Real-time single-molecule electronic DNA sequencing by synthesis using polymer-tagged nucleotides on a nanopore array

Carl W. Fuller[a,b,c], Shiv Kumar[a,b], Mintu Porel[a,b], Minchen Chien[a,b], Arek Bibillo[c], P. Benjamin Stranges[d], Michael Dorwart[c], Chuanjuan Tao[a,b], Zengmin Li[a,b], Wenjing Guo[a,b], Shundi Shi[a,b], Daniel Korenblum[c], Andrew Trans[c], Anne Aguirre[c], Edward Liu[c], Eric T. Harada[c], James Pollard[c], Ashwini Bhat[c], Cynthia Cech[c], Alexander Yang[c], Cleoma Arnold[c], Mirkó Palla[d], Jennifer Hovis[c], Roger Chen[c], Irina Morozova[a,b], Sergey Kalachikov[a,b], James J. Russo[a,b], John J. Kasianowicz[e], Randy Davis[c], Stefan Roever[c], George M. Church[d,1], and Jingyue Ju[a,b,f,1]

[a]Center for Genome Technology and Biomolecular Engineering, Columbia University, New York, NY 10027; [b]Department of Chemical Engineering, Columbia University, New York, NY 10027; [c]Genia Technologies, Santa Clara, CA 95050; [d]Department of Genetics, Harvard Medical School, Boston, MA 02115; [e]Semiconductor and Dimension Metrology Division, National Institute of Standards and Technology, Gaithersburg, MD 20899; and [f]Department of Pharmacology, Columbia University College of Physicians and Surgeons, New York, NY 10032

Contributed by George M. Church, March 18, 2016 (sent for review February 4, 2016; reviewed by Cynthia J. Burrows, Xiaohua Huang, and Andre Marziali)

DNA sequencing by synthesis (SBS) offers a robust platform to decipher nucleic acid sequences. Recently, we reported a single-molecule nanopore-based SBS strategy that accurately distinguishes four bases by electronically detecting and differentiating four different polymer tags attached to the 5′-phosphate of the nucleotides during their incorporation into a growing DNA strand catalyzed by DNA polymerase. Further developing this approach, we report here the use of nucleotides tagged at the terminal phosphate with oligonucleotide-based polymers to perform nanopore SBS on an α-hemolysin nanopore array platform. We designed and synthesized several polymer-tagged nucleotides using tags that produce different electrical current blockade levels and verified they are active substrates for DNA polymerase. A highly processive DNA polymerase was conjugated to the nanopore, and the conjugates were complexed with primer/template DNA and inserted into lipid bilayers over individually addressable electrodes of the nanopore chip. When an incoming complementary-tagged nucleotide forms a tight ternary complex with the primer/template and polymerase, the tag enters the pore, and the current blockade level is measured. The levels displayed by the four nucleotides tagged with four different polymers captured in the nanopore in such ternary complexes were clearly distinguishable and sequence-specific, enabling continuous sequence determination during the polymerase reaction. Thus, real-time single-molecule electronic DNA sequencing data with single-base resolution were obtained. The use of these polymer-tagged nucleotides, combined with polymerase tethering to nanopores and multiplexed nanopore sensors, should lead to new high-throughput sequencing methods.

single-molecule sequencing | nanopore | DNA sequencing by synthesis | polymer-tagged nucleotides | chip array

The importance of DNA sequencing has increased dramatically from its inception four decades ago. It is recognized as a crucial technology for most areas of biology and medicine and as the underpinning for the new paradigm of personalized and precision medicine. Information on individuals' genomes and epigenomes can help reveal their propensity for disease, clinical prognosis, and response to therapeutics, but routine application of genome sequencing in medicine will require comprehensive data delivered in a timely and cost-effective manner (1). Although 35 years of technological advances have improved sequence throughput and have reduced costs exponentially, genome analysis still takes several days and thousands of dollars to complete (1, 2). To realize the potential of personalized medicine fully, the speed and cost of sequencing must be brought down another order of magnitude while increasing sequencing accuracy and read length. Single-molecule approaches are thought to be essential to meet these requirements and offer the additional benefit of eliminating amplification bias (3, 4). Although optical methods for single-molecule sequencing have been achieved and commercialized, the most successful, Pacific Biosciences' single molecule real-time (SMRT) sequencing by synthesis (SBS) approach, requires expensive instrumentation and the use of fluorescently tagged nucleotides (4, 5).

In the last two decades, there has been great interest in taking advantage of nanopores, naturally occurring or solid-state ion channels, for polymer characterization and distinguishing the bases of DNA in a low-cost, rapid, single-molecule manner (6–9). Three nanopore sequencing approaches have been pursued: strand sequencing in which the bases of DNA are identified as they pass sequentially through a nanopore (6, 7), exonuclease-based nanopore

Significance

Efficient cost-effective single-molecule sequencing platforms will facilitate deciphering complete genome sequences, determining haplotypes, and identifying alternatively spliced mRNAs. We demonstrate a single-molecule nanopore-based sequencing by synthesis approach that accurately distinguishes four DNA bases by electronically detecting and differentiating four different polymer tags attached to the terminal phosphate of the nucleotides during their incorporation into a growing DNA strand in the polymerase reaction. With nanopore detection, the distinct polymer tags are much easier to differentiate than natural nucleotides. After tag release, growing DNA chains consist of natural nucleotides allowing long reads. Sequencing is realized on an electronic chip containing an array of independently addressable electrodes, each with a single polymerase–nanopore complex, potentially offering the high throughput required for precision medicine.

