Nucleic Acids Research, 2003, Vol. 31, No. 4 ***1311–1318***
DOI: 10.1093/nar/gkg218

Discrimination among individual Watson–Crick base pairs at the termini of single DNA hairpin molecules

Wenonah A. Vercoutere[1,2], Stephen Winters-Hilt[1,3], Veronica S. DeGuzman[1,2], David Deamer[1,2], Sam E. Ridino[1], Joseph T. Rodgers[1], Hugh E. Olsen[2], Andre Marziali[4] and Mark Akeson[1,2,3,*]

[1]Center for Biomolecular Science and Engineering, [2]Department of Chemistry and Biochemistry, and [3]Howard Hughes Medical Institute, University of California, Santa Cruz, CA 95064, USA and [4]Department of Physics and Astronomy, University of British Columbia, Vancouver, Canada

Received September 4, 2002; Revised and Accepted December 10, 2002

ABSTRACT

Nanoscale α-hemolysin pores can be used to analyze individual DNA or RNA molecules. Serial examination of hundreds to thousands of molecules per minute is possible using ionic current impedance as the measured property. In a recent report, we showed that a nanopore device coupled with machine learning algorithms could automatically discriminate among the four combinations of Watson–Crick base pairs and their orientations at the ends of individual DNA hairpin molecules. Here we use kinetic analysis to demonstrate that ionic current signatures caused by these hairpin molecules depend on the number of hydrogen bonds within the terminal base pair, stacking between the terminal base pair and its nearest neighbor, and 5′ versus 3′ orientation of the terminal bases independent of their nearest neighbors. This report constitutes evidence that single Watson–Crick base pairs can be identified within individual unmodified DNA hairpin molecules based on their dynamic behavior in a nanoscale pore.

INTRODUCTION

Atomic force microscopy, laser tweezers and fluorescence microscopy are established tools for examining single biological macromolecules (1–9). A new strategy has recently emerged in which individual DNA or RNA molecules are captured and examined within a single nanoscale pore formed by the α-hemolysin protein channel using impedance of monovalent ion current as the measured variable (10–20). Instruments based on this principle (Fig. 1A and B) permit serial examination of unmodified polynucleotide molecules at high speed (16). Discrimination has been achieved among individual poly(C), poly(A) and poly(U) RNA homopolymers (17) and among individual poly(dA)100 and poly(dC)100 DNA homopolymers (11). Segments within individual poly(A)30(C)70 RNA block copolymers have also been read (17) based on differences in secondary structure between the purine and pyrimidine blocks. It has been proposed that a more refined instrument may permit sequencing of long native DNA molecules at high speed (21,22).

We recently used an α-hemolysin nanopore to detect single nucleotide substitutions in duplex DNA using a series of blunt-ended hairpin molecules (14). The hairpin stems ranged in length from 3 to 9 bp. A downward trend in current amplitude was observed from ~82 pA for a 3 bp stem to ~38 pA for a 9 bp stem given an open channel current (I_o) of 120 pA. Beginning with the 8 bp hairpin, but more pronounced for the 9 bp hairpin, individual blockades gated between several discrete conductance states (Fig. 1C and D). We used a combination of Hidden Markov Models and Support Vector Machines to analyze these gating patterns (23), which allowed us to detect the identity and orientation of Watson–Crick base pairs at the termini of individual DNA hairpin molecules.

Here we use kinetic analysis to determine which DNA properties influence the observed sequence-dependent channel gating. Our data indicate that ionic current signatures caused by a given DNA hairpin molecule depend on the number of hydrogen bonds within the terminal base pair, stacking between the terminal base pair and its nearest neighbor, and 5′ versus 3′ orientation of the terminal bases, independent of their nearest neighbors. We propose a working model in which the low-frequency current transitions are due to orientation and binding of hairpins in the α-hemolysin vestibule, while the fast current spikes are due to fraying of the duplex terminus and subsequent penetration of one frayed strand into the pore-limiting aperture.

MATERIALS AND METHODS

DNA hairpin design, synthesis and purification

DNA oligonucleotides were synthesized using an ABI 392 Synthesizer, purified by PAGE and stored at –70°C in TE buffer. The sequence of the standard 9 bp hairpin was d(CTTCGAACGTTTTCGTTCGAAG). The base-pairing region is underlined. The prediction that each hairpin would adopt one base-paired structure was tested and confirmed using the DNA Mfold server (http://www.bioinfo.rpi.edu/

*To whom correspondence should be addressed. Tel: +1 831 459 5157; Fax: +1 831 459 2935; Email: makeson@chemistry.ucsc.edu

applications/mfold/old/dna/), which is based in part on data from SantaLucia (24).

Formation of α-hemolysin pores in horizontal bilayers

Each experiment was conducted using one α-hemolysin channel inserted into a diphytanoyl-phosphatidylcholine/hexadecane bilayer across a 25-micron-diameter horizontal Teflon aperture, as described previously (17). Seventy microliter chambers on either side of the bilayer contained 1.0 M KCl buffered at pH 8.0 (10 mM HEPES/KOH). Voltage was applied across the bilayer between Ag-AgCl electrodes. DNA was added to the *cis* chamber at a final concentration of 10 or 20 μM. All experiments were maintained at room temperature (23 ± 0.1°C) using a Peltier device.

