Automated Forward and Reverse Ratcheting of DNA in a Nanopore at Five Angstrom Precision[1]

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

Nanopore Group, MS SOE2, Department of Biomolecular Engineering University of California, Santa Cruz, CA 95064.

Abstract

Single-molecule techniques have been developed for commercial DNA sequencing[1,2]. One emerging strategy uses a nanopore to analyze DNA molecules as they are driven electrophoretically in single file order past a sensor[3-5]. However, uncontrolled DNA strand electrophoresis through nanopores is too fast for accurate base reads[6]. A proposed solution would employ processive enzymes to deliver DNA through the pore at a slower average rate[7]. Here, we describe forward and reverse ratcheting of DNA templates through the α–hemolysin (α-HL) nanopore controlled by wild-type phi29 DNA polymerase (phi29 DNAP). DNA strands were examined in single file order at one nucleotide spatial precision in real time. The registry error probability (either an insertion or deletion during one pass along a template strand) ranged from 10% to 24.5% absent optimization. This general strategy facilitates multiple reads of individual template strands and is transferrable to other nanopore devices for implementation of DNA sequence analysis.

A commercial nanopore device for *de novo* DNA sequencing 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 extensive DNA sample preparation common to other sequencing technologies and thus fully exploit the speed of nanopore sequencing; ii) systematic spatial control at *circa* 5 Å precision; iii) temporal control at 0.1-1000 ms per nucleotide; iv) absence of complex active voltage control. This is necessary to avoid crosstalk in a compact electronic array of thousands of nanopores; v) a sensor that can determine single base identity; and vi) a counter that can identify transitions between nucleotides along homopolymeric regions.

Devices based on the pore-forming protein α-HL are common in nanopore technology [3-5, 8]. Briefly, a single α-HL nanopore is inserted in a lipid bilayer that separates two wells that

contain 100 µl each of a buffered KCl solution (**Fig. 1a**). 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 electrophorese through the nanopore. This event results in a brief current blockade that is influenced by DNA strand length[3] and base composition[5].

A consensus has emerged that the average rate of ssDNA electrophoresis through nanopores (approximately 3 µs nt^{-1} at 120 mV for α-HL) is too fast to allow accurate base identification [6]. Therefore a functional nanopore sequencing device will require a means to systematically slow DNA template movement. One proposed strategy involves coupling a protein motor to the nanopore [7]. 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 [9]. In addition, as the polymerase drives 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[10].

Thus far most progress has been made with 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 [11]. Replication was blocked in bulk phase using synthetic DNA 'blocking oligomers', and was initiated only after each DNA template was captured by the nanopore with attendant removal of the block by voltage-driven unzipping. Although an important proof of concept, this method could not be used in a practical sequencing device for two reasons: 1) at most three sequential ionic current steps could be observed before the T7 DNAP dissociated from the DNA template under load; and 2) 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 DNAP, remained bound to DNA captured in the α-HL pore approximately 10,000-fold longer than did A family polymerases[12]. 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, consistent with previously measured rates [13,14]. However, these experiments relied upon transient chemical protection of the DNA primer terminus to prevent elongation and excision in bulk phase [12]. This approach permitted only a 20 minute window to capture unmodified DNA from bulk phase, and therefore is not suitable for a commercial sequencing device.

In the work presented here we combine phi29 DNAP-dependent template control with an improved blocking oligomer strategy for capture, activation and electrophoresis of up to 500 DNA molecules in single file order through individual nanopores. Importantly, we discovered that blocking oligomers potentiated formation of phi29 DNA polymerase-DNA complexes in bulk phase that were quiescent for at least five hours. This eliminated the need for active voltage control to effect polymerase binding to DNA. It also enabled automated

forward and reverse ratcheting of each DNA strand through the pore, exploiting phi29 DNAP in two distinct mechanistic modes for template strand analysis.

Figure 1b illustrates features of the blocking oligomer designed for use with phi29 DNAP. The DNA substrate is a 23-nucleotide (nt) primer annealed to a synthetic 70-nt DNA template (**Supplementary Fig. 1**). To protect the DNA primer from phi29 DNAP-dependent extension and digestion in bulk phase, a blocking oligomer (dashed line, **Fig. 1b, i**) is annealed immediately adjacent to the DNA primer/template (p/t) junction. The blocking oligomer includes a ~25-nt complement to the template strand and two acridine (z) residues attached to the 5′ end (**Fig. 1b, ii**). One of these acridines substitutes for a nucleotide and abuts the primer terminus; the other is an added 5′-overhang that is presumed to intercalate into the DNA duplex. The blocking oligomer was appended with a 3′- three carbon spacer (s) followed by seven abasic (1′, 2′-H) residues (x's). This appended segment has two functions: protection of the blocking oligomer against exonucleolysis by phi29 DNAP (**Supplementary Fig. 2**), and facilitation of blocking oligomer removal as the phi29 DNAP-DNA complex is pulled into the nanopore by an applied voltage.

