Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision

Gerald M Cherf, Kate R Lieberman, Hytham Rashid, Christopher E Lam, Kevin Karplus & Mark Akeson

An emerging DNA sequencing technique uses protein or solid-state pores to analyze individual strands as they are driven in single-file order past a nanoscale sensor. However, uncontrolled electrophoresis of DNA through these nanopores is too fast for accurate base reads. Here, we describe forward and reverse ratcheting of DNA templates through the α-hemolysin nanopore controlled by phi29 DNA polymerase without the need for active voltage control. DNA strands were ratcheted through the pore at median rates of 2.5–40 nucleotides per second and were examined at one nucleotide spatial precision in real time. Up to 500 molecules were processed at ~130 molecules per hour through one pore. The probability of a registry error (an insertion or deletion) at individual positions during one pass along the template strand ranged from 10% to 24.5% without optimization. This strategy facilitates multiple reads of individual strands and is transferable to other nanopore devices for implementation of DNA sequence analysis.

Single-molecule techniques have been developed for commercial DNA sequencing. A promising new nanopore-based strategy will require integration of six features: (i) automated capture and processing of genomic DNA templates in single-file order from a heterogeneous mixture over many hours. This is essential to eliminate the extensive DNA sample preparation common to other sequencing technologies and thus to fully exploit the speed of nanopore sequencing; (ii) systematic spatial control so that DNA moves through the nanopore in ~5-Å increments; (iii) temporal control so that each nucleotide resides in the pore sensor for 0.1–1,000 ms; (iv) absence of complex active voltage control (necessary to avoid crosstalk in a compact electronic array of thousands of pores); (v) a sensor that can determine the identity of single bases; and (vi) a counter that can identify transitions between nucleotides in homopolymeric regions. Here we describe a system that implements the first four features.

Devices based on the pore-forming protein α-hemolysin (α-HL) are common in nanopore technology. Briefly, a single α-HL nanopore is inserted in a lipid bilayer that separates two wells, each containing 100 μl of a buffered KCl solution. Negatively charged single-stranded DNA (ssDNA) is added to the cis well. A voltage applied between the wells (trans side +) causes the ssDNA to enter and electrophores through the nanopore. This event results in a brief current blockade that is influenced by DNA strand length and base composition.

A consensus has emerged that the average rate of ssDNA electrophoresis through nanopores (~3 nt·ms⁻¹ at 120 mV for α-HL) is too fast to allow bases to be accurately identified. Therefore, a functional nanopore sequencing device will require a means to systematically slow DNA template movement. One proposed strategy involves coupling an enzyme motor to the nanopore. This strategy is attractive because many processive enzymes, including polymerases, ratchet along DNA strands one nucleotide at a time, up to tens of thousands of times in succession in bulk phase. Systematic, enzyme-driven displacement of a captured DNA strand relative to the nanopore would be anticipated at milliseconds per nucleotide step. In addition, as the enzyme motor pulls the DNA strand through the nanopore sensor, the force of the electric field acting in the opposite direction is predicted to hold the strand taut, and therefore reduce base read errors caused by Brownian motion.

Thus far, most of the enzyme motors coupled to nanopores have been DNA polymerases. T7 DNA polymerase (T7 DNAP) was shown to bind to DNA captured in the α-HL pore and then catalyze nucleotide additions that advanced the template strand through the pore against an 80-mV applied voltage. Replication was blocked in bulk phase using synthetic DNA complementary to the template strand. Although important as a proof of concept, this method could not be used in a practical sequencing device for two reasons: (i) at most, three sequential ionic current steps could be observed before the T7 DNAP dissociated from the DNA template under load; and (ii) to remove the blocking oligomer and subsequently bind T7 DNAP, the DNA was tethered in the pore and driven back and forth by reversing voltage polarity at 10-ms intervals. This would result in an unacceptable level of crosstalk between pores in a compact commercial array.

More recent experiments showed that a B-family polymerase, phi29 DNA polymerase (phi29 DNAP), remained bound to DNA captured in the α-HL pore ~10,000 times longer than did A-family polymerases. DNA replication by phi29 DNAP controlled sequential movement of at least 50 bases through the α-HL pore from a precise starting point on a DNA primer strand without active voltage control. The rate of elongation and template displacement was tens of milliseconds per nucleotide, con-

Nanopore Group, Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, California, USA. Correspondence should be addressed to M.A. (makeson@soe.ucsc.edu).