Author contributions: C.W.F., S. Kumar, A. Bibillo, R.C., S. Kalachikov, J.J.K., R.D., S.R., G.M.C., and J.J. designed research; C.W.F., S. Kumar, M. Porel, M.C., A. Bibillo, P.B.S., M.D., C.T., Z.L., W.G., S.S., A.T., A.A., E.L., E.T.H., J.P., A. Bhat, C.C., A.Y., C.A., M. Palla, J.H., R.C., S. Kalachikov, and R.D. performed research; S. Kumar, A. Bibillo, P.B.S., M.D., Z.L., D.K., R.C., and S. Kalachikov contributed new reagents/analytic tools; C.W.F., D.K., A.T., I.M., and J.J.R. analyzed data; and C.W.F., S. Kumar, M.C., I.M., S. Kalachikov, J.J.R., J.J.K., and J.J. wrote the paper.

Reviewers: C.J.B., University of Utah; X.H., University of California, San Diego; and A.M., University of British Columbia.

Conflict of interest statement: The nanopore SBS technology has been exclusively licensed by Genia. In accordance with the policy of Columbia University and the National Institute of Standards and Technology, the coinventors (S. Kumar, M.C., C.T., Z.L., S. Kalachikov, J.J.R., J.J.K., and J.J.) are entitled to royalties through this license. G.M.C. is a member of the Scientific Advisory Board of Genia; other potential conflicts are described at arep.med.harvard.edu/gmc/tech.html.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. Email: gchurch@genetics.med.harvard.edu or dj222@columbia.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1601782113/-/DCSupplemental.

sequencing in which nucleotides are enzymatically cleaved one-by-one from a DNA molecule and monitored as they are captured by and pass through the nanopore (10), and a nanopore SBS approach, described by us, in which identifiable polymer tags are attached to nucleotides and registered in nanopores during enzyme-catalyzed DNA synthesis (11). Common to all these methods is the need for precise control of the reaction rates so that each base is determined in order and without fail. Strand sequencing requires a method for slowing down the passage of the DNA through the nanopore and decoding a plurality of bases within the channel; complex ratcheting approaches, taking advantage of molecular motors, have been developed for this purpose (12, 13) and were used recently to assemble the sequence of phiX174 (14). Exonuclease-based sequencing requires the release of each nucleotide close enough to the pore to guarantee its capture and its transit through the pore at a rate slow enough to obtain a valid ionic current signal (10, 15). In addition, both of these methods rely on distinctions among the four natural bases, two relatively similar purines and two similar pyrimidines. Our nanopore SBS approach has the advantage of using synthetic polymer tags attached to the nucleotides that are designed specifically to produce unique and readily distinguishable ionic current blockade signatures for sequence determination (11). Direct attachment of the polymerase to the nanopore ensures that there is ample time for the polymer tags on the incoming nucleotide to be captured in the nanopore before completion of the DNA polymerase catalytic cycle. The rates of exonucleases are sensitive to sequence context and secondary structure (10, 15–17). In contrast, polymerases have less variable rates than exonucleases and therefore will yield more consistent base identification in nanopore SBS.

We recently demonstrated proof-of-principle for the nanopore SBS sequencing method (11), which combines SBS (18–20) with nanopore-based identification of different-sized polymer tags (8, 9) attached to the nucleotides. One of four different-length PEG tags was attached to the terminal phosphate of each nucleotide. Despite having long tags with 16–36 PEG monomer units, these tagged nucleotides were incorporated efficiently by DNA polymerase. During the phosphoryl transfer step of the DNA polymerase reaction, the tag is released as part of the polyphosphate byproduct, so only the natural nucleotide remains in the growing DNA strand. This tag was detected and identified by monitoring pore current as it passed through a single-protein nanopore (α-hemolysin, hereafter αHL) embedded in a lipid membrane under a voltage gradient. Depending on the length of the PEG tag, the pore current was reduced to different levels, and translocation required different times (11), allowing discrimination of such tags and enabling the identification of each nucleotide incorporated in the SBS process.

The measurements with PEG tags were made using high salt concentrations, which are not ideal for polymerase reactions. To develop the nanopore SBS approach further and to optimize the tags, we report here the design and synthesis of nucleotides tagged with modified oligonucleotides and their application for nanopore SBS. These tags have structural modifications that create distinct current blockades, measured using an electronic chip-based array of nanopores embedded in lipid bilayer membranes. The tags are attached to the terminal phosphate of 2′-deoxynucleoside-5′-hexaphosphates using Huisgen cycloaddition azide/alkyne coupling chemistry (21). Tagged nucleoside hexaphosphates were chosen because they generally have better activity with DNA polymerases than tagged nucleoside triphosphates (22, 23). With these tagged nucleotides, we demonstrate continuous single-molecule electronic DNA sequencing with single-base resolution by nanopore SBS. The measurement of current is made during the polymerase catalytic cycle when the complementary tagged nucleotide is bound within the complex of DNA polymerase, primer/template, and divalent metal cation and lasts until the completion of the polymerase catalytic step with the release of the tagged polyphosphate product. Once this product is released, the polymer tag is free to leave the pore, ending the blockade signal for that particular DNA synthesis step. To increase the likelihood that each tag will be measured in sequential order, a single polymerase molecule is covalently attached to the nanopore at an appropriate distance to allow fast capture of the tag by the nanopore (Fig. 1*A*). Each of the four tags has a distinctive structure that interacts with the narrowest constriction in the αHL channel, thereby reducing the ionic current across the channel to different extents (24, 25).