The blocking oligomer protected the DNA primer strand from phi29 DNAP-dependent extension and digestion in bulk phase (**Fig. 1c**). In these experiments, DNA substrates and phi29 DNAP were incubated in nanopore buffer (0.3 M KCl, 10 mM Hepes/KOH pH 8, 1 mM EDTA, 1 mM DTT) supplemented with 10 mM Mg^{2+} for 5 hours at 23 °C. The products were then analyzed by denaturing polyacrylamide gel electrophoresis. Absent the blocking oligomer, phi29 DNAP digested the DNA primer strands (-dNTP, lane 3) or extended them (+dNTP, lane 4). In contrast, when protected by the blocking oligomer, the primer strands were neither digested (-dNTP, lanes 6), nor extended (+dNTP, lanes 7) by phi29 DNAP.

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 (p/t) junction. Initially we considered a proven strategy wherein active voltage control was 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 [11,15]. This complex method proved to be unnecessary. Figure 2 shows an experiment using a 94mer 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 ~4ms and an average residual current of 22.5 pA that are similar to translocation events for DNA substrates bearing short duplex regions described previously[18]. Events greater than 200 ms duration were rare. By comparison, Figure 2 shows an ionic current trace typical of 200 events when phi29 DNAP and dNTPs were subsequently added (see also **Supplemental Videos 1 & 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 trans-membrane pore. This effect is especially pronounced as the abasic residues enter then pass through the α–HL limiting aperture circumscribed by lys147 [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 also 178 events where the ionic current series began as shown in Figure 2, but did not progress completely through the two peaks either due to 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 due to the presence of the blocking oligomers. Activation of a given complex was achieved only upon nanopore capture. Successive stages of this hypothetical process are illustrated in Figure 2c: 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 ratchets the DNA template 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 ratchets 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 due to 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 Mg^{2+} 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 absent free Mg^{2+}, 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 Mg^{2+}, supporting the second prediction of the model: because traversal of the second 35 pA ionic current peak requires DNA replication, it should be dependent upon the presence of both Mg^{2+} 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 described 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 p/t junction is positioned in the polymerase active site (iv in **Fig 2b,c**). This prediction also proved to be correct (**Fig. 2d**). That is, using a substrate bearing a 3′-H terminated primer, the first 35 pA peak was traversed as it was with the substrate bearing a 3′-OH terminus, due 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. 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 [12].

Together these experiments indicate that phi29 DNAP can be used to control forward and reverse ratcheting of individual DNA templates through the α-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 enzyme. In the reverse direction, replication by phi29 DNAP extends the primer strand and thus biases movement of the template in the 3′-to-5′ direction relative to the pore. To quantify the average rate of DNA template movement per nucleotide in each mode, it was necessary 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 10 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 correspond 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 μM (**Supplementary Fig. 4**). As anticipated, when these concentration-dependent pauses were assembled in logical order, 25 nucleotide 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 nucleotide DNA segment, the median rate of replication was 40 nt s^{-1} (IQR = 12 nt s^{-1}, n=200). The ionic current states during voltage-driven unzipping (states 0 to -16) mirrored the states observed during replication. If we assume 25-nucleotide 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 nucleotides (in the context of the blocking oligomer design shown in Figure 1b, ii). When annealed to a p/t 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 occur when using this phi29 DNAP control strategy (**Fig. 4a-d**). There are two types: 1) errors that occur when the rate of strand displacement past the pore sensor exceeds the rate of data acquisition (the

nucleotide at a given position would be missed or 'deleted' (compare **Fig. 4a. i & ii**)); and 2) errors that occur when the strand slips back and forth so that a given position is read more than once (a nucleotide would be falsely 'inserted' into the sequence order (compare **Fig 4a i & iii**)). Together these would result in registry-dependent "indel" errors - insertion or deletion of bases relative to the correct sequential series.