Received 27 September 2011; accepted 7 February 2012; published online 14 February 2012; doi:10.1038/nbt.2147
Figure 1 Experimental set-up. (a) Nanopore device. A single α-HL nanopore is inserted in a lipid bilayer that separates two wells, each containing 100 μl of a buffered KCl solution. DNA bearing a ssDNA segment is added to the cis well. A voltage applied between the wells (trans side +) results in an ionic current that is modulated as individual ssDNA molecules traverse the nanopore. (b) The blocking oligomer strategy. (i) The blocking oligomer (dashed line) protects a DNA substrate composed of a 23 nt primer annealed to a 70-mer template strand (Supplementary Fig. 1). (ii) The blocking oligomer contains a 25-mer segment complementary to the DNA template near the primer-template junction. It is capped at its 5′-end by two acridines (z), and has a 3′-tail composed of seven abasic residues (x) capped by a three-carbon spacer at its terminus (a) (Supplementary Fig. 2). (c) Test of DNA substrate protection in bulk phase using blocking oligomers. DNA substrates were incubated at 23 °C for 5 h in nanopore buffer in the presence of phi29 DNAP, dNTPs and Mg²⁺ as noted, with or without blocking oligomers. Fluorescently tagged primers were separated by denaturing PAGE. The arrows mark the full-length product (+56-nt extension), blocking oligomer, the starting primer (23-nt primers) and products of exonucleolysis (degradation products).

Our next objective was to remove the blocking oligomer from each individual DNA template captured in the nanopore so that phi29 DNAP could bind at the primer-template junction. Initially we considered a proven strategy whereby active voltage control is used to unzip the blocking oligomer from the DNA template upon capture, followed by a voltage polarity reversal to drive the newly exposed DNA primer-template junction into the cis well to fish for a polymerase molecule.¹¹,¹³

This complex method proved to be unnecessary as illustrated in the following experiment. The substrate was a 94-mer DNA template strand bearing five abasic (1', 2'-H) residues spanning positions +25 to +29 from the n = 0 position (Fig. 2a). This abasic insert serves as an ionic current reporter during strand displacement through the α-HL pore.¹¹,¹²,¹⁶,¹⁷ The DNA template was annealed to a 23-nt primer, and the 3′-terminus of the primer strand was protected from bulk phase modification by the blocking oligomer described above (Fig. 1b, ii).

Addition of this DNA construct alone to the nanopore cis chamber resulted in ionic current blockades with a median duration of ~4 ms and an average residual current of 22.5 pA, which are similar to translocation events for DNA substrates bearing short duplex regions described previously.¹⁸ Events with a duration >200 ms were rare. By comparison, Figure 2b shows an ionic current trace typical of 200 events when phi29 DNAP and dNTPs were subsequently added (see also Supplementary Videos 1 and 2). Capture of a DNA substrate molecule (Fig. 2b, i) resulted in a 23–24 pA residual ionic current that lasted several seconds (Fig. 2b, ii). Under a sustained 180-mV load, the ionic current then stepped through a series of discrete levels that traversed a 35-pA maximum (Fig. 2b, iii) before dropping to 22 pA and settling at a characteristic 25-pA amplitude (Fig. 2b, iv). These current levels were caused by sequential movement of the five abasic residues of the template strand through the α-HL transmembrane pore. This effect is especially pronounced as the abasic residues enter, then pass through, the α-HL limiting aperture circumscribed by lys147 (ref. 12). Upon reaching the 25-pA amplitude, the ionic current steps reversed direction and retraced the 35-pA peak (Fig. 2b, v) at about ten times the speed that the first peak was traversed, before stalling at 24 pA (Fig. 2b, vi). In this experiment, there were 178 additional events where the ionic current series began as shown in Figure 2b, but did not progress completely through the two peaks owing to either enzyme dissociation (112/178), a stall within the first ionic current peak (23/178) or a stall in the second ionic current peak (44/178).

These nanopore data suggest that when phi29 DNAP was added to the nanopore bath in the presence of the protected DNA substrates,
it formed stable but enzymatically inactive complexes owing to the presence of the blocking oligomers. Activation of a given complex was achieved only upon capture by the nanopore. Figure 2c illustrates successive stages of this hypothetical process: (i) the open channel; (ii) nanopore capture of a polymerase-DNA complex with a blocking oligomer bound; (iii) mechanical unzipping of the blocking oligomer promoted by the applied voltage, which ratches the DNA template forward through the nanopore (this gives rise to the first 35-pA current peak as the abasic insert traverses the major pore constriction); (iv) release of the blocking oligomer, which exposes the 3′-OH terminus of the DNA primer within the polymerase active site; (v) DNA replication by phi29 DNAP, which ratches the template in the reverse direction through the nanopore, giving rise to the second 35-pA current peak; (vi) stalling of DNA replication when the abasic residues of the template strand reach the catalytic site of phi29 DNAP.

