions have worked as designed, even with minimal optimization.

[0178] DNA made as described in examples 2 and 3 is recovered and sequenced, using a commercial nanopore sequencer (MinION from Oxford Nanopore), confirming that the desired sequence is obtained.

Example 4 – DNA synthesis: block oligonucleotide addition, using a different restriction enzyme

[0179] The following synthesis is carried analogously to Example 3, but using the restriction enzyme MluI, which cuts at ‘ACGCGT’ to form:

```

...NNNA          CGCGTNNN...
...NNNTGCGC      ANNN...

```

In this example TOPO is charged to form a complex with sequence complementarity that will enable the charged TOPO to transfer DNA to DNA cut with MluI :

```

5' pCACGTCAGGCGTATCCATCCCTTCGCGTTCACGTACTCGCAATGCTGTAG
3' GTGCAGTCCGCATAGGTAGGGAAGCGC AGTGCATGAGCGTTACGAGATCb

```

+

TOPO

=

```

5' pCACGTCAGGCGTATCCATCCCTT*          (
3' GTGCAGTCCGCATAGGTAGGGAAGCGC
(** indicates TOPO bound at 3' phosphate)

```

+

CGCGTTCACGTACTCGCAATGCTGTAG
AGTGCATGAGCGTTACGAGATCb

(b=biotin. This can be removed with streptavidin)

[0180] By a process analogous to the preceding example, the charged TOPO is then used to add the oligomer to the 5' end of strand being synthesized, having a complementary acceptor sequence, thereby releasing the TOPO, and the strand is then "deprotected" using the MluI, and the cycle repeated until the desired sequence of oligomers is obtained.

Example 5 – Addition of single base using topoisomerase strategy

[0181] We have found that the topoisomerase system can also be designed to add single bases to a single stranded DNA chain (in comparison to Example 3, which describes adding 'cassettes'). The DNA bit to be 'added' is contained in a short DNA sequence conjugated to vaccinia topoisomerase I (topo). In the presence of a suitable single stranded 'deprotected' 'acceptor' DNA, the topo-charged DNA is enzymatically and covalently ligated ('added') to the acceptor by the topoisomerase, which in the process becomes removed from the DNA. A type IIS restriction enzyme can then cleave all of the added DNA with the exception of a single base (the base which is being 'added'). This process of deprotect-add is repeated to add additional bases (bits).

[0182] Topo Charging: A generic charging protocol is as follows, similar to Example 3:

...N-N-N-N-N-N-N-N-N-C-C-C-T-T-N-N-N-N-N-N-N-N-N-N-N-N-N-N...
...N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-I-I-I-I-I N-N-N-N-N-N-N-N...biotin

+

Topoisomerase (*)

=

N-N
N-N-N-N-N-N-N-N-N-N-N...biotin
(by-product)

+

...N-N-N-N-N-N-N-N-N-C-C-C-T-T*
...N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-I-I-I-I-I
(topo charged)

As in Example 3, the N's on top of one another are complementary. I is inosine. The biotin is

used to remove unreacted product and byproduct. Addition of a single base is carried out as follows

N-N-N-N-N-N-N-N-N-N...

(acceptor sequence nucleotides indicated in italics)

+

...N-N-N-N-N-N-N-N-N-C-C-C-T-T*

...N-N-N-N-N-N-N-N-N-N-N-N-N-N-I-I-I-I-I

(topo charged)

=

...N-N-N-N-N-N-N-N-N-C-C-C-T-T-*N-N-N-N-N-N-N-N-N-N...*

...N-N-N-N-N-N-N-N-N-N-N-N-N-N-I-I-I-I-I

+

*

(free topo)

Deprotection is illustrated as follows, using BciVI restriction enzyme (site in bold):

...N-**G-T-A-T-C-C**-N-N-C-C-C-T-T-*N-N-N-N-N-N-N-N-N-N...*

...N-**N-N-N-N-N-N**-N-N-N-N-N-N-N-N-I-I-I-I-I

+

BciVI

(restriction enzyme)

=

T-N-N-N-N-N-N-N-N-N-N...

(note a 'T' has been added to the 5' of the acceptor DNA)

÷

N-N-I-I-I-I-I

(dissociated*)

+

...N-G-T-A-T-C-C-N-N-C-C-C-T

...N-N-N-N-N-N-N-N-N-N-N-N

T-N-N-N-N-N-N-N-N-N-N...