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

The model we used to quantify registry errors is illustrated in Figure 4d(i-iii) for the replication-driven ratchet with examples in Figure 4a (i-iii). A correct read was called when the current advanced from ionic current state 7 to state 8 and resided there for at least 3 ms before advancing to ionic current state 9 (**Fig. 4d,i**, green arrows; & **Fig. 4a,i**). A deletion error was called when the current advanced from ionic current state 7 directly to state 9 without residing in state 8 for at least 3 ms (**Fig. 4d,ii**, red dashed arrow; & **Fig. 4a,ii**). An insertion error was called when the current advanced from ionic current state 7 to state 8 and resided there for at least 3 ms, but then returned to state 7 for at least 3ms before advancing once again through state 8 and then to state 9 (**Fig. 4d,iii** , grey dashed arrow; & **Fig. 4a,iii**). Corresponding arrows centered at states -8 and -11 in Figure 4c (voltage-driven zipper) and state 11 in Figure 4d (replication-driven ratchet) have the same meaning.

Error estimates for individual molecules read once in each direction are summarized in Figure 4e. We found that the insertion error probability at one nucleotide spatial precision ranged from 5% to 10.5%, and the deletion error probability ranged from 5% to 15%. The combined error probability (either an insertion or deletion at a given position in a single pass) ranged from 10% to 24.5%. Errors at position 11/-11 were about ½ as frequent as errors at positions 8/-8.

We conclude that DNA substrates pre-bound to phi29 DNAP can be protected from enzymatic modification for many hours using blocking oligomers, and that activation of each complex for replication occurs upon capture on the α–HL pore. Because the DNA molecules are enzymatically modified only at the nanopore, it is possible to combine all components of the replication reaction (i.e. DNA template, dNTP, Mg^{2+}, polymerase) in the nanopore chamber at time zero and run a lengthy analysis of many DNA templates without further user intervention. Thus, DNA strand sequence could be determined for an unknown template in a homogeneous preparation by analyzing a series of single molecules. Alternatively, DNA strand sequencing could be performed on molecules captured one-byone from a heterogeneous mixture (e.g. a genomic digest) with multiple reads per individual strand to achieve consensus.

Finally, processive *circa* 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 due to uncontrolled strand motion at the single nucleotide scale. These DNA template registry errors (10% to 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 absent improvements, proof of concept DNA sequencing is conceivable using this robust control strategy when coupled to nanopores capable of nucleotide discrimination[18-20].

ONLINE METHODS

Materials

Heptameric α-HL was provided by Oxford Nanopore Technologies. Diphytanoylphosphatidylcholine (DPyPC) lipid was purchased from Avanti Polar Lipids. Wild-type phi29 DNAP (833,000 U ml^{-1}; specific activity 83,000 U mg^{-1}) was supplied by Enzymatics Corporation. Phosphoramidites (including abasic residue (dSpacer), 3′-spacer C3 CPG, and acridine) were from Glen Research. DNA oligonucleotides were synthesized at the Stanford University Protein and Nucleic Acid (PAN) Facility. Lyophilized oligonucleotides received from PAN were re-suspended in 7M urea, 0.1× TBE, purified by denaturing polyacrylamide gel electrophoresis, and quantified using a Nanodrop™ 1000 (Thermo Scientific).

Bulk phase assay of DNA elongation and digestion by phi29 DNAP in the presence or absence of blocking oligomer

2 μM each of a 70mer DNA template (**Supplementary Fig. 1**) and a 23-nt DNA primer bearing a 5′ fluorescein (6-FAM) label, were combined in the presence or absence of blocking oligomer (**Fig 1b(ii)**) at 2.4 μM final concentration. This mixture was incubated at 90 °C for three minutes in 1X TE , 100 mM KCl buffer, followed by slow-cooling to room temperature. 1 μM of preannealed DNA substrate was then incubated for 5 hours at room temperature (approximately 23°C) in nanopore buffer (22.5 μL final volume) supplemented with 10 mM MgCl$_2$, 0.75 μM phi29 DNAP, and 100 μM dNTPs as indicated (**Fig. 1**). Reactions were terminated by the addition of buffer-saturated phenol. Following phenol/chloroform extraction and ethanol precipitation, dried DNA pellets were dissolved in 7M urea, 0.1X TBE and resolved by denaturing polyacrylamide gel electrophoresis. Gels were 17% acrylamide:bisacrylamide (19:1), 7 M urea, 1× TBE, and run at 18 W. 6-FAM labeled DNA primer products were visualized on a UVP Gel Documentation device using a Sybr Gold filter.