This model makes three testable predictions. First, traversal of the first 35-pA peak resulting from voltage-driven unzipping of the blocking oligomer should be independent of phi29 DNAP catalytic capability. Therefore it should be observed in the absence of the Mg2+ ions required for both polymerase and exonuclease function. In experiments in which complexes of phi29 DNAP and the substrate shown in Figure 2a were captured in the absence of free Mg2+, the first 35-pA peak was indeed traversed, followed by a stall at the 25-pA level and eventual voltage-promoted dissociation of the complex (Supplementary Fig. 3). The second 35-pA peak was not observed in the absence of Mg2+, supporting the second prediction of the model, that because traversal of the second 35-pA ion current peak requires DNA replication, it should be dependent upon the presence of both Mg2+ and a full complement of dNTP substrates. As an additional test, traversal of this second peak should stall if one of the required dNTP substrates was withheld. Results consistent with this prediction are shown in Supplementary Figure 3.

The third prediction of this model is that progression into the proposed replication-dependent peak should be influenced by the chemical identity of the DNA primer terminus. In particular, substitution of the 3′-OH terminus with a 3′-H terminus should delay appearance of the second 35-pA current peak by causing a stall as the primer-template junction is positioned in the polymerase active site (Fig. 2b, iv and 2c, iv). This prediction also proved to be correct (Fig. 2d). That is, using a substrate bearing a primer with a 3′-H terminus, the first 35-pA peak was traversed as it was with the substrate bearing a 3′-OH terminus.
owing to voltage-driven unzipping of the blocking oligomer. The ionic current then stalled for several seconds at 25 pA (red horizontal arrow, Fig. 2d). Eventually, traversal of the second 35-pA peak was observed. Previous work\textsuperscript{12} has shown that this recovery was due to excision of the 3’-H terminated residue by the phi29 DNAP exonuclease and subsequent strand elongation beginning at the newly exposed 3’-OH of the neighboring dGMP nucleotide.

Together these experiments indicate that phi29 DNAP can be used to control forward and reverse ratcheting of individual DNA templates through the \( \alpha \)-HL pore. In the forward direction, the template strand is driven 5’ to 3’ through the nanopore by applied voltage as its complementary blocking oligomer is unzipped at the bound phi29 DNAP. In the reverse direction, replication by phi29 DNAP extends the primer strand and thus rectifies movement of the template in the 3’-to-5’ direction relative to the pore.

To quantify the average rate of DNA template movement in each mode, we sought to account for all 25 template nucleotides in each direction along the trace. We first established an ionic current map by building a composite derived from ten traces that traversed both amplitude peaks (Fig. 3a,b). Thirty-two reproducible amplitudes were resolved for more than one-half of the replicated DNA templates using a 3-ms minimum cutoff. These amplitude steps were symmetric around a 25-pA midpoint (ionic current state 0) except for state 1, which was not observed in a majority of traces. To confirm that the 16 ionic current states in the replication-dependent peak corresponded to displacement of 25 template nucleotides, we measured translocation pauses when one dNTP substrate at a time was reduced to 100 nM in the nanopore buffer while all other dNTPs were held at 100 \( \mu \)M (Supplementary Fig. 4). As anticipated, when these concentration-dependent pauses were assembled in logical order, 25-nt additions to the DNA primer strand could be accounted for within ionic current states 0 to 16 of the map (n0 to n24, Fig. 3b). Along this 25-nt DNA segment, the median rate of replication was 40 nt s\(^{-1}\) (interquartile range (IQR) = 12 nt s\(^{-1}\), \( n = 200 \)). The ionic current states during voltage-driven unzipping (states –16 to 0) mirrored the states observed during replication. If we assume 25-nt displacement during that process as well, the median rate of translocation during unzipping was measurably slower (median = 2.5 nt s\(^{-1}\), IQR = 3.2 nt s\(^{-1}\), \( n = 200 \)).