N-N-I-I-I-I-I

=

$$T-N-N-N-N-N-N-N-N-N-N... + N-N-I-I-I-I-I-I$$

(NNIIIII dissociates from the single strand with added base)

[0183] The following oligonucleotides are synthesized commercially (B=biotin, P-phosphate, I = inosine):

NAT1 CACGTCAGGCGTATCCATCCCTTCACGTACTCGCAATGCTGTATGGCGAT
 NAT1b P-CACGTCAGGCGTATCCATCCCTTCACGTACTCGCAATGCTGTATGGCGAT-B
 NAT9cI P-IIIIIIAAGGGATGGATACGCCTGACGTG
 NAT9x P-ATCGCCATACAGCATTGCGAG
 NAT9 ACGTGAAGGGATGGATACGCCTGACGTG
 Nat9Acc CACGTAGCAGCAAACAGTGCCTAGACTATCG
 Nat1P CACGTCAGGCGTATCCATCC
 FP4 CGATAGTCTAGGCACTGTTTG

[0184] The oligonucleotides are solubilized to 100 μ M in TE buffer and stored at -20 °C.

[0185] Hybridization: The following hybridized oligonucleotides are made by mixing the oligonucleotides as described, heating to 95 °C for 5 minutes, and then dropping the temperature by 5 °C every 3' until the temperature reaches 20 °C. Hybridized oligonucleotides are stored at 4 °C or -20 °C.

NAT1b/NAT9cI/NAT9x

8 μ L NAT1B
 10 μ L NAT9cI
 10 μ L NAT9x
 48 μ L TE
 4 μ L 5M NaCl

NAT1/NAT9cI

10 μ L NAT1
 10 μ L NAT9cI
 80 μ L PBS

NAT1/NAT9

10 μ L NAT1

10 μ L NAT9

80 μ L PBS

[0186] Buffers & Enzymes: The following buffers are used:

TE: 10M Tris pH 8.0, 1mM EDTA, pH 8.0

PBS: phosphate buffered saline (137mM NaCl, 2.7mM KCl, 10mM Na₂HPO₄, 1.8mM KH₂PO₄)(pH 7.4)

10x Cutsmart: 500 mM KAc, 200 mM Tris-Ac, 100 mM Mg-Ac, 1 mg/mL BSA pH 7.9

BciVI is purchased from NEB and Streptavidin-coated magnetic beads (s-MagBeads) are purchased from ThermoFisher

[0187] The addition reaction is carried out as follows.

[0188] 1. Topo charge: The reagents are assembled as per table:

	Experiment	(-) control #1	(-) control #2
10x topo buffer	3	3	3
Water	17	21	23
NAT1B/NAT9cl/NAT9x	6	6	-
Topo	4	-	4

The reagents are then incubated at 37 °C for 30 minutes. The byproducts are removed using streptavidin magnetic beads (5 μ L) in 1x topo buffer after 10 minutes at room temperature to allow binding.

[0189] 2. Reaction: The reagents are assembled as per table:

	Experiment	(-) control #1	(-) control #2
From A.1.c	1	1	1

NAT9Acc	1	1	1
10x topo buffer	1	1	1
water	7	7	7

The reagents are then incubated at 37 °C for 30 minutes. The addition reaction is expected to proceed as follows:

NAT1B 5' p-CACGTCAGGCGTATCCATCCC**TT**CACGTACTCGCAATGCTGTATGCCGAT-B
NAT9cI 3' GTGCAGTCCGCATAGGTAGGGAIIIII GAGCGTTACGACATACCGCTA-p **NAT9x**

+ TOPO

=

5' CACGTCAGGCGTATCCATCCC**TT*** CACGTACTCGCAATGCTGTATGCCGAT-B
 3' GTGCAGTCCGCATAGGTAGGGAIIIII GAGCGTTACGACATACCGCTA-p

[0190] The asterisk (*) represents topoisomerase. Note that NAT9cI is phosphorylated, but this isn't shown for illustration purposes.