Nanopore experiments

Setup of the nanopore device and insertion of an α-HL nanopore into a lipid bilayer have been described [4]. Briefly, a single α-HL nanopore was inserted into a lipid bilayer that separates two wells that each contained 100 μl of nanopore buffer (pH 8). A 180 mV

potential was applied across the bilayer and ionic current was measure through the nanopore between AgCl electrodes in series with an integrating patch clamp amplifier (Axopatch 200B, Molecular Devices) in voltage clamp mode. Data were recorded using an analog-to-digital converter (Digidata 1440A, Molecular Devices) at 100 kHz bandwidth in whole-cell configuration then filtered at 5 kHz using a analogue low-pass Bessel filter. Experiments were conducted at 23°C with 1 μM DNA substrate preannealed to blocking oligomer in nanopore buffer with 0.75 μM phi29 DNAP, 10 mM Mg^{2+}, and 100 μM of each dNTP added to the nanopore *cis* well unless otherwise noted. A single nanopore experiment is defined as the time during which ionic current data were acquired from one α–HL nanopore in an intact bilayer before termination by bilayer rupture, loss of channel conductance, or completion of a preset number of translocation events. Successful forward and reverse DNA template translocations ranged for 50 to 500 per experiment .The ratio of fast (<200 ms) DNA alone events to polymerase-bound DNA events was approximately 10 to 1 unless otherwise noted.

Nanopore data analysis

Ionic current output from the nanopore device was analyzed using Clampfit 10.2 software (Axon Instruments) after additional smoothing of data at 2 kHz using a low-pass Gaussian filter unless otherwise noted. The standardized ionic current map in Figure 3b was compiled from ten translocation events from the experiment in Figure 2. These ten events each started at 23 pA, traversed two peaks at 35 pA, and ended by stalling at 23-24 pA. Ionic current steps within these ten translocation events were counted as discrete states if their durations were 3 ms or greater. For the data summarized in Supplementary Figure 4, sequence-dependent pauses were scored if the duration of a discrete ionic current state between 0 and 16 was 200 ms or greater, or if fluctuations between any two discrete states between 0 and 16 was 200 ms or greater.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEGEMENTS

REFERENCES

1. Eid J, et al. Real-time DNA sequencing from single polymerase molecules. Science. 2009; 323:133–138. [PubMed: 19023044]

2. Harris TD, et al. Single-molecule DNA sequencing of a viral genome. Science. 2008; 320:106–109. [PubMed: 18388294]

3. Kasianowicz JJ, Brandin E, Branton D, Deamer DW. Characterization of individual polynucleotide molecules using a membrane channel. Proc. Natl. Acad. Sci. USA. 1996; 93:13770–13773. [PubMed: 8943010]

4. Akeson M, Branton D, Kasianowicz JJ, Brandin E, Deamer DW. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. Biophys. J. 1999; 77:3227–3233. [PubMed: 10585944]

5. Meller A, Nivon L, Brandin E, Golovchenko J, Branton D. Rapid nanopore discrimination between single polynucleotide molecules. Proc. Natl. Acad. Sci. USA. 2000; 97:1079–1084. [PubMed: 10655487]

6. Branton D, et al. The potential and challenges of nanopore sequencing. Nat. Biotechnol. 2008; 10:1146–1153. [PubMed: 18846088]

7. Church GM, Deamer DW, Branton D, Baldarelli R, Kasianowicz J. 1998 ("US patent # 5,795,782 (filed March 1995) Characterization of individual polymer molecules based on monomer-interface interaction.

8. Movileanu L, Cheley S, Bayley H. Partitioning of individual flexible polymers into a nanoscopic protein pore. Biophys. J. 2003; 85:897–910. [PubMed: 12885637]

9. Blanco L, et al. Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. J. Biol. Chem. 1989; 264:8935–8940. [PubMed: 2498321]

10. Lu B, Albertorio F, Hoogerheide DP, Golovchenko JA. Origins and consequences of velocity fluctuations during DNA passage through a nanopore. Biophys. J. 2011; 101:70–79. [PubMed: 21723816]

11. Olasagasti F, et al. Replication of individual DNA molecules under electronic control using a protein nanopore. Nat. Nanotechnol. 2010; 11:798–806. [PubMed: 20871614]

12. Lieberman KR, et al. Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase. J. Am. Chem. Soc. 2010; 132:17961–17972. [PubMed: 21121604]

13. Soengas MS, Gutierrez C, Salas M. Helix-destabilizing activity of phi 29 single-stranded DNA binding protein: effect on the elongation rate during strand displacement DNA replication. J. Mol. Biol. 1995; 253:517–29. [PubMed: 7473731]

14. Ibarra B, et al. Proofreading dynamics of a processive DNA polymerase. EMBO J. 2009; 28:2794–2802. [PubMed: 19661923]