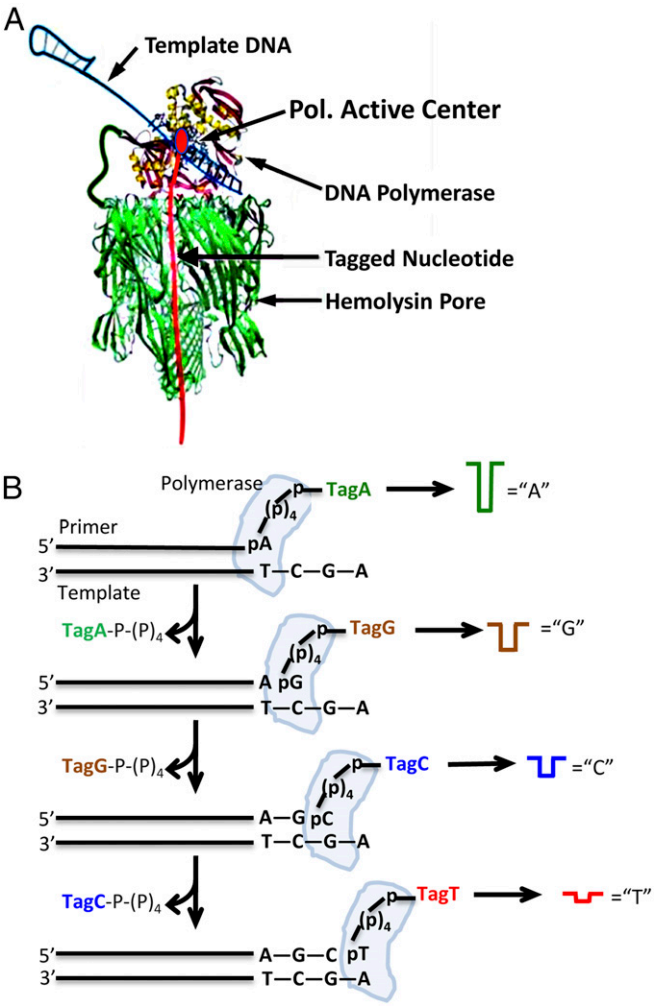

Fig. 1. Principle of nanopore SBS. (*A*) Nanopore–polymerase sequencing engine. A single DNA polymerase molecule is covalently attached to an αHL nanopore heptamer. Primer and template DNA (shown as a double-hairpin conformation) along with the tagged nucleotide form a ternary complex with the polymerase. (*B*) SBS schematic showing the sequential capture and detection of tagged nucleotides by the nanopore as they are incorporated into the growing DNA strand in the polymerase reaction to produce nucleotide-specific current blockades.

Results and Discussion

We previously demonstrated the ability to identify PEG tags of different lengths by their distinct blockade currents in membranes traversed by a single αHL nanopore (8, 11). The instrumentation and testing method used made it possible to record rapid, submillisecond deviations from open current readings (9). Here we used a complementary metal-oxide semiconductor (CMOS) chip containing an independently addressable array of 264 electrodes supporting independent bilayers. In the prototype chip, there was a minimum measurement period of ∼1 ms, which required a revised approach for tag detection.

Fig. 1*A* shows the nanopore–polymerase sequencing engine with a single Phi29 DNA polymerase molecule (26) covalently attached to αHL (27). Polymerase binds primer and template DNA, along with the complementary tagged nucleotide, allowing placement of the tag within the nanopore. The schematic in Fig. 1*B* shows the sequential capture and detection of tagged nucleotides by the nanopore as they are being incorporated into the growing primer strand during the polymerase reaction. Preliminary experiments using streptavidin-bound 3′-biotinylated oligonucleotides following published procedures (25) confirmed that it is possible to capture and hold oligonucleotides in chip-based pores and to measure current blockade levels over a period of ∼50–2,000 ms or longer. A survey of oligonucleotide tags, all based on modifications of oligo dT_{30}, was conducted, yielding the four tags shown in Fig. S1 that generate distinguishable current blockade signatures. For sequencing we decided to mimic the experiment that generates the current blockade data shown in Fig. S1 by replacing the streptavidin–tag complex with a DNA polymerase ternary complex that holds the tagged nucleotides (Fig. 2), allowing the monitoring of current while the tags are still attached to the nucleotides. Thus, detection of the tag occurs before nucleotide incorporation by phosphoryl transfer, rather than after the tags are detached from their nucleotides to pass through the nanopore channel. The tag detection can be accomplished if the rate of the nucleotide incorporation is slower than the combined rate of nucleotide tag capture by the nanopore and current blockade measurement. The use of appropriate polymerase mutants and refined reaction conditions can increase the time the tagged nucleotide is present in the ternary complex, and close tethering, via covalent attachment, of the polymerase to the nanopore can result in rapid tag captures on the order of microseconds.