Data acquisition

Ionic current was filtered at 10 kHz bandwidth using an analog low-pass Bessel filter and recorded at 20 μs intervals (50 kHz) using an Axopatch 200B amplifier (Axon Instruments, Foster City, CA) coupled to an Axon Digidata 1200 digitizer. The applied potential was 120 mV (*trans* side positive) unless otherwise noted. Semi-automated analysis of transition level blockade current and duration was performed using Fetchan 6.01 (Axon Instruments). Fetchan superimposes and records an idealized trace on the data, approximating amplitude and duration of each transition. After setting threshold levels, we used a non-stop level detection mode to measure and acquire events. A change in amplitude was accepted as a transition to another level if it had a magnitude greater than or equal to five times the root mean square (r.m.s.) noise at 10 kHz bandwidth of the original level (25). Based on a Gaussian distribution of signal noise, the probability that random noise would cause the signal to exceed its r.m.s. by 5-fold was <1%.

Data analysis

Amplitudes for each experiment were determined from a first-order Gaussian fit to a histogram of at least 1000 events for each level, binned at 0.2 pA/bin. In 1.0 M KCl at 23.0°C, I_o was 120 ± 3 pA. In unusual cases (<10% of channels), the toggle rate and amplitude difference between conductance states caused by the 9 bp hairpins were measurably different from the representative events shown in Figure 2. Data from those channels were not included in the analysis. Mean lifetimes were calculated using Clampfit (Axon Instruments). To calculate the average lifetime of each state, duration measurements were plotted in semi-log frequency histograms

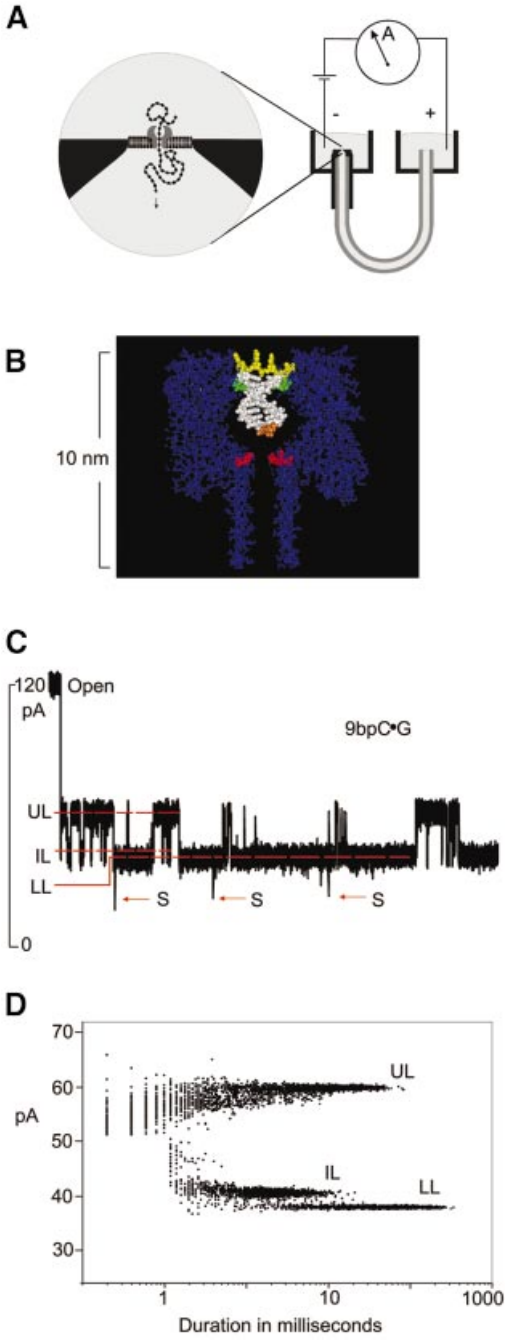

Figure 1. A nanopore device based on the α-hemolysin channel. (**A**) Diagram of a horizontal bilayer apparatus developed in our laboratory. One α-hemolysin channel is intercalated in a horizontal diphytanoyl phosphatidylcholine bilayer. The bilayer is supported on a 25-micron-diameter conical aperture at the end of a U-shaped Teflon tube. The tube connects two 70 μl vol baths filled with 1 M KCl buffered at pH 8.0 in 10 mM HEPES/KOH. Voltage is applied between two Ag-AgCl electrodes that are connected to an Axon 200B amplifier. (**B**) Two-dimensional diagram of a 9 bp hairpin captured in the pore vestibule. The stick figure in blue is a two-dimensional section of the α-hemolysin pore derived from X-ray crystallographic data (27). A ring of lysines that circumscribe a 1.5-nm-limiting aperture of the channel pore is highlighted in red. A ring of threonines that circumscribe the narrowest 2.3-nm-diameter section of the pore mouth is highlighted in green. In our working model, the four dT hairpin loop (yellow) is perched on this narrow ring of threonines, suspending the duplex stem in the pore vestibule. The terminal base pair (brown) dangles near the limiting aperture. The structure of the 9 bp hairpin shown here was rendered to scale using WebLab ViewerPro. (**C**) Representative blockade of ionic current caused by a 9 bp DNA hairpin (9bp**C**•G). I_o is typically 120 pA at 120 mV and 23.0°C, but this may vary by ±3 pA. Capture of a DNA hairpin causes a rapid current decrease. In the case of 9 bp hairpins, the residual current transitions between four levels: an upper conductance level (UL), an intermediate level (IL), a lower level (LL) and a transient downward spike (S). (**D**) A two-dimensional plot of duration versus current amplitude for UL, IL and LL conductance states.