15. Wilson NA, et al. Electronic control of DNA polymerase binding and unbinding to single DNA molecules. ACS Nano. 2009; 3:995–1003. [PubMed: 19338283]

16. Hurt N, Wang H, Akeson M, Lieberman KR. Specific nucleotide binding and rebinding to individual DNA polymerase complexes captured on a nanopore. J. Am. Chem. Soc. 2009; 131:3772–3778. [PubMed: 19275265]

17. Gyarfas B, et al. Mapping the position of DNA polymerase-bound DNA templates in a nanopore at 5 A resolution. ACS Nano. 2009; 3:1457–1466. [PubMed: 19489560]

18. Benner S, et al. Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore. Nature Nanotech. 2007; 2:718–724.

19. Derrington IM, et al. Nanopore DNA sequencing with MspA. Proc. Natl. Acad. Sci. USA. 2010; 107:16060–16065. [PubMed: 20798343]

20. Stoddart D, et al. Nucleobase recognition in ssDNA at the central constriction of the alpha-hemolysin pore. Nano Lett. 2010; 10:3633–3637. [PubMed: 20704324]

21. Tsutsui M, Taniguchi M, Yokota K, Kawai T. Identifying single nucleotides by tunnelling current. Nat. Nanotechnol. 2010; 4:286–290. [PubMed: 20305643]
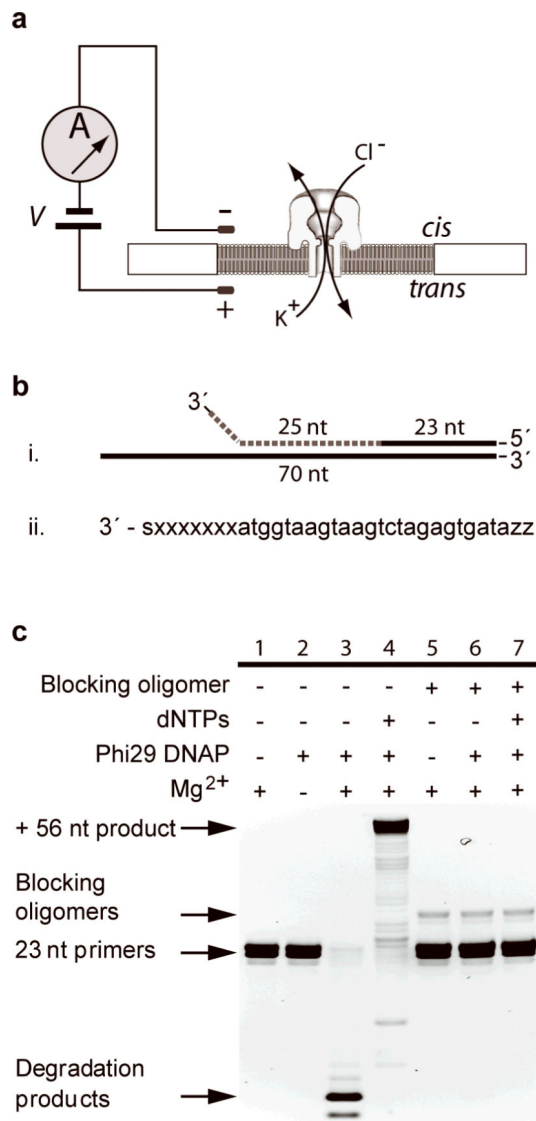
Figure 1.
Experimental set-up. (**a**) Nanopore device. A single α-HL nanopore is inserted in a lipid bilayer that separates two wells, each containing 100 μl of a buffered KCl solution. DNA bearing a ssDNA segment is added to the *cis* well. A voltage applied between the wells (*trans* side +) results 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 nucleotide primer annealed to a 70mer template strand (**Supplementary Fig. 1**). (ii) The blocking oligomer is composed of two parts: 1) a 25mer segment complementary to the DNA template near the p/t junction capped at its 5′ end by two acridines (z); and 2) a 3′- tail composed of 7 abasic residues (x) capped by a three-carbon spacer at its terminus (s) (**Supplementary Fig. 2**). (**c**) Test of DNA substrate protection in bulk phase using blocking oligomers. DNA substrates were incubated at 23 °C for five hours in nanopore buffer in the presence of phi29 DNAP, dNTPs, and Mg^{2+} as noted, with or without blocking oligomers. Fluorescently tagged primers were