To increase throughput, we reduced the length of the blocking oligomer segment annealed to the target DNA template from 25 to 15 nt (in the context of the blocking oligomer design shown in Fig. 1b, ii). When annealed to a primer-template substrate (Supplementary Fig. 5a), this blocking oligomer still afforded protection of DNA in bulk phase (Supplementary Fig. 5b) but allowed faster removal of the blocking oligomer at the nanopore. With this shorter blocking oligomer, up to 500 molecules were processed in single-file order at a rate of 130 per hour through one pore.

We next determined the probability of template registry errors that occurred when using this phi29 DNAP control strategy (Fig. 4a–e). There were two types of errors. The first occurred when the rate of strand displacement past the pore sensor exceeded the rate of data acquisition (the nucleotide at a given position would be missed or ‘deleted’ (compare Fig. 4a, i and ii)). The second type occurred when the strand slipped back and forth so that a given position was read more than once (a nucleotide would be falsely ‘inserted’ into the sequence order (compare Fig. 4a, i and iii)). Both types of error resulted in register-dependent ‘indel’ errors—that is, insertion or deletion of bases relative to the correct sequential series.

To accurately measure the frequency of these registry errors, we focused on ionic current states that correspond to single-nucleotide positions bounded by clearly discernible neighboring states. The data in Figure 3b suggest that state 8 (bounded by states 7 and 9) and state 11 (bounded by states 10 and 12) within the replication-dependent peak satisfied these criteria. This was confirmed using a DNA mapping strategy described previously\textsuperscript{12–14} (Fig. 4b and Supplementary Fig. 6). By extension, the mirror-image ionic current states –8 and –11 of the voltage-driven peak were also used.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4}
\caption{Estimating DNA template registry errors in the nanopore during phi29 DNAP–controlled translocation. The DNA template and experimental conditions are identical to those described in Figure 2. (a) Example of ionic current traces. The traces shown are for the replication-driven ratchet moving DNA templates between ionic current state 7 (29 pA), state 8 (29 pA) and state 9 (34 pA). Movement from state 8 in either direction is equivalent to a 1-nt displacement. (i) Correct read. The current rises from 27 pA to 29 pA and resides there for at least 3 ms before advancing to 34 pA. (ii) Deletion. The ionic current advances directly from 27 pA to 34 pA and fails to reside at 29 pA for at least 3 ms (arrow). (iii) Insertion. The ionic current trace advances from 27 pA to 29 pA. It resides at 29 pA for at least 3 ms but then slips back to 27 pA for at least 3 ms (arrow). These flickers between states may occur more than once before advancing through 29 pA and on to 34 pA, but were scored as a single insertion in this analysis. (b) All ionic current states (black circles) in the replication-dependent peak were confirmed using a DNA mapping strategy described previously\textsuperscript{12,14}. Numbers n10 to n17 (red circles) correspond to nucleotide positions along the template strand as in Figure 3b. Details on how this map was prepared are given in Supplementary Figure 6. (c,d) Model used for error estimates. The black circles are the same ionic current states as those represented in Figure 3b for –12 to –7 (c, voltage-driven zipper peak) and 7 to 12 (d, replication-driven ratchet peak). Ionic current state 0 is shown at the center of the map for reference and to emphasize symmetry. States –11, –8, 8 and 11 are the center points of error measurements and are represented by open circles. The green arrows represent correct ionic current progression through those central ionic current states in the absence of ‘indel’ errors. The black dotted lines indicate ‘insertions’. The red dotted lines represent ‘deletions’. The traces labeled i–iii in a correspond to events of types i–iii passing through ionic current state 8 in panel d. (e) Probability of single-nucleotide registry errors for individual molecules ratcheting through the \( \alpha \)-HL nanopore controlled by phi29 DNAP. Estimates are based on 200 transitions through each of states –11, –8, 8 and 11 for the DNA construct illustrated in Figure 2a.}
\end{figure}
We used models of the insertion and deletion events (Fig. 4a) to quantify registry errors for the voltage-driven zipper (Fig. 4c) and replication-driven ratchet (Fig. 4d). Using the ratchet as an example, a correct read was called when the current advanced from ionic current state 7 to state 8 and remained there for at least 3 ms before advancing to ionic current state 9 (Fig. 4d, i, green arrows, and Fig. 4a, i). A deletion error was called when the current advanced from ionic current state 7 directly to state 9 without remaining in state 8 for at least 3 ms (Fig. 4d, ii, red dashed arrow, and Fig. 4a, ii). An insertion error was called when the current advanced from ionic current state 7 to state 8 and remained there for at least 3 ms, but then returned to state 7 for at least 3 ms before advancing once again through state 8 and then to state 9 (Fig. 4d, iii, gray dashed arrow, and Fig. 4a, iii). Corresponding arrows centered at states –8 and –11 of the voltage-driven zipper (Fig. 4c) and state 11 of the replication-driven ratchet (Fig. 4d) have the same meaning.