[0191] When the charged topo is in the presence of an acceptor sequence, it undergoes the following reaction:

5' p-CACGTCAGGCGTATCCATCCC**TT***
 GTGCAGTCCGCATAGGTAGGGAIIIII
 +
 CACGTAGCAGCAAACAGTGCCTAGACTATCG
 =
 5' p-CACGTCAGGCGTATCCATCCC**TT**CACGTAGCAGCAAACAGTGCCTAGACTATCG
 GTGCAGTCCGCATAGGTAGGGAIIIII

[0192] PCR amplification and measurement of the molecular weights of the product on agarose gel confirms the expected product is produced. See Figure 30, depicting correct sized band in lane 1 (experiment), no bands in negative controls.

[0193] B. Deprotection Reaction: The reagents are assembled as per table:

	1	2	3	4
NAT1/NAT9	1	1	-	-
NAT1/NAT9cI	-	-	1	1
10x cutsmart	2	2	2	2
water	17	16	17	16
BciVI	0	1	0	1

The reagents are incubated at 37 °C for 90 minutes. For the deprotection reaction, a representative product of an addition reaction is created using purchased oligonucleotides, and tested for digestion with the BciVI restriction enzyme:

NAT1 5' CACGTCAGGCGTATCCATCCC**TT**CACGTA**CTCGCAATGCTGTATGGCGAT**
 NAT9cI 3' GTGCAGTCCGCATAGGTAGGGAAIIIII

+

BciVI

=

NAT1 5' CACGTCAGGCGTATCCATCCC**T** **T**CACGTA**CTCGCAATGCTGTATGGCGAT**
 NAT9cI 3' GTGCAGTCCGCATAGGTAGGG AIIIII

It was not known whether the the restriction enzyme would cut the DNA as intended, given that 3' of the cut site are a series of inosines as opposed to 'regular' bases. As a positive control, the 'appropriately' base-paired equivalent of NAT1/NAT9cI is made (NAT1/NAT9c):

NAT1 5' CACGTCAGGCGTATCCATCCC**TT**CACGTA**CTCGCAATGCTGTATGGCGAT**
 NAT9c 3' GTGCAGTCCGCATAGGTAGGGAAAGTGCA

PCR amplification of the product followed by measurement of molecular weight on agarose gels

(Figure 31) shows that the enzyme works as intended. For the positive control, a larger band is observed when undigested (lane 1), but a smaller band/s are observed with digestion. The same pattern is observed with NAT1/NAT9cI, showing that the presence of inosines does not negate or interfere with digestion. A small amount of undigested product seems to remain with NAT1/NAT9cI, suggesting that the cleavage is not as effective, at least under these conditions, as with NAT1/NAT9c. Cleavage efficiency may be improved by altering buffer conditions and/or addition of more inosines at the 5' end of NAT9cI.

[0194] The foregoing example demonstrates that it is feasible to use a Topo / TypeIIS restriction enzyme combination to add a single nucleotide to the 5' end of a target single stranded DNA. A related topoisomerase, SVF, that recognizes the sequence CCCTG (<https://www.ncbi.nlm.nih.gov/pubmed/8661446>) is used to add a 'G' instead of a 'T', using an analogous process, thus allowing construction of a sequence encoding binary information with T and G.

[0195] As noted above, where dsDNA is generated using topoisomerase strategies, nicks in DNA on the opposing strand can be repaired using a ligase together with ATP. But when doing the single nucleotide addition, as in this example, we are building a single stranded DNA, so there are no nicks that need to be repaired and no need to use ligase.

Example 6 – Addition of single base using topoisomerase strategy couple with 5' phosphate coupling

[0196] In another approach to single base addition, we use a 5'phosphate as a blocking group to provide single base pair addition in the 3' to 5' direction. The charging reaction charges the topoisomerase with a single T (or G, or other nucleotide as desired), having a 5' phosphate group. When the charged topoisomerase 'sees' a free 5' unblocked (unphosphorylated) single stranded DNA chain it will add the T to that chain, providing a DNA with a T added to the 5'. This addition is facilitated by the presence of an adapter DNA having sequences to which the topoisomerase and the single stranded acceptor DNA can bind. (Note that the adapter DNA is catalytic – it can be reused as a template in repeated reactions.) The added nucleotide has a 5' phosphate on it, so it won't be a substrate for further addition until it is exposed to a phosphatase, which removes the 5' phosphate. The process is repeated, using Topo to add a single "T" to the 5' end of a target single stranded DNA and SVF topoisomerase to add a single 'G', thus allowing

construction of a sequence encoding binary information with T and G. The process is depicted schematically as follows:

*****ALTERNATE TRANSFER MECHANISM*****

```

T-TOPO  (T is 5' phosphorylated)

+

N-N-N-N-N-N-N-N-N-N-N  (acceptor) (5' N has 5' OH)

+

N-N-N-N-N-N-N-C-C-C-T
N-N-N-N-N-N-N-N-N-N-N-A-I-I-I-I-I
(adapter)

=

TOPO

+

N-N-N-N-N-N-N-C-C-C-T T-N-N-N-N-N-N-N-N-N-N-N
N-N-N-N-N-N-N-N-N-N-N-A-I-I-I-I-I
    this transient intermediate that breaks down to -->

T-N-N-N-N-N-N-N-N-N-N-N  (T is 5' phosphorylated)

+

N-N-N-N-N-N-N-C-C-C-T
N-N-N-N-N-N-N-N-N-N-N-A-I-I-I-I-I

```

Example 7 – Using DNA origami to aid in attaching DNA adjacent to nanopore

[0197] A DNA strand with a large origami structure on one end is captured in a nanopore, and immobilized to surface-conjugated streptavidin through a terminal biotin moiety on the DNA. After restriction enzyme cleavage of the origami structure, the immobilized DNA can be moved back and forth through the pore, as confirmed by current disruption. The immobilization enables a controlled movement of a single DNA molecule through the pore, which in turn enables both the ‘reading’ and ‘writing’ of information to DNA.

[0198] As depicted in Figure 35, a bulky double-stranded DNA unit is formed, which is too large to fit through the nanopore, with a single stranded region, linked to the bulky portion by two short double stranded regions having which serves to anchor the DNA to be added to in the

synthesis. The single stranded region can then be detached and anchored to the surface adjacent to the nanopore, and the origami structure released. See Figure 33.

[0199] Nanopores are formed in 3mm chips with 20nm SiO₂, with 50 * 50 µm windows. Chip are provided by Nanopore Solutions. Nanopore cassette holders and flow cells are provided by Nanopore Solutions. The amplifier is a Tecella Pico 2 amplifier. This is a usb-powered amplifier that uses a usb-computer interface for control. Tecella supplies (Windows) software to control the amplifier. The multimeter is a FLUKE 17B+ Digital Multimeter, capable of detecting current as low as 0.1 uA. For screening of radiofrequency noise we use a Concentric Technology Solutions TC-5916A Shield Box (Faraday Cage) with USB interface. Oligonucleotides are obtained from IDT.com. "PS" is Propargyl Silane - O-(PROPARGYL)-N-(TRIETHOXSILYLPROPYL) CARBAMATE from <http://www.gelest.com/product/o-propargyl-n-triethoxysilylpropylcarbamate-90/>.

[0200] The origami structure is based on single-stranded m13 with a 'honeycomb' cube origami structure which is ~ 20nm on one side. There are double stranded regions adjacent to the honeycomb each containing a unique restriction site. One of those sites is used to attach modified DNA to enable attachment near the nanopore, the other is used for cleaving off the origami structure once the DNA is attached.

[0201] *Nanopore Formation:* Nanopores are formed in the chips using dielectric breakdown, as follows:

1. Chips are carefully mounted in the cassettes
2. Wetting: 100% ethanol is carefully pipetted on the chip. Bubbles must be removed. However, direct pipetting of solution on the chip should be avoided or the chip can crack (SiO₂ is only 20nm).
3. Surface treatment: ethanol is removed, and freshly prepared Piranha solution (75% sulfuric acid, 25% hydrogen peroxide (30%)) is pipetted onto the chip. (let piranha solution come to room temperature). Leave on for 30 minutes.
4. Rinse 4 times with distilled water.
5. Rinse 2 times with HK buffer (10mM HEPES pH 8, 1M KCl)
6. Assemble cassette into flow cell.
7. Add 700 µL HK buffer to each chamber of the flow cell.
8. Insert silver electrodes attached to the amplifier and close the Faraday cage.

9. Test resistance with 300 mV. No current should be detected. If it is detected, the chip is likely cracked and one must start again.

10. Connect electrodes to a DC current of 6 V and test the current with a multimeter. Current should be low and should not change. Increase voltage by 1.5 V and hold the voltage until the resistance increases. If resistance does not increase after 5-10 minutes, increase the voltage another 1.5 V and try again. Repeat until resistance increases, at which point the applied voltage should be stopped immediately. (with sufficient voltage, dielectric breakdown occurs and a 'hole' is created in the SiO₂ membrane. When initially created the hole is small, but will increase in size as the voltage is maintained.)