separated by denaturing PAGE (see Online Methods). The arrows mark the full length product ('+56 nt extension'), the starting primer ('23nt primers') and products of exonucleolysis ('degradation products').
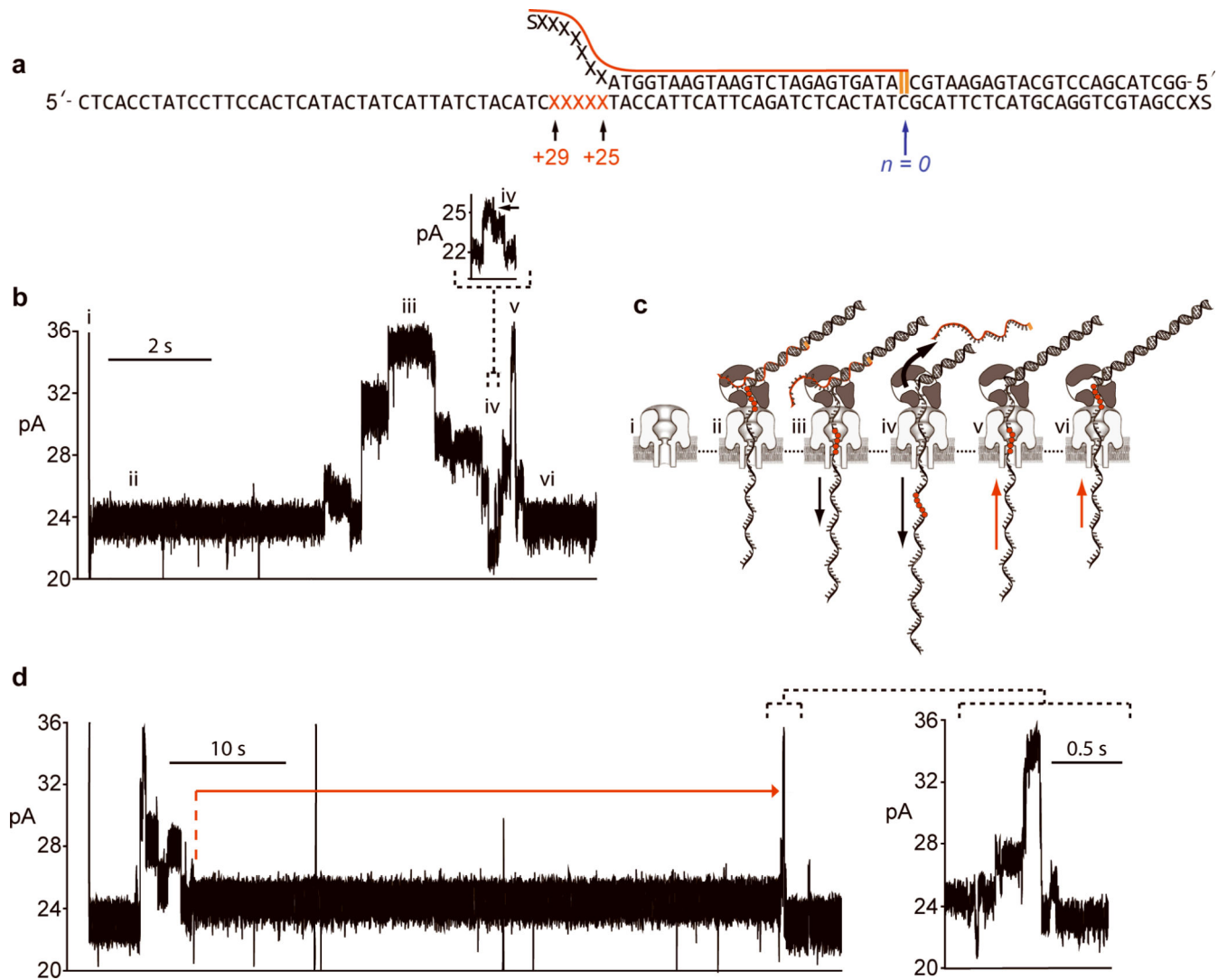
Figure 2.
Forward and reverse ratcheting of DNA through the nanopore. (**a**) DNA substrate protected by a blocking oligomer. A blocking oligomer (red line) capped by two acridine residues at its 5′ end protects the primer from catalysis in bulk phase. A 94mer DNA template contains five abasic residues (X's at positions 25-to-29) that cause an ionic current increase as they traverse the nanopore. (**b, c**) Forward and reverse ratcheting of DNA through the nanopore. Roman numerals i-vi of the ionic current trace in panel (b) correspond to cartoons i-vi in panel (c). See text for details. In the cartoon panels, the unzipped strand is shown passing through the enzyme, but this has not been demonstrated unequivocally. (**d**) Primer terminus chemistry controls replication and traversal of the second current peak. Conditions of this experiment were identical to those in panel (b) except that a 2′,3′-dideoxycytidine monophosphate (ddCMP) was substituted for 2′-deoxycytidine monophosphate (dCMP) at the primer 3′ terminus. The *ca.* 50 second segment of the current trace at 25 pA (horizontal red arrow) corresponds to position iv in panels (b) and (c). Replication fails to progress and advance the DNA template until phi29 DNAP excises and replaces the ddCMP nucleotide