Analysis of 200 current traces yielded error estimates for individual molecules read once in each direction (Fig. 4e). We found that the insertion error probability at 1-nm spatial precision was 5–10.5%, and the deletion error probability was 5–15%. The combined error probability (either an insertion or deletion at a given position in a single pass) was 10–24.5%. Errors at position 11 or –11 were about half as frequent as errors at positions 8 or –8.

We conclude that DNA substrates pre-bound to phi29 DNA polymerase (DNAP) can be protected from enzymatic modification for many hours using blocking oligomers, and that activation of each complex for replication occurs protected from enzymatic modification for many hours using blocking molecules read once in each direction (Fig. 4d) have the same meaning. Driven ratchet (–11 of the voltage-driven zipper (Fig. 4d) and state 11 of the replication-state 11 of the replication-driven ratchet (Fig. 4d) have the same meaning.

Analysis of 200 current traces yielded error estimates for individual molecules read once in each direction (Fig. 4e). We found that the insertion error probability at 1-nm spatial precision was 5–10.5%, and the deletion error probability was 5–15%. The combined error probability (either an insertion or deletion at a given position in a single pass) was 10–24.5%. Errors at position 11 or –11 were about half as frequent as errors at positions 8 or –8.

We conclude that DNA substrates pre-bound to phi29 DNA polymerase (DNAP) can be protected from enzymatic modification for many hours using blocking oligomers, and that activation of each complex for replication occurs protected from enzymatic modification for many hours using blocking molecules read once in each direction (Fig. 4d) have the same meaning. Driven ratchet (–11 of the voltage-driven zipper (Fig. 4d) and state 11 of the replication-driven ratchet (Fig. 4d) have the same meaning.

Analysis of 200 current traces yielded error estimates for individual molecules read once in each direction (Fig. 4e). We found that the insertion error probability at 1-nm spatial precision was 5–10.5%, and the deletion error probability was 5–15%. The combined error probability (either an insertion or deletion at a given position in a single pass) was 10–24.5%. Errors at position 11 or –11 were about half as frequent as errors at positions 8 or –8.

We conclude that DNA substrates pre-bound to phi29 DNA polymerase (DNAP) can be protected from enzymatic modification for many hours using blocking oligomers, and that activation of each complex for replication occurs protected from enzymatic modification for many hours using blocking molecules read once in each direction (Fig. 4d) have the same meaning. Driven ratchet (–11 of the voltage-driven zipper (Fig. 4d) and state 11 of the replication-driven ratchet (Fig. 4d) have the same meaning.

Analysis of 200 current traces yielded error estimates for individual molecules read once in each direction (Fig. 4e). We found that the insertion error probability at 1-nm spatial precision was 5–10.5%, and the deletion error probability was 5–15%. The combined error probability (either an insertion or deletion at a given position in a single pass) was 10–24.5%. Errors at position 11 or –11 were about half as frequent as errors at positions 8 or –8.

We conclude that DNA substrates pre-bound to phi29 DNA polymerase (DNAP) can be protected from enzymatic modification for many hours using blocking oligomers, and that activation of each complex for replication occurs protected from enzymatic modification for many hours using blocking molecules read once in each direction (Fig. 4d) have the same meaning. Driven ratchet (–11 of the voltage-driven zipper (Fig. 4d) and state 11 of the replication-driven ratchet (Fig. 4d) have the same meaning.

Finally, processive –5-Å DNA template displacement was documented at two positions each for the voltage-driven zipper and for the replication-driven ratchet. This allowed us to make initial probability estimates for registry errors resulting from uncontrolled strand motion at the single-nucleotide scale. These DNA template registry errors (10–24.5% combined probability for insertions and deletions at a given position) must be reduced for a commercial nanopore sequencing device. This will be a fruitful area of research because several factors that influence template motion and the ionic current signal-to-noise ratio have not been optimized. These include ionic strength, polymerase mechanics, enzyme identity, applied voltage and composition of the trans-side buffer. Even without further improvements, this robust control strategy should enable DNA sequencing when coupled to nanopores capable of nucleotide discrimination.19–21.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/