Synthesis and Testing of Tagged Nucleotides for Nanopore SBS Reactions. We designed and synthesized a wide variety of tagged nucleotides. After characterization of their properties and performance, we selected a subset of four tagged nucleotides that produce distinct and reproducible reduction in nanopore ionic currents with minimum background. The structures of these tagged nucleotides are shown in Fig. 2. The synthetic strategy is to attach an azide group via a chemical linker to the terminal phosphate of the

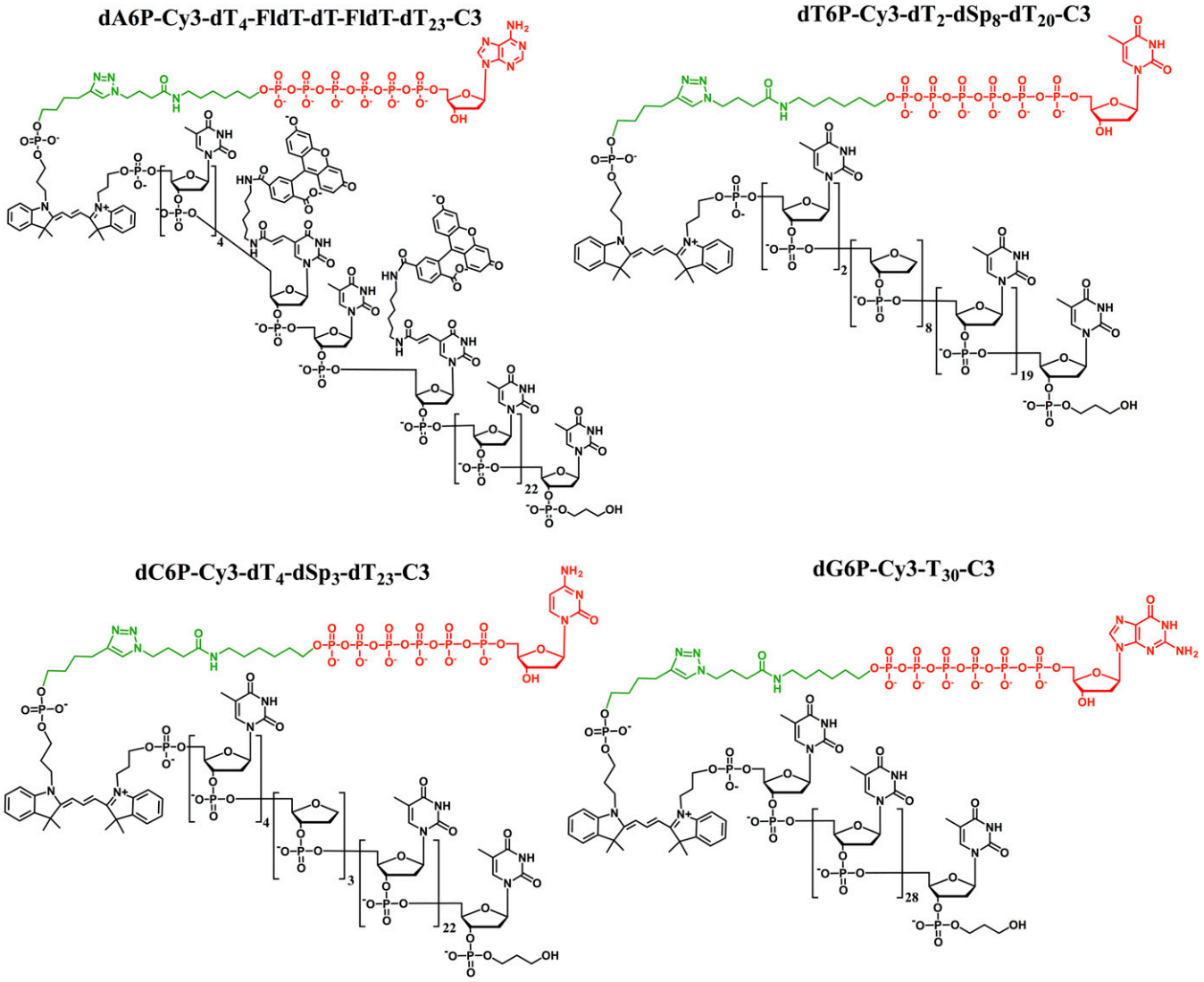

Fig. 2. Structures of the four polymer-tagged nucleotides. Nucleotides used in this study are 5′-nucleoside hexaphosphates (red) connected to a common linker (green) and an oligonucleotide tag (black) consisting of Cy3 and either an unmodified oligonucleotide chain (dT_{30}) (in dG6P-Cy3-dT_{30}-C3) or oligonucleotides with a variety of modifications including runs of abasic nucleotides (dSp_8 and dSp_3) (in dT6P-Cy3-dT_2-dSp_8-dT_{20}-C3 and dC6P-Cy3-dT_4-dSp_3-dT_{23}-C3) or two thymidines modified with fluorescein (in dA6P-Cy3-dT_4-FldT-dT-FldT-dT_{23}-C3).

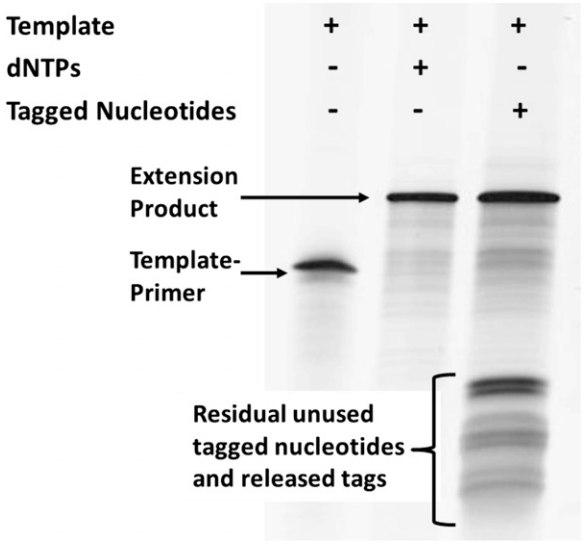

Fig. 3. Full-length polymerase extension products with dNTPs (center lane) or tagged nucleotides (right lane). Reactions were conducted with 3 pmol template (83 bases) and 75 pmol dNTPs or tagged dN6Ps, for 2–10 min at 37 °C, and the products were analyzed on a denaturing polyacrylamide gel.