with a dCMP on the primer terminus. At that point, replication proceeds (insert, dashed lines) generating a current peak as in position v of panel (b) and (c). These experiments were performed at 23 °C and 180 mV applied potential in nanopore buffer supplemented with 0.75 μM phi29 DNAP, 1 μM DNA substrate, 100 μM (each) dNTP, 10 mM Mg^{2+}, 1 mM DTT, and 1 mM EDTA.
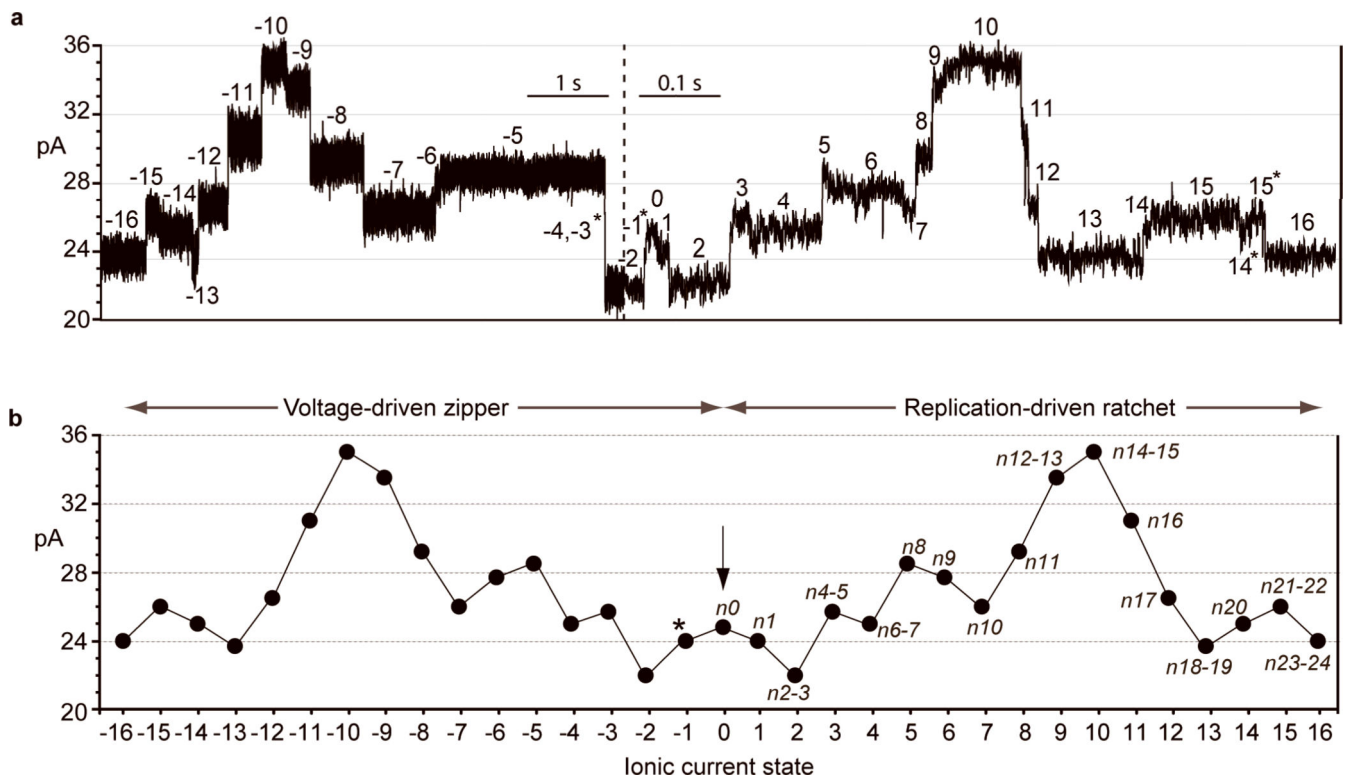
Figure 3.
Reproducible ionic current states as DNA is ratcheted through the nanopore. The DNA template and experimental conditions are identical to those in Figure 2. (**a**) Representative ionic current trace as a single DNA molecule is ratcheted through the pore. Thirty-three discrete amplitudes are highlighted: -16 to -1 for steps that occur during non-catalytic, voltage-driven unzipping of the blocking oligomer, 0 for the midpoint at 25 pA, and 1 to 16 for steps that occur during DNA replication. The vertical dashed line marks a timescale change required to see steps during DNA replication which is faster than mechanical unzipping. * denotes the -1 position which was absent in most traces; # denotes positions that were misread in this specific trace. (**b**) Reference map used to identify the current amplitudes shown in (a). Details on how the map was established are described in Supplementary Figure 4. All 25 nucleotide positions along the DNA template were accounted for within the 16 ionic current amplitudes of the replication-dependent peak. This was achieved by compiling ionic current amplitude pauses induced by independently diluting each of the dNTP substrates. The correlation between template positions and ionic current states is noted as n0 to n24 in italics.