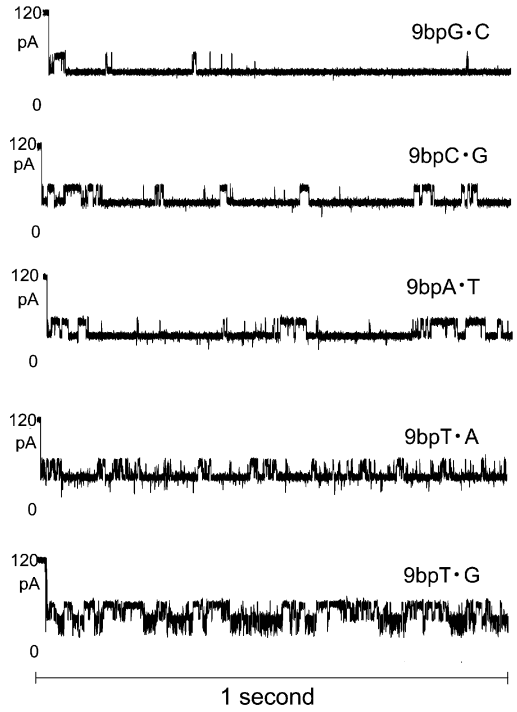

Figure 2. Blockade of the α-hemolysin pore by 9 bp DNA hairpins in which the terminal base pair is varied. Blockade events were acquired at 120 mV applied potential and 23.0°C (Materials and Methods). Each signature shown is caused by a single hairpin molecule captured in the pore vestibule, and is representative of several thousand single molecule events.

with 20 bins per decade. The minimum duration of a transition that can be accurately measured in a signal is determined by the rise time, T_r (~$0.33/f_c$), which is ~33 μs at 10 kHz (26); therefore, durations <40 μs were ignored in this analysis. At least 1000 measurements of duration were used for each plot. Curve fitting to determine the probability density function was performed using the Levenberg–Marquardt method, with no function weighting. The R^2 value and standard deviation of the function were used to evaluate goodness of fit.

Molecular modeling

The molecular model of α-hemolysin combined with the 9 bp hairpin shown in Figure 1B was prepared using the crystal structure of α-hemolysin (27) and Chem-Site (28).

RESULTS

The DNA hairpins examined in this study are shown in Table 1. Among these are 9 bp hairpins with the four possible combinations of terminal base-pair identity and orientation (9bpG•C, 9bpC•G, 9bpA•T and 9bpT•A). In our nomenclature, the letter before the dot is the 5′ nucleotide of the single-stranded DNA 22mer that forms each 9 bp DNA hairpin, and the letter after the dot is the 3′ nucleotide of the same strand. Together, those two nucleotides constitute the base pair at the terminus of the hairpin duplex stem. Current blockade signatures selected from thousands of events caused by each of these four classes of molecules are shown in Figure 2. Four conductance states are readily observed (Fig. 1C and D): (i) an intermediate level (IL) that initiates all 9 bp hairpin events; (ii) an upper conductance level (UL); (iii) a lower conductance level (LL) that must be preceded by the upper level; and (iv) spikes down from the lower level that indicate close proximity of the terminal base pair to the pore-limiting aperture. A working model to explain these conductance states is presented in the Discussion.

Terminal base-pair identity can be determined by kinetic analysis of dwell times in these conductance states. In particular, the average dwell time in the lower conductance level (LL in Fig. 1C and D) and the frequency of downward current spikes from the lower conductance level (S in Fig. 1C) are highly dependent upon the presence of a stable base pair at the duplex terminus. This is illustrated in Figure 3, where neither a 5′ dC dangling nucleotide nor a 3′ dG dangling nucleotide alone stabilized ionic current in the lower level (I = 38 pA, I_o = 120 pA), whereas both nucleotides together (the C•G pair) did. We reasoned that the presence of any two nucleotides at the terminus of the hairpin stem might account for this current stabilization. However, two weakly paired thymine bases at the blunt-end terminus of a 9 bp hairpin stem resulted in an unstable blockade signature (Fig. 3), confirming the requirement for a Watson–Crick base pair to yield a steady lower level conductance. In practice, the LL has the added advantage that transitions to the UL are stochastic, and that one first-order exponential can be fitted to the dwell time distribution giving a time constant (τ_{LL}) in the millisecond range.

The sensitivity of the lower conductance state to Watson–Crick base-pair identity was tested by measuring

Table 1. DNA hairpins used in this study

TT T T G C C G A T A T G C C G T A T A G C 5′ 3′	TT T T G C C G A T A T G C C G T A T A C G	TT T T G C C G A T A T G C C G T A A T	TT T T G C C G A T A T G C C G T A T A	TT T T G C C G A T A T G C C G T A T A T G	TT T T G C C G A T A T G C C G T A T T	TT T T G C C G A T A T G C C G T A F A	TT T T G C C G A T A T G C C G T A A T G C	TT T T G C C G A T A T G C C G T A A T C G	TT T T G C C G A T A T G C C G T A A T A T	TT T T G C C G T T T C G A A T T A
9bpG•C	9bpC•G	9bpA•T	9bpT•A	9bpT•G	9bpT•T	9bpF•A	9bp(AT)G•C	9bp(AT)C•G	9bp(AT)A•T	9bp(AT)T•A

Primary sequence reads from the 5′ end at bottom left to the 3′ end at bottom right. Each hairpin has a 9 bp long stem and a four dT loop. The terminal base pair at the end of each hairpin stem is in close proximity to the pore-limiting aperture when a given hairpin is captured in the α-hemolysin vestibule.