deoxynucleoside-5′-hexaphosphate (dA6P, dC6P, dG6P, and dT6P, the dN6Ps) (Fig. S2). The tags were oligonucleotide-based polymers synthesized with a 5′-alkyne moiety that reacts readily with the desired azido-dN6P by azide-alkyne Huisgen cycloaddition (21, 28, 29) to produce the tagged nucleotides (Fig. 2). The tagged nucleotides were purified by HPLC, and their molecular weights were determined by MALDI-TOF MS (Fig. S3). The tags consisted of oligodeoxynucleotides (e.g., dT_{30}), some of which are interspersed with modified phosphodiester building blocks having specific base or backbone modifications. These modifications include bulky, charged dyes on the nucleotide base and runs of three or eight abasic (dSpacer, hereafter "dSp") sites (30, 31). Tags were protected at their 3′ end from exonuclease activity with 3′-phospho-propanol. The modifications in the dT_{30} tag sequence were positioned to be at or near the narrowest constriction in the αHL channel (24), where they can directly and specifically influence ionic current through the pore when the tagged nucleotide is held by polymerase in the ternary complex attached to the pore.

To demonstrate that the nucleotides modified with the polymeric tags were still good substrates for DNA polymerase, single-base extension reactions were performed using Therminator γ DNA polymerase. Primer–loop templates bearing complementary bases opposite the site of incorporation on the primer were used (11). The reaction products were desalted and subjected to MALDI-TOF MS analysis. As shown in Fig. S4, successful incorporations of the tagged nucleotides were revealed by the presence of new peaks with the expected mass of the primer-extension products and peaks indicative of the released pentaphosphate tag, along with peaks representing remaining reactants (primer, tagged nucleotide).

Additional experiments monitoring DNA synthesis on denaturing polyacrylamide gels show that primers can be extended by more than 40 bases using only tagged nucleotides and any of several DNA polymerases. In Fig. 3, we show the result of extension reactions using *Bst* 2.0 DNA polymerase (an in silico-designed homologue of *Bacillus stearothermophilus* DNA polymerase I, large fragment from New England Biolabs) at 37 °C, with reactants and products separated on a polyacrylamide-urea gel. The primer/template DNA (83-mer) is self-priming and can be extended by 47 nucleotides. Full-length products were obtained with either all four natural dNTPs or with the four tagged nucleotides depicted in Fig. 2.

Assembly of αHL–Polymerase Complexes and Insertion into Membranes on an Array Chip. To achieve continuous sequencing, we constructed a sequencing engine having a single DNA polymerase molecule attached directly to the heptameric αHL nanopore (Fig. 1*A*) using a recombinant technique to make two kinds of αHL monomers, one wild type and the other differing only in carrying a specific peptide domain (32) for attaching polymerase. These two types of monomers were mixed in appropriate ratios, allowing pore heptamers to self-assemble, followed by isolation of the heptamer pore carrying a single-peptide domain. When polymerase engineered to possess the mating protein fragment was added in excess to this heterologous pore, it reacted with the pore monomer tagged with the peptide domain, yielding the desired nanopore–polymerase conjugate (the procedure is described in *SI Methods*). During nanopore SBS, the polymerase attached to the pore tightly binds the DNA template/primer, completing a single structure that integrates itself into a lipid bilayer configured for monitoring transmembrane conductance and ready for sequencing DNA. When each complementary tagged nucleotide binds the polymerase–primer–template complex, its tag rapidly enters the nanopore channel in response to applied voltage to yield a unique current blockade that is recognized for sequence determination.

Sequencing experiments were performed using a CMOS nanopore chip (Fig. 4) that has an array of 264 Ag-AgCl electrodes (5-μm diameter) within shallow wells that support lipid bilayer membranes. The electrodes are individually addressable by a computer interface, allowing voltages to be applied only to the membranes in specific wells and thus effectively permitting independent sequence reads at these locations. All reagents are introduced into a simple flow cell above the chip using a computer-controlled syringe pump. Electrical measurements are made asynchronously at least once every millisecond and are recorded on the computer.

An 83-base synthetic primer/template, which self-anneals at both ends to prevent its entry into the nanopore, was added to the pore–polymerase complex and pumped over the membranes in a sensor chip flow cell in Hepes buffer (pH 7.5) containing 150 mM KCl and 3 mM $SrCl_2$. The applied voltage is adjusted to ensure that, in a majority of cases, one and only one pore is inserted into the membranes of each well. For preliminary experiments using only a single tagged nucleotide and the noncatalytic divalent metal ion (Sr^{2+}), the tagged nucleotide mix was added to the nanopore chip at room temperature, and current

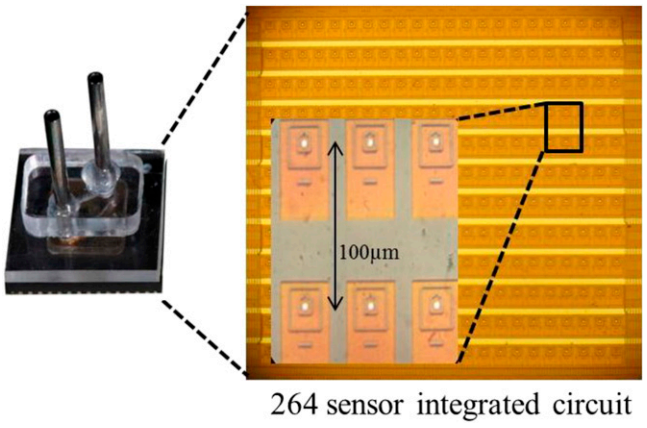

Fig. 4. CMOS chip for nanopore SBS. The CMOS chip (~1 × 1 mm) consists of an integrated circuit with an array of 264 silver electrodes (5-μm diameter). The chip supports analog-to-digital conversion, reporting electrical measurements from all electrodes independently at a rate of more than 1,000 per second. The flow cell (*Left*) surrounding the chip (*Right*) includes sample input and output ports for reagents. The *Inset* shows details of the architecture of the electrode array and individual sensing cells.