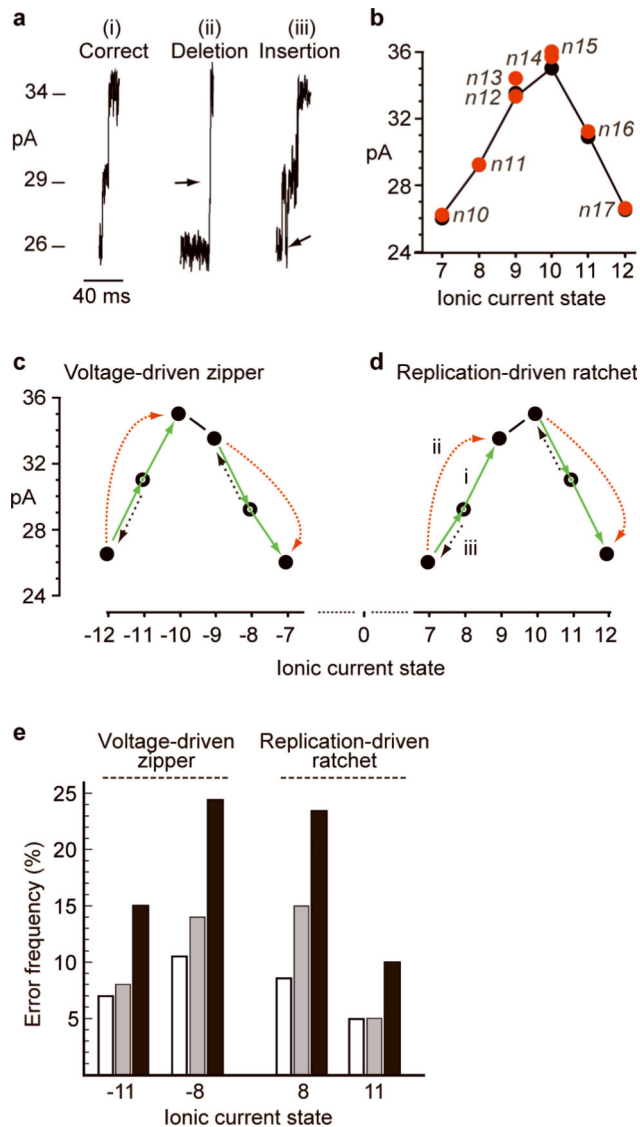
Figure 4.
Estimating DNA template registry errors in the nanopore during phi29 DNAP controlled translocation. The DNA template and experimental conditions are identical to those described in Figure 2. **(a)** Example ionic current traces. The traces shown are for the replication-driven ratchet moving DNA templates between ionic current state 7 (26 pA), state 8 (29 pA), and state 9 (34 pA). Movement from state 8 in either direction is equivalent to a one-nucleotide displacement. The data were filtered using a 2kHz low pass Bessel filter. (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 3ms (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 in the replication-dependent peak were confirmed

using a DNA mapping strategy described previously [12,14]. These values are represented by red circles. Numbers n10 to n17 (italics) 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 (voltage-driven zipper peak) and 7 to 12 (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 absent 'indel' errors. The grey dashed lines are when the current slips back one state before advancing properly. This results in an 'insertion'. The red dashed lines represent 'deletions' when the ionic current trace advances but does not register at the target state. 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 α–HL nanopore controlled by phi29 DNAP. Estimates are based on 200 events for the DNA construct illustrated in Figure 2a. Insertion error frequency (white bars); deletion error frequency (light grey bars); combined error frequency (dark grey bars).