11. Test the pore using the amplifier. At 300 mV one should see current of a few to several nA. The more current, the larger the pore.

[0202] Figure 34 depicts a basic functioning nanopore. In each panel, the y-axis is current (nA) and the x-axis is time (s). The left panel "Screening of RF Noise" illustrates the utility of the Faraday cage. A chip with no nanopore is placed in the flow cell and 300mV applied. When the lid of the Faraday cage is closed (first arrow) the noise reduction can be seen. A small spike occurs when the latch is closed (second arrow). Notice the current is ~0 nA. After pore manufacture (middle panel), application of 300mV (arrow) results in a current of ~3.5 nA. When DNA is applied to the ground chamber and +300mV is applied DNA translocations (right panel) can be observed as transient decreases in the current. (Note, in this case the TS buffer is used: 50mM Tris, pH 8, 1M NaCl). Lambda DNA is used for this DNA translocation experiment.

[0203] *Silver Chloride Electrodes:*

1. Silver wire is soldered to insulated copper wire.
2. Copper wire is grounded, and silver is dipped into fresh 30% sodium hypochlorite for 30 minutes.
3. Silver should acquire a dark gray coating (silver chloride).
4. Silver wire is rinsed extensively in distilled water and dried.
5. It is now ready for use.

[0204] *Silanization of Beads:* The silanization method is initially developed/tested on SiO₂ coated magnetic beads (GBioscience). The following protocol is adopted:

1. Pretreat beads in fresh pirannah solution for 30 minutes.

2. Wash 3x with distilled water.
3. Wash 2x in methanol.
4. Dilute APTES stock 1:500 in methanol.
5. Add diluted APTES to beads, incubate at RT for 45 minutes.
6. Rinse with methanol.
7. 100° C for 30 minutes.
8. Store under vacuum.

[0205] *Silanization of silicon chip*

1. Mount chip with nanopore in a cassette.
2. Rinse with methanol, carefully removing any air bubbles.
3. Add fresh pirannah solution (equilibrated to room temperature) and incubate for 30 minutes.
4. Wash 4x with distilled water.
5. Wash 3x with methanol.
6. Dilute APTES stock 1:500 in methanol and use to wash chip 2x. Incubate at RT for 45 minutes.
7. Rinse 2x with methanol.
8. Dry under and air stream.
9. Store under vacuum overnight.

[0206] *Streptavidin conjugation of beads:* The streptavidin conjugation is initially developed/tested on the silanized beads prepared above.

1. Wash silanized beads with Modified Phosphate-Buffered Saline (MPBS)
2. Make a fresh solution of 1.25% glutaraldehyde in MPBS (using 50% glutaraldehyde stock, stored frozen).
3. Add 1.25% glutaraldehyde to beads and let stand for 60' with gentle up-down pipetting every 15 minutes.
4. Wash 2x with MPBS.
5. Wash 2x with water.
6. Let dry under vacuum.

7. Add streptavidin (500 µg/mL in MPBS) to beads and incubate 60 minutes. (For negative control beads, use bovine serum albumin (BSA) (2mg/mL in MPBS) in place of streptavidin).
8. Remove streptavidin and add BSA (2mg/mL in MPBS). Incubate 60 minutes.
9. Wash 2x in MPBS.
10. Store at 4° C.

[0207] *Streptavidin conjugation of silicon chip*

1. Rinse silanized chip with ethanol 2x
2. Rinse silanized chip with MPBS 2x
3. Make a fresh solution of 1.25% glutaraldehyde in MPBS (using 50% glutaraldehyde stock, stored frozen).
4. Rinse chip with 1.25% glutaraldehyde 2x, let stand for 60' with gentle up-down pipetting every 15 minutes
5. Wash 2x with MPBS
6. Wash 2x with water
7. Let dry under air stream
8. To one half of the chip add BSA (2 mg/mL in MPBS), and to the other add streptavidin (500 µg/mL in MPBS). Incubate 60 minutes. Make a marking on the cassette to indicate which half of the chip is streptavidin modified.
9. Rinse both halves of the chip with BSA (2 mg/mL in MPBS). Incubate 60 minutes.
10. Wash in MPBS.

[0208] The buffers used herein are made as follows:

MPBS: 8 g/L NaCl, 0.2 g/L KCl, 1.44 g/L disodium-phosphate, 0.240 g/L potassium phosphate, 0.2 g/L polysorbate-20 (pH 7.2)

PBS: 8 g/L NaCl, 0.2 g/L KCl, 1.44 g/L disodium-phosphate, 0.240 g/L potassium phosphate

TS: 50 mM Tris pH 8.0, 1M NaCl

HK: 10 mM HEPES pH 8.0, 1M KCl

TE: 10 mM Tris, 1 mM EDTA, pH 8.0

Pirannah Solution: 75% hydrogen peroxide (30%) + 25% sulfuric acid

APTES stock: 50% methanol, 47.5% APTES, 2.5% nanopure water. Age at 4°C for at

least 1 hour. Store at 4° C.

PDC stock: 0.5% w/v 1,4-phenylene diisothiocyanate in DMSO

[0209] Oligonucleotides (5' TO 3') are ordered:

o1 CTGGAACGGTAAATTCAGAGACTGCGCTTTCATTCTGGCTTTAATG

o3 GGAAAGCGCAGTCTCTGAATTTAC

N1 CTTACTGGAACGGCTATCGATATCGCAGCAGGACAGA

BN1 Biotin-CTTACTGGAACGGCTATCGATATCGCAGCAGGACAGA

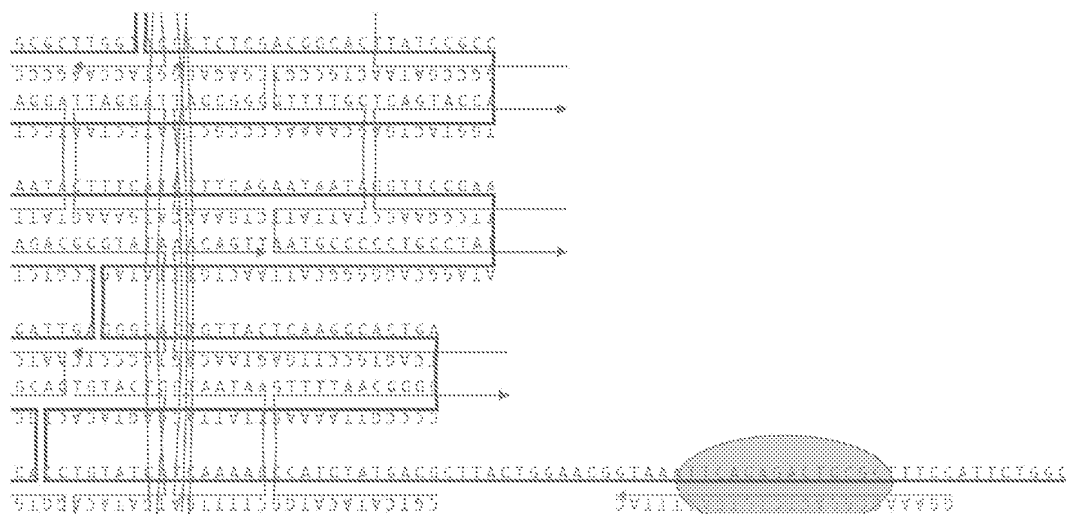
N2 GTCCTGCTGCGATATCGATAGCCGTTCCAGTAAG

[0210] Oligonucleotide pair hybridization is carried out as follows:

1. Make stock solutions of oligo's at 100uM concentration in TE buffer
2. Dilute oligos to 10 μ M in PBS
3. Heat to 85° C for 5' in a thermal cycler
4. Ramp heat down by 5° C every 3' until 25° C
5. Store at 4° C or -20° C

[0211] *Streptavidin Conjugation:* Streptavidin conjugation to SiO₂ is developed and tested using SiO₂ coated magnetic beads, and the protocols were then adapted for SiO₂ chips. Binding of biotinylated oligos to both streptavidin and BSA conjugated beads are tested. As expected, negligible binding is observed with BSA-conjugated beads, while strong binding is observed with streptavidin conjugated beads. See Figure 38. Since it would be more convenient to perform the conjugation in high salt (DNA movement is performed in high salt), the ability of the beads to bind in HK buffer is also tested. Binding in HK buffer is comparable to binding in MPBS buffer (Figure 39).

[0212] Origami constructs are made and confirmed operable as described above in Figures 35-37. Biotinylation of the origami structure is tested using oligonucleotides. We already know from the 'Origami' results described above for Figure 37 that the AlwNI site is active. An oligonucleotide pair that recreates a segment of the exact sequence in the origami DNA is used below (o1/o3). The origami molecule is depicted as follows:



[0213] The oligo pair o1/o3 is

CTGGAACGGTAAATTTCAGAGACTGCGCTTTCCATTCTGGCTTTAATG o1
 CATTTAAGTCTCTGACGCGAAAGG o3

[0214] The DNA is digested with AlwNI in the presence of T4 DNA ligase, and a biotinylated oligo that is complementary to the overhang on the 3' side of the origami sequence (which itself is attached to a long ssDNA sequence which itself is attached to the other side of the origami), according to the following reaction:

```

CTGGAACGGTAAATTCAGAGACTGCGCTTTCCATTCTGGCTTTAATG   o1
      CATTTAAGGTCTCTGACGCGAAAGG                           o3
          + AlwNI
          ...
CTGGAACGGTAAATTCAGAGA          CTGCGCTTTCCATTCTGGCTTTAATG
      CATTTAAGTC          TCTGACGCGAAAGG
          +
B-CTTACTGGAACGGCTATCGATATCGCAGCAGGACAGA          BN1
      GAATGACCTTGCCGATAGCTATAGCGTCGTCCTG          N2
          + ligase
          ...
B-CTTACTGGAACGGCTATCGATATCGCAGCAGGACAGACTGCGCTTTCCATTCTGGCTTTAATG
      GAATGACCTTGCCGATAGCTATAGCGTCGTCCTGTCTGACGCGAAAGG

```

In this strategy, AlwNI cleaves the target DNA. When the ligase is added it is possible for this DNA to be religated, but the restriction enzyme will cut it again. However, if/when the (right) fragment (of o1/o3) binds to BN1/N2, the restriction site is NOT recreated, thus this product will not be cut. Specific attachment is confirmed, by testing with and without the restriction enzyme:

	1	2
o1/o3	1	1
n1/bn2	1	1
10x lig buf	2	2
water	15	14.5
AlwNI		.5
Ligase	1	1

All reagents except ligase are added and solution is incubated at 37° C for 60 minutes. Ligase is added and solution incubated overnight at 16° C. 10x lig buff refers to NEB 10x T4 DNA ligase buffer. Ligase in NEB T4 DNA ligase. o1/o3 and n1/n2 refer to annealed oligo pairs, as depicted above. Units are microliters. Agarose gel analysis confirms that in the presence of the AlwNI, a larger product is formed, corresponding to the biotinylated oligonucleotide attached to the long ssDNA arm attached to the origami structure. A similar strategy is used for 3' biotinylation, where desired.

[0215] Above we demonstrate the ability to form and use a nanopore to detect voltage induced transit of DNA across the pore, the creation of an origami molecule with a long ss region attached at its' far end to a biotin, and the conjugation of streptavidin to silicon dioxide, and to use that to capture biotinylated DNA. These tools are used to attach and control the movement of a single DNA molecule next to a nanopore.

[0216] The first step is to conjugate streptavidin to one surface of an SiO₂ nanopore (and BSA to the other side). This is accomplished according to the protocol above. The resulting pores tend to have a lower current than they initially have. After some brief 6v pulses, the currents return to be near their original current. A functioning nanopore at this point is shown in Figure 40.

[0217] Next, the origami DNA is inserted. When the origami DNA is added to the appropriate chamber and the current turned on, the origami will insert into the chamber. A representation of this is shown in Figure 41. Experimental results when the origami is introduced at a final concentration of 50 pM confirm that the DNA with the origami inserts into the pore relatively soon (typically in seconds), which is detectable by the resulting reduction of current flow across the nanopore (e.g., in these experiments, current before origami insertion is ~3nA, and ~2.5nA after insertion). If the current is allowed to run for longer times, double insertions can be observed. If higher concentrations are used, insertion occurs too quickly to be observed.