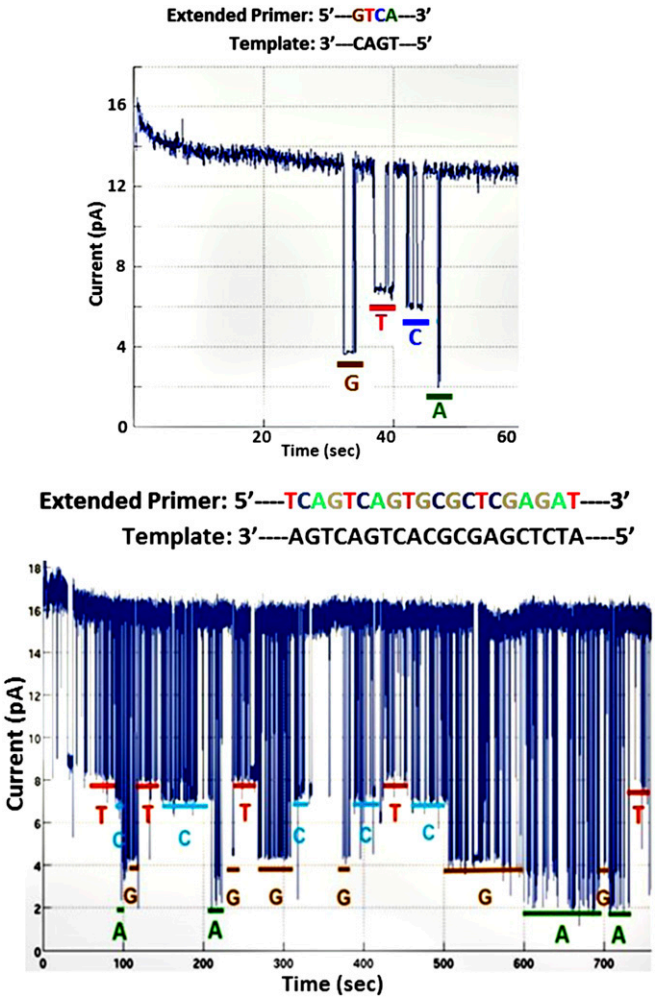

Fig. 5. Sequencing results on nanopore array chips. Sequencing reactions were performed with inserted αHL pores conjugated to a single Phi29 DNA polymerase molecule, template, and the four tagged nucleotides. (*Upper*) Four bases are clearly distinguished (3 mM $SrCl_2$ on the *trans* side and 0.1 mM $MnCl_2$ on the *cis* side of the membrane). (*Lower*) A longer sequence read (3 mM $SrCl_2$ *trans*, 3 mM $MgCl_2$, 0.7 mM $SrCl_2$ *cis*).

recording was started. These experiments established that binding of the tagged nucleotides with the polymerase–nanopore conjugate along with the primer and template to form a ternary complex is sequence-dependent and that the current levels for the four nucleotides are readily distinguished. Representative data for a chemically conjugated polymerase–pore complex are shown in Fig. S5*B* and Table S1.

Nanopore SBS. For sequencing experiments, a buffer solution containing 150 mM KCl, 3 mM $SrCl_2$, and either 0.1 mM $MnCl_2$ or 3 mM $MgCl_2$ (catalytic divalent cations) and 3 μM each of the four tagged nucleotides was added to the polymerase–nanopore chip constructed using the peptide-based conjugation method to start the SBS process. Ionic currents were measured by applying 80–120 mV across the membranes for 3–10 min. The sequencing data are shown in Fig. 5. Four distinct transient current levels substantially below the open-pore current level [~13 pA (Fig. 5, *Upper*) and ~16 pA (Fig. 5, *Lower*)] are observed, indicating pore captures of the tags associated with each nucleotide during each cycle of the polymerase reaction. These data are in accord with the levels observed when only a single tagged nucleotide and only noncatalytic divalent metal (Sr^{2+}) were used (Table S1). In these traces, the correct sequence order for four (Fig. 5, *Upper*) and 20 bases (Fig. 5, *Lower*), respectively, was observed.

As expected, using the data in Fig. 5, *Lower* as an example, the bulkiest tag, in dA6P-Cy3-dT_4-FldT-dT-FldT-dT_{23}-C3, containing two fluorescein-modified dTs (FldT), gave the lowest current (~2 pA); the unmodified dT_{30} tag in dG6P-Cy3-dT_{30}-C3 produced an intermediate current (~4 pA), whereas higher currents were observed with the less bulky tags containing stretches of an abasic nucleotide, dSp, in dC6P-Cy3-dT_4-dSp_3-dT_{23}-C3 (~7 pA) and dT6P-Cy3-dT_2-dSp_8-dT_{20}-C3 (~8 pA).