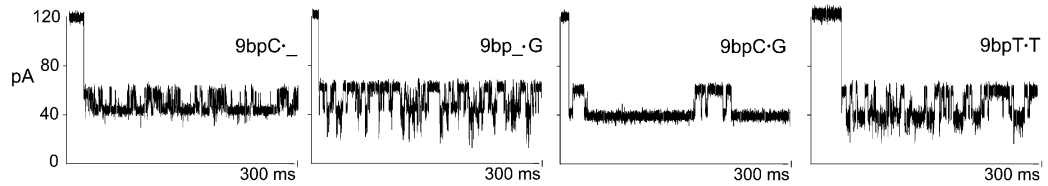

Figure 3. Comparison of blockade signatures caused by DNA hairpins with dangling and blunt ends. All hairpins were built onto a core 8 bp DNA hairpin with the primary sequence d(TTCGAACGTTTTCGTTCGAA). 9bpC•_ shows a blockade signature caused by a hairpin with a dangling 5′-C nucleotide. 9bp_•G shows a blockade signature caused by a dangling 3′-G nucleotide. 9bpC•G shows a blockade signature for a hairpin in which both terminal nucleotides are present forming a 5′-C•G-3′ terminal Watson–Crick base pair. 9bpT•T shows a typical blockade signature for a blunt-ended 9 bp hairpin in which the terminal 5′-T•T-3′ pair is weakly associated. Experimental conditions are described in Materials and Methods.

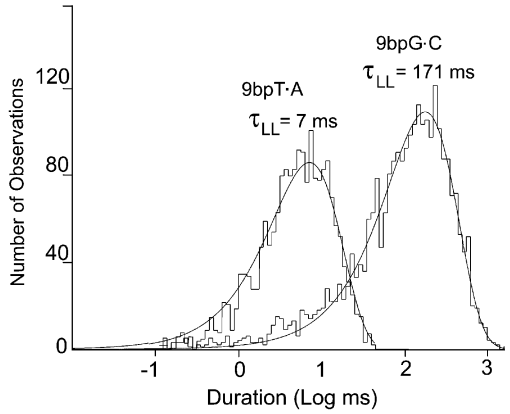

Figure 4. Dwell time histograms for LL blockade events. Duration measurements were plotted in semi-log frequency histograms with 20 bins per decade. At least 1000 measurements of duration were used for each plot. To determine the probability density function and the average event lifetime, τ_{LL}, curves were fitted to each histogram using the Levenberg–Marquardt method. 9bpT•A is the standard 9 bp hairpin with a 5′-T•A-3′ terminus and 9bpG•C is a 9 bp hairpin with a 5′-G•C-3′ terminus.

τ_{LL} and spike frequency for the four 9 bp hairpins whose blockade signatures are illustrated in Figure 2. Dwell time histograms for the lower conductance state caused by 9bpG•C and 9bpT•A are shown in Figure 4. First-order exponentials fitted to similar histograms for all four permutations of Watson–Crick base-pair terminal ends reveal τ_{LL} values ranging from 160 to 7 ms in the order 9bpG•C > 9bpC•G > 9bpA•T > 9bpT•A (Table 2). The reverse order is observed for spike frequency ranging from 4 spikes s^{-1} (9bpG•C) to 91 spikes s^{-1} (9bpT•A). Thus, two kinetic parameters can be used to discriminate among Watson–Crick base pairs on single DNA hairpin molecules, confirming the pattern recognition results established previously (23).

One of the more difficult base pairs to recognize using conventional hybridization assays is a terminal mismatch, in particular a TG wobble pair. We tested the sensitivity of the nanopore to this mismatch by comparing blockade signatures caused by a hairpin composed of the sequence 9bpT•G with blockade signatures caused by the perfectly complementary sequences 9bpC•G and 9bpT•A (Fig. 2). In this experiment, all individual blockades that exhibited the characteristic four current level signature could be identified as one of these molecules. Quantitative examination of the data revealed that spike frequency was the key diagnostic parameter: there was a significant difference between spike frequencies caused by each of the three termini, i.e. 12 spikes s^{-1} (9bpC•G), 91 spikes s^{-1} (9bpT•A) and 1400 spikes s^{-1} (9bpT•G) (Table 2). In contrast, τ_{LL} values were statistically different between 9bpT•G and 9bpC•G termini, but not between 9bpT•G and 9bpT•A termini (Table 2). It appears that τ_{LL} values plateau in the low millisecond time range.