[0218] *Binding of inserted DNA to chip.* After the origami is inserted into the nanopore, 15 minutes are allowed to elapse before voltage is applied again. The end of the ssDNA region of the origami contains a biotin, and streptavidin is conjugated to the surface of the nanopore. Streptavidin binds to avidin with an affinity constant that approaches that of a covalent bond. The 15 minute time allows the DNA to diffuse and for the biotin end to find and bind to the streptavidin. If the DNA has in fact become attached to the surface, when the voltage is reversed the observed current should be slightly less than what was seen previously. Also, switching the current back and forth should result in currents in both directions that are lower than that seen with a free pore.. In the example shown here, a free pore shows a current of ~3nA. Figure 42 shows a representation of attached DNA, and Figure 43 shows experimental results of voltage switching the attached origami DNA. Note that the currents seen in both directions are ~±2.5 nA, which is lower than the ~±3nA observed with a free pore. If the DNA hasn't bound to the surface, the original current will be recovered when the voltage is switched (Figure 44).

[0219] In order to remove the origami structure, the buffer in the flow cell chamber containing the origami structure was removed and replaced with 1x Swa1 buffer with 1uL Swa1/20 °L. The buffer in the other flow cell chamber is replaced with 1x Swa1 buffer without Swa1. This is incubated at room temperature for 60 minutes, then washed with HK buffer, and voltage applied. Movement of the DNA back and forth as represented in Figure 45 is confirmed by the experimental data in Figure 46, showing controlled movement of immobilized DNA through a SiO₂ nanopore.

Example 8 – Alternative means to attach the polymer to the surface adjacent to the nanopore

[0220] The foregoing examples describe attachment of DNA to the surface adjacent to the nanopore by biotinylating the DNA and coating the attachment surface with streptavidin. Some alternative means of attaching the polymer are depicted in Figure 47.

[0221] a) *DNA Hybridization*: In one method, the DNA which is extended in the methods of the invention is hybridized to a short oligonucleotide which is attached near the nanopore. Once the synthesis is complete, the synthesized DNA can be easily removed without a need for restriction enzymes, or alternatively the double strand formed by the bound oligonucleotide and the synthesized DNA can provide a substrate for a restriction enzyme. In this example, the short oligomers are conjugated to the surface using biotin-streptavidin, or ligated using 1,4-phenylene diisothiocyanate as follows:

Conjugation of biotinylated DNA to SiO₂:

A. SILANIZE:

1. Pre-treatment: nha solution for 30 minutes, wash with double distilled H₂O (ddH₂O)
2. Prepare APTES stock: 50% MeOH, 47.5% APTES, 2.5% nanopure H₂O: age >1hr 4°C
3. dilute APTES stock 1:500 in MeOH
4. incubate chips at room temperature
5. rinse MeOH
6. dry
7. heat at 110° C for 30 minutes

CONJUGATE:

1. Treat chip with PDC stock 5h (room temperature) (PDC stock: 0.5% w/v 1,4-phenylene diisothiocyanate in DMSO)

2. 2 washes in DMSO (brief!)
3. 2 washes in ddH₂O (brief!)
4. 100 nM amino-modified DNA in ddH₂O (pH 8) O/N 37° C
5. 2 washes 28% ammonia solution (deactivate)
6. 2 washes ddH₂O

[0222] Single stranded DNA having a terminal sequence complementary to the attached oligonucleotides is introduced as described above and allowed to hybridize with the attached oligonucleotides.

[0223] b) *Click chemistry*: Click chemistry is a general term for reactions that are simple and thermodynamically efficient, do not create toxic or highly reactive byproducts, and operate in water or biocompatible solvents, and are often used to join substrates of choice with specific biomolecules. The click conjugation in this case uses similar chemistry as used in a) to attach the oligonucleotides, only it is here used to attach the polymer which is extended in the course of synthesis in the methods of the invention. While in this example, DNA is the polymer, this chemistry would work to attach other polymers which have been functionalized by addition of a compatible azide group.

SILANIZE:

1. Pre-treatment: piranha solution for 30', wash ddH₂O
2. Prepare PS (propargyl silane) stock: 50% MeOH, 47.5% PS, 2.5% nanopure H₂O:
age >1hr 4C
3. Dilute APTES stock 1:500 in MeOH
4. Incubate chips at room temperature
5. Rinse MeOH
6. Dry
7. Heat at 110° C for 30 minutes

[0224] DNA which is terminated in an azide functional group will covalently bind to this surface (as shown in Figure 47). Azide terminated oligos are ordered, and attached to the longer origami DNA, as described for the biotin addition to the DNA previously.