In many cases (Fig. 5, *Lower*), a nucleotide tag appears to be captured several times in succession ("stuttering" at a given blockade current level) before the tag of the next nucleotide in the sequence is observed. The stuttering is presumably caused by the tagged nucleotide being repeatedly bound to and released from the polymerase before incorporation. Such stuttering can be reduced by changing experimental conditions. As seen in Fig. 6, the use of a 100-mV transmembrane potential and 150 mM KCl, along with 3 mM $MgCl_2$ on the *cis* side of the membrane and 3 mM $SrCl_2$ on the *trans* but not on the *cis* side of the membrane results in less stutter and also decreases the average duration of nucleotide captures. Such adjustment in the relative concentration of catalytic and noncatalytic metal cofactors is one means of making the tagged nucleotide incorporation rate slower than the combined rate of tag capture and detection by the pore. In this example, using a template with alternating homopolymer stretches and carefully considering the duration of each nucleotide capture event, we have obtained nearly perfect sequence reads with minimal stutter. Additional data on homopolymer sequencing are shown in Fig. S6. Optimization of the average duration of nucleotide capture events should allow further increases in sequencing accuracy. Base calling was carried out by manual inspection of the current level of each deflection, ignoring ones with dwell times less than 10 ms.

In summary, the sequencing data (Figs. 5 and 6 and Fig. S6) clearly indicate that our nanopore SBS method is capable of generating clear, sequence-dependent electronic signal streams at the single-molecule level in a continuous process that requires no addition or removal of reagents, sequencing an array of multiple DNA templates

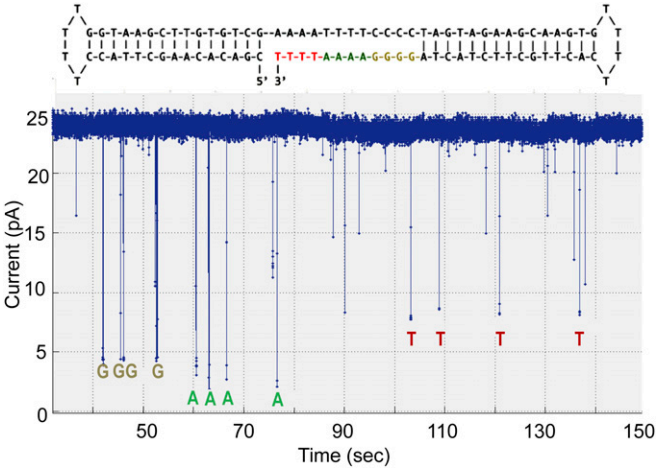

Fig. 6. Nanopore SBS results for a 12-base homopolymeric region of DNA. (*Upper*) A double-hairpin primer template was used to prevent template ends from entering the nanopore. The sequence immediately downstream from the 3′ end encodes 5′-GGGGAAAATTTT-3′. (*Lower*) Brief reductions in current are indicative of tag capture within the pore during the polymerase reaction, with the depth of the deflection characteristic of the different structures of the four tags. Other deflections in this figure, including the one at center, are brief (<10 ms) and were ignored as noise. The current trace shown is raw data output without noise reduction. Additional representative sequence data are shown in Fig. S6.

in parallel in a short time. Although the rate of sequencing and overall accuracy of nanopore SBS remain to be established, the system has many parameters that can be optimized further. These include synthesis of additional tagged nucleotides that offer better-resolved current blockade signatures, improvement of reaction conditions, and the testing of a variety of DNA polymerases that have improved rates, without the use of $SrCl_2$ to moderate the progression of DNA synthesis. With careful attention to these details, along with the construction of integrated circuits with many more electrodes and faster measurement capabilities, we expect to be able to create a high-throughput single-molecule electronic DNA sequencing system.

Methods

Synthesis of the Tagged Nucleotides Shown in Fig. 2 by Coupling 2′-Deoxynucleoside-5′-Hexaphosphate-Azides to 5′-Hexynyl Oligonucleotide Tags. The overall synthetic scheme for the synthesis of tagged nucleotides by coupling 2′-deoxynucleoside-5′-hexaphosphate-azides (dN6P-N_3) to 5′-hexynyl oligonucleotide tags is presented in Fig. S2. The detailed synthetic methods and characterization are presented in *SI Methods*. DNA polymerase extension reactions using these tagged nucleotides are described in *SI Methods*.

Nanopore SBS with Tagged Nucleotides. Sequencing reactions on nanopore chips were performed with inserted αHL pores conjugated to a single Phi29 DNA polymerase molecule, any of a variety of primer/template DNAs similar to that shown in Fig. S5*A*, the four tagged nucleotides (Fig. 2), 150 mM KCl, and 20 mM Hepes buffer (pH 7.5) at room temperature. Nucleotides consist of dA6P-Cy3-dT_4-FldT-dT-FldT-dT_{23}-C3, dG6P-Cy3-dT_{30}-C3, dC6P-Cy3-dT_4-dSp_3-dT_{23}-C3, and dT6P-Cy3-dT_2-dSp_8-dT_{20}-C3 (3 μM each), with 3 mM $SrCl_2$ on the *trans* side and 0.1 mM $MnCl_2$ on the *cis* side of the membrane (e.g., Fig. 5, *Upper*), 3 mM $SrCl_2$ *trans*, 3 mM $MgCl_2$, 0.7 mM $SrCl_2$ *cis* (Fig. 5, *Lower*), or 3 mM $SrCl_2$ *trans*, 3 mM $MgCl_2$ *cis* (Fig. 6). DC voltage of 100 mV was applied during sequencing.