Table 2. Comparison between single DNA hairpin kinetic parameters and $\Delta\Delta G°$ for terminal base pairs

Identity	τ_{LL} (ms)	Spike frequency (s^{-1})	$\Delta\Delta G°_{23}$ (kcal/mol)
9bpG•C	160 ± 23	4 ± 1	–1.9
9bpC•G	50 ± 4	12 ± 4	–1.8
9bpA•T	43 ± 5	34 ± 10	–1.2
9bpT•A	7 ± 1	91 ± 47	–1.3
9bpT•G	6 ± 2	1300 ± 400	–0.3

$\Delta\Delta G°_{23}$ values are the difference between calculated $\Delta G°$ of duplex formation for 9 bp DNA hairpins and calculated $\Delta G°$ of duplex formation for core 8 bp hairpins that lack the terminal base pair. Calculations assumed 23.0°C and 1 M KCl. They were performed using Mfold (http://www.bioinfo.rpi.edu/applications/mfold/old/dna/) which is based on data from SantaLucia (24). Spike frequency and τ_{LL} values are mean ± SE for at least three experiments using different individual channels.

It was important to determine whether the rankings of spike frequency and τ_{LL} correlated with conventional estimates of terminal base-pair stability. Table 2 lists free energy values for terminal base pairs ($\Delta\Delta G°_{23}$) calculated using the online computational tool Mfold (http://www.bioinfo.rpi.edu/applications/mfold/old/dna/), which is based on a nearest neighbor model of duplex stability (24). This model is particularly strong because it considers data from seven independent studies (29–35). In Table 2, the $\Delta\Delta G°_{23}$ values are the difference between the free energy of duplex formation for a given 9 bp hairpin and the free energy of duplex formation of a common 8 bp core hairpin sequence at 23°C. Among Watson–Crick base pairs, $\Delta\Delta G°_{23}$ values ranged from –1.9 kcal/mol for 9bpG•C to –1.2 kcal/mol for 9bpA•T. $\Delta\Delta G°_{23}$ for the T•G wobble pair was calculated to be –0.3 kcal/mol. In general, the rank of spike frequency and τ_{LL} correlated with $\Delta\Delta G°_{23}$; however, the correlation is imperfect in that the expected order of 9bpT•A and 9bpA•T was reversed.

Hydrogen bond and base stacking each influence τ_{LL} and spike frequency

Having established a general correlation between the nanopore data and classical measures of base-pair stability, we determined whether non-covalent forces that contribute to

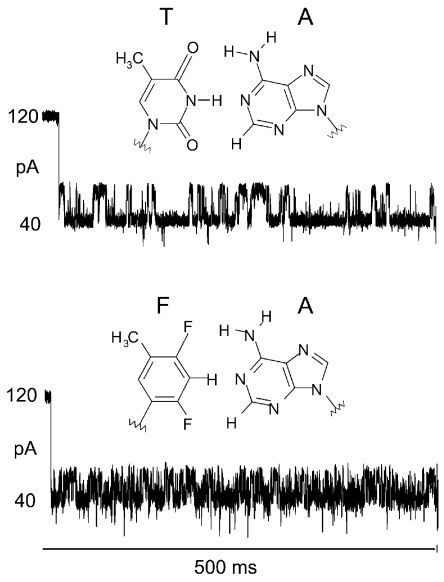

Figure 5. Effect of difluorotoluene (F) substitution for thymine (T) on blockades caused by 9 bp hairpins. The blockade signature at the top is caused by a 9 bp hairpin with a 5′-T•A-3′ terminus (9bpT•A in Table 1). The blockade signature at the bottom is caused by a nearly identical 9 bp hairpin in which the 5′ thymine is replaced by difluorotoluene (9bpF•A in Table 1) giving a 5′-F•A-3′ terminus that lacks hydrogen bonds. The blockade signatures shown are representative of thousands of single molecule events acquired under standard conditions (Materials and Methods).

DNA duplex stability could be discerned. Forces that stabilize DNA duplexes include hydrogen bonding between bases, and base stacking. Forces that destabilize DNA duplexes include hydrogen bonding between water molecules and nucleotide bases, and electrostatic repulsion between phosphodiester anions in the DNA backbone. Steric effects may stabilize or destabilize the duplex depending upon sequence context (36).

Initial inspection of the data in Table 2 suggests that hydrogen bonding plays a significant role in spike frequency and τ_{LL}: i.e. terminal base pairs that are known to form three hydrogen bonds when paired (5′-G•C-3′ and 5′-C•G-3′) have lower spike frequencies and greater τ_{LL} values than terminal base pairs that form two hydrogen bonds when paired (5′-A•T-3′, 5′-T•A-3′ and 5′-T•G-3′). To test directly the influence of hydrogen bonding on these kinetic parameters, we compared current signatures caused by 9bpT•A with those caused by a hairpin with a difluorotoluene (F)/adenine terminus (9bpF•A). Difluorotoluene is a good structural mimic of thymine (36,37) that is recognized by DNA polymerases despite the presumed absence of hydrogen bonding to paired adenines (38,39). Blockade current signatures are illustrated in Figure 5. Reduction of hydrogen bonding by the T•A→F•A substitution caused destabilization of the lower conductance state reflected in a decreased average dwell time in that state (τ_{LL} = 2 ms) and an increased spike frequency (290 ± 10 s^{-1}) relative to the T•A control (Table 2).

The data in Table 2 also indicate that orientation of the bases in the terminal pair influences spike frequency and τ_{LL}: i.e. flipping the terminal base pair so that a purine is on the 5′ side and a pyrimidine is on the 3′ side (9bpC•G→9bpG•C and 9bpT•A→9bpA•T) consistently increased τ_{LL} and decreased spike frequency. Among Watson–Crick base pairs, the size of this effect equals or exceeds the effect of increasing the hydrogen bond number (Table 2). Classical thermodynamic studies suggest two possible explanations: (i) stacking forces with the neighboring base pair are altered when the terminal base pair is flipped (24,33); and (ii) stacking of bases at the 5′ position of a duplex can be different from those at the 3′ position independent of the neighboring base pair (40). A third possible explanation is that the terminal nucleotide on the 3′ or 5′ side of the duplex stem interacts with an amino acid in the vestibule wall and that this interaction is nucleotide dependent.

Table 3. Effect of penultimate base-pair orientation on τ_{LL} for 9 bp hairpins with different Watson–Crick base-pair termini and orientations

	Penultimate base pair	
	5′-T•A-3′	5′-A•T-3′
Terminal base pair	τ_{LL} ± SE (ms)	
5′-T•A-3′	7 ± 1	20 ± 4
5′-A•T-3′	43 ± 5	30 ± 6
5′-G•C-3′	160 ± 23	210 ± 90
5′-C•G-3′	50 ± 4	66 ± 20

Values shown represent mean ± SE for at least three different individual channels. Experimental conditions are described in Materials and Methods.

To test the first explanation, we compared τ_{LL} for the standard 9 bp hairpins containing the four possible Watson–Crick termini (Table 1, left) with their counterparts in which the penultimate TA base pair was flipped, i.e. hairpins 9bp(AT)T•A, 9bp(AT)A•T, 9bp(AT)C•G and 9bp(AT)G•C (Table 1, right). 9bpT•A was the least stable of the original sequences with τ_{LL} equal to 7 ms. By making the penultimate base-pair substitution 9bp(TA)T•A→9bp(AT)T•A, τ_{LL} was increased ~3-fold to 20 ms (Table 3). Conversely, 9bpA•T was the most stable of the thymidine/adenine termini with τ_{LL} equal to 43 ms. By making the same alteration of the neighboring base pair as in the previous experiment, 9bp(TA)A•T→9bp(AT)A•T, τ_{LL} decreased to 30 ms. Thus, stacking against the neighboring base pair did account for some of the stability difference associated with orientation of the thymine/adenine termini. The independent effect of placing adenine at the 5′ position was also important by inference. For the guanine/cytosine termini, the outcome was different (Table 3). In those cases, flipping penultimate base pairs did not significantly influence τ_{LL}. Thus, the 3-fold difference in τ_{LL} for 5′-G•C-3′ versus 5′-C•G-3′ is due to an end-specific effect independent of the neighboring base pair. Importantly, the general trend in spike frequency correlated with calculated terminal base pair stability even when the penultimate base pair was reversed (Fig. 6).

DISCUSSION

The main point of this study is that kinetic analysis can be used to discriminate between Watson–Crick base pairs at the termini of individual DNA hairpin molecules captured in a nanoscale pore. The orientation of each terminal base pair is also distinguishable, i.e. 5′-G•C-3′ versus 5′-C•G-3′ and 5′-A•T-3′ versus 5′-T•A-3′. Calculated base-pair stabilities correlate reasonably well with the kinetic properties we

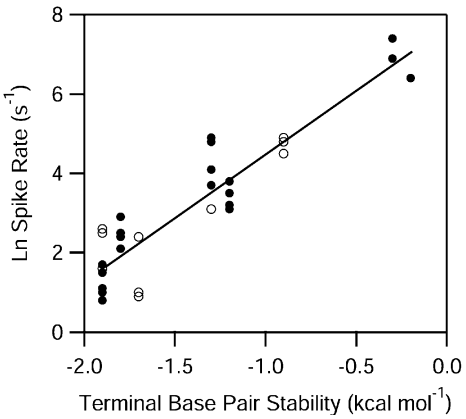

Figure 6. Correlation between the stability of hairpin terminal base pairs and the frequency of downward current spikes. Terminal base-pair stabilities used are $\Delta\Delta G^{\circ}_{23}$ values calculated by taking the difference between the free energy of duplex formation for a given 9 bp hairpin and the free energy of duplex formation of the appropriate 8 bp core hairpin sequence. These free energy values were calculated using Mfold, assuming 23°C and 1 M salt. $\Delta\Delta G^{\circ}_{23}$ values for hairpins with penultimate T•A base pairs (filled circles) are as follows: 9bp(TA)A•T = –1.2 kcal/mol; 9bp(TA)T•A = –1.3 kcal/mol; 9bp(TA) C•G = –1.8 kcal/mol; 9bp(TA)G•C = –1.9 kcal/mol; 9bp(TA)T•G = –0.3 kcal/mol; 9bp(TA) T•T = –0.2 kcal/mol. $\Delta\Delta G^{\circ}_{23}$ values for hairpins with penultimate A•T base pairs (open circles) are as follows: 9bp(AT)A•T = –1.3 kcal/mol; 9bp(AT)T•A = –0.9 kcal/mol; 9bp(AT)C•G = –1.9 kcal/mol; 9bp(AT)G•C = –1.7 kcal/mol. Each data point represents the spike rate for a given hairpin molecule for a channel prepared on a separate day. The line represents a least squares fit to the transformed data ($y = 3.2x + 7.7$, $R^2 = 0.85$).

examined (Tables 2 and 3; Fig. 6), however the correlation is imperfect. For example, the expected order of 9bpT•A and 9bpA•T based on Mfold stabilities was reversed for the nanopore data. Possible explanations for this discrepancy include uncertainty surrounding the predicted stability of terminal 5′-A•T-3′ and 5′-T•A-3′ pairs (24,33), and limits on the precision of optical melting curves that underlie the free energy calculations. We also note that τ_{LL} and spike frequency are likely to be influenced by the electric field, and by interactions between the duplex terminus and amino acids in the vestibule wall. The magnitude of these effects could be sequence dependent, thus altering the stability ranking in the nanopore assay relative to a bulk solution assay.

Our data support the premise that hydrogen bonding between bases and base stacking each influence stability of the terminal base pair captured in the pore. However, in practice it is difficult to quantitatively separate these two properties. This is illustrated by comparing the TG wobble pair and the CG Watson–Crick base pair. In substituting a 5′ thymidine for a 5′ cytosine at the 9 bp hairpin terminus, the hydrogen bond number is reduced from three to two, but stacking energy is stabilized by –0.1 kcal/mol as shown by melting curves for DNA duplexes with dangling ends (40). Although small, this change in stacking energy is comparable to calculated differences in $\Delta\Delta G^{\circ}_{23}$ between some of the terminal base pairs in Table 2. Thus, the change in blockade signature associated with the CG→TG terminal substitution results from the combined effect of added stacking stabilization due to thymine and destabilization due to loss of hydrogen bond energy. Competing effects are also likely when the thymine in the terminal TA base pair is replaced by difluorotoluene (9bpTT/AA→9bpFT/AA). This isostere would be an ideal tool to assess the influence of hydrogen bonding on τ_{LL} and spike frequency; however, difluorotoluene is non-polar and its stacking interaction with neighboring bases stabilizes the duplex by –1.5 kcal/mol relative to thymine (41). Thus, the blockade signature due to 9bpT•A→9bpF•A (Fig. 5) is a conservative measure of destabilization due to loss of hydrogen bonding because it is partially offset by increased stacking stability.

We have developed a working model to explain the current transitions caused by hairpin capture (Fig. 7). Because this model includes DNA interaction with the pore wall and with the electric field, it predicts that dynamic properties of DNA base pairs measured in the α-hemolysin pore (e.g. fraying) will differ somewhat from those dynamic properties in bulk phase. In our model, each 9 bp hairpin is captured so that the terminal base pair can interact with amino acids in the vestibule wall near the limiting aperture formed by lysine 147 and glutamate 111 of α-hemolysin. Circular dichroism assays (data not shown) indicate that the 9 bp hairpin stem is a B-form duplex in bulk phase. The length per base pair of B-form DNA is 3.4 Å, therefore the total stem length would be 30.6 Å. The distance between the narrowest part of the vestibule mouth at threonine 9 and the pore-limiting aperture at lysine 147/glutamate 111 is 33 Å. Therefore, if the hairpin loop is perched at the ring formed by threonine 9, the 9 bp stem would reach at least within 3 Å of the limiting aperture. Given our uncertainty about the exact position of the hairpin loop, and the finite precision of the α-hemolysin X-ray crystal structure [1.9 Å (27)], the estimated position of the hairpin terminus is probably accurate within ±1 bp or ±3.4 Å.

Upon capture of a 9 bp blunt hairpin, the IL state is caused by orientation and immobilization of the hairpin terminus due to an electrostatic interaction between the terminal base pair and residues in the vestibule wall (Fig. 7A). The dwell time, τ_{IL}, for this intermediate conductance state is largely independent of base-pair identity or orientation (data not shown), presumably because the bases are hydrogen bonded to one another and the interaction with the surface is electrostatic. The IL state invariably transitions to the upper conductance state, UL (Fig. 7B). In the model, this state corresponds to desorption of the terminal base pair from the protein wall and thermal motion of the hairpin stem in orientations that allow greater ion current flow through the limiting aperture. These orientations may be angular displacement of the hairpin terminus away from the limiting aperture, or azimuthal orientation of the molecule allowing ion current to flow along the major groove of the duplex stem.

From the UL conductance state, the hairpin may return to the IL state or it may transition into a third conductance state, LL (Fig. 7C). Residence time in this state is dependent upon terminal base-pair identity and orientation. In this state, we hypothesize that one or both terminal nucleotides of the duplex stem are adsorbed to the protein in close proximity to the pore-limiting aperture. Thus, the terminal base pair interacts strongly with the ion current through the limiting aperture. When the duplex end frays from this bound state (Fig. 7D), one strand may extend and penetrate the limiting aperture resulting in a transient spike (Fig. 7E).

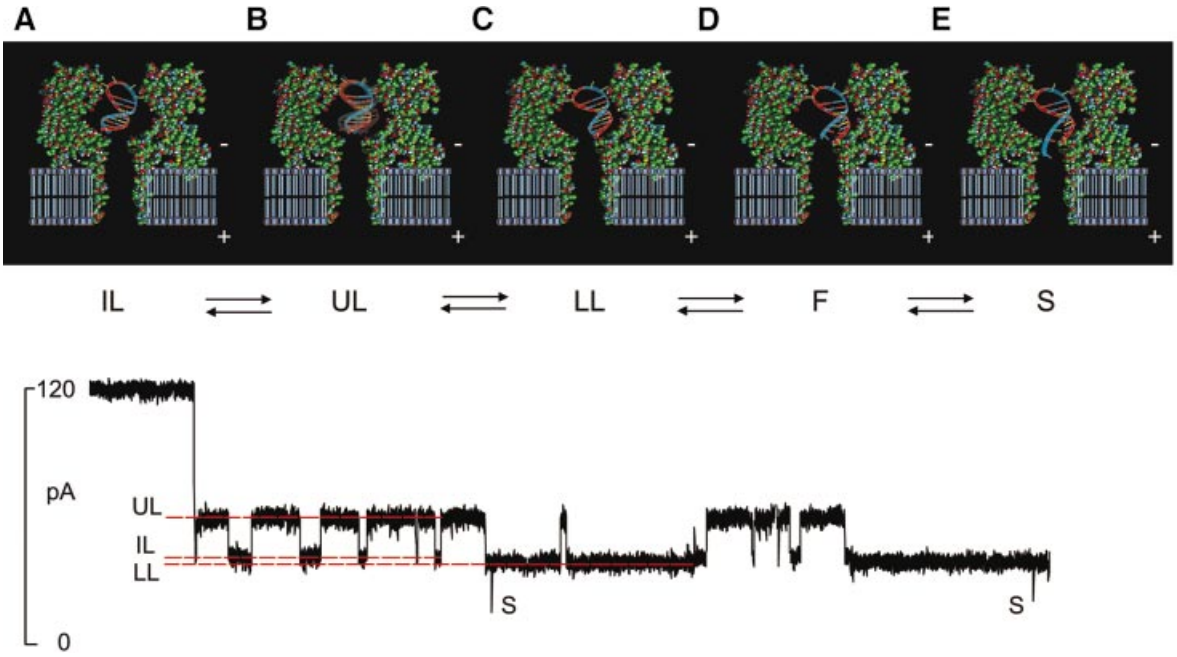

Figure 7. Working model of 9 bp DNA hairpin capture in the α-hemolysin vestibule under an applied voltage. (**A**) The IL conductance state. One of the terminal nucleotides interacts with amino acids in the vestibule wall. (**B**) The UL conductance level, where the terminus has desorbed from the pore wall and is free to explore orientations within the vestibule that allow increased current through the pore. (**C**) The LL conductance level. In this state, one or both nucleotides of the duplex terminus interact with amino acids such that one or both nucleotides are poised directly over the pore-limiting aperture. When one or more base pairs at the duplex terminus fray from this position (**D**), one of the two strands extends from its helical conformation and penetrates the pore-limiting aperture (**E**). This causes transient spikes to ~14 pA residual current.

CONCLUSIONS

The use of nanopores to characterize single nucleic acid molecules has grown substantially since the strategy was first reported in 1996 (16). This report demonstrates that impedance of ionic current through a nanoscale pore formed by α-hemolysin can be used to read the identity and orientation of single Watson–Crick base pairs. Nanopore technology may have practical uses for DNA analysis, including high speed sequencing (21,22).

ACKNOWLEDGEMENTS

We wish to thank Sutter Instruments (Novato, CA) for help in manufacturing quartz glass tips used in preparing the bilayer apertures. This work was supported by grants from the National Human Genome Research Institute and from the Howard Hughes Medical Institute. A gift from Agilent Corporation was used to purchase key components of the nanopore instrument.

REFERENCES

1. Bennink,M.L., Scharer,O.D., Kanaar,R., Sakata-Sogawa,K., Schins,J.M., Kanger,J.S., de Grooth,B.G. and Greve,J. (1999) Single-molecule manipulation of double-stranded DNA using optical tweezers: interaction studies of DNA with RecA and YOYO-1. *Cytometry*, **36**, 200–208.
2. Goodwin,P.M., Johnson,M.E., Martin,J.C., Ambrose,W.P., Marrone,B.L., Jett,J.H. and Keller,R.A. (1993) Rapid sizing of individual fluorescently stained DNA fragments by flow cytometry. *Nucleic Acids Res.*, **21**, 803–806.
3. Perkins,T.T., Quake,S.R., Smith,D.E. and Chu,S. (1994) Relaxation of a single DNA molecule observed by optical microscopy. *Science*, **264**, 822–826.
4. Perkins,T.T., Smith,D.E. and Chu,S. (1994) Direct observation of tube-like motion of a single polymer chain. *Science*, **264**, 819–822.
5. Perkins,T.T., Smith,D.E., Larson,R.G. and Chu,S. (1995) Stretching of a single tethered polymer in a uniform flow. *Science*, **268**, 83–87.
6. Rief,M., Clausen-Schaumann,H. and Gaub,H.E. (1999) Sequence-dependent mechanics of single DNA molecules. *Nature Struct. Biol.*, **6**, 346–349.
7. Smith,S.B., Cui,Y. and Bustamante,C. (1996) Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, **271**, 795–799.
8. Smith,D.E., Babcock,H.P. and Chu,S. (1999) Single-polymer dynamics in steady shear flow. *Science*, **283**, 1724–1727.
9. Wang,M.D., Yin,H., Landick,R., Gelles,J. and Block,S.M. (1997) Stretching DNA with optical tweezers. *Biophys. J.*, **72**, 1335–1346.
10. Meller,A. and Branton,D. (2002) Single molecule measurements of DNA transport through a nanopore. *Electrophoresis*, **23**, 2583–2591.
11. Meller,A., Nivon,L., Brandin,E., Golovchenko,J. and Branton,D. (2000) Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl Acad. Sci. USA*, **97**, 1079–1084.
12. Meller,A., Nivon,L. and Branton,D. (2001) Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.*