ACKNOWLEDGMENTS. This work was supported by NIH Grant R01 HG007415 and Genia Technologies, Inc.

1. Soon WW, Hariharan M, Snyder MP (2013) High-throughput sequencing for biology and medicine. *Mol Syst Biol* 9:640.
2. Schatz MC, Delcher AL, Salzberg SL (2010) Assembly of large genomes using second-generation sequencing. *Genome Res* 20(9):1165–1173.
3. Harris TD, et al. (2008) Single-molecule DNA sequencing of a viral genome. *Science* 320(5872):106–109.
4. Eid J, et al. (2009) Real-time DNA sequencing from single polymerase molecules. *Science* 323(5910):133–138.
5. Korlach J, et al. (2010) Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol* 472:431–455.
6. Kasianowicz JJ, Brandin E, Branton D, Deamer DW (1996) Characterization of individual polynucleotide molecules using a membrane channel. *Proc Natl Acad Sci USA* 93(24):13770–13773.
7. Maitra RD, Kim J, Dunbar WB (2012) Recent advances in nanopore sequencing. *Electrophoresis* 33(23):3418–3428.
8. Robertson JW, et al. (2007) Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc Natl Acad Sci USA* 104(20):8207–8211.
9. Reiner JE, Kasianowicz JJ, Nablo BJ, Robertson JW (2010) Theory for polymer analysis using nanopore-based single-molecule mass spectrometry. *Proc Natl Acad Sci USA* 107(27):12080–12085.
10. Clarke J, et al. (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol* 4(4):265–270.
11. Kumar S, et al. (2012) PEG-labeled nucleotides and nanopore detection for single molecule DNA sequencing by synthesis. *Sci Rep* 2:684.
12. Cherf GM, et al. (2012) Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nat Biotechnol* 30(4):344–348.
13. Manrao EA, et al. (2012) Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat Biotechnol* 30(4):349–353.
14. Laszlo AH, et al. (2014) Decoding long nanopore sequencing reads of natural DNA. *Nat Biotechnol* 32(8):829–833.
15. Reiner JE, et al. (2012) The effects of diffusion on an exonuclease/nanopore-based DNA sequencing engine. *J Chem Phys* 137(21):214903.
16. Panwar AS, Muthukumar M (2009) Enzyme-modulated DNA translocation through a nanopore. *J Am Chem Soc* 131(51):18563–18570.
17. van Oijen AM, et al. (2003) Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science* 301(5637):1235–1238.
18. Ju J, et al. (2006) Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proc Natl Acad Sci USA* 103(52):19635–19640.
19. Guo J, et al. (2008) Four-color DNA sequencing with 3′-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc Natl Acad Sci USA* 105(27):9145–9150.
20. Bentley DR, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456(7218):53–59.
21. Lallana E, Riguera R, Fernandez-Megia E (2011) Reliable and efficient procedures for the conjugation of biomolecules through Huisgen azide-alkyne cycloadditions. *Angew Chem Int Ed Engl* 50(38):8794–8804.
22. Kumar S, et al. (2005) Terminal phosphate labeled nucleotides: Synthesis, applications, and linker effect on incorporation by DNA polymerases. *Nucleosides Nucleotides Nucleic Acids* 24(5-7):401–408.
23. Sood A, et al. (2005) Terminal phosphate-labeled nucleotides with improved substrate properties for homogeneous nucleic acid assays. *J Am Chem Soc* 127(8):2394–2395.
24. Movileanu L, Cheley S, Howorka S, Braha O, Bayley H (2001) Location of a constriction in the lumen of a transmembrane pore by targeted covalent attachment of polymer molecules. *J Gen Physiol* 117(3):239–252.
25. Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H (2009) Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc Natl Acad Sci USA* 106(19):7702–7707.
26. Berman AJ, et al. (2007) Structures of phi29 DNA polymerase complexed with substrate: The mechanism of translocation in B-family polymerases. *EMBO J* 26(14):3494–3505.
27. Tanaka Y, et al. (2011) 2-Methyl-2,4-pentanediol induces spontaneous assembly of staphylococcal α-hemolysin into heptameric pore structure. *Protein Sci* 20(2):448–456.
28. Kolb HC, Finn MG, Sharpless KB (2001) Click chemistry: Diverse chemical function from a few good reactions. *Angew Chem Int Ed Engl* 40(11):2004–2021.
29. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB (2002) A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew Chem Int Ed Engl* 41(14):2596–2599.
30. Takeshita M, Chang CN, Johnson F, Will S, Grollman AP (1987) Oligodeoxynucleotides containing synthetic abasic sites. Model substrates for DNA polymerases and apurinic/apyrimidinic endonucleases. *J Biol Chem* 262(21):10171–10179.
31. Gyarfas B, et al. (2009) Mapping the position of DNA polymerase-bound DNA templates in a nanopore at 5 A resolution. *ACS Nano* 3(6):1457–1466.
32. Zakeri B, et al. (2012) Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci USA* 109(12):E690–E697.
33. Davis R, Chen R, Bibillo A, Korenblum D, Dorwart M (2014) Nucleic acid sequencing using tags. Patent Application US20140134616.
34. Fuller CW, Kumar S, Ju J, Davis R, Chen R (2015) Chemical methods for producing tagged nucleotides. US Patent Appl US20150368710.
35. Blackman ML, Royzen M, Fox JM (2008) Tetrazine ligation: Fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J Am Chem Soc* 130(41):13518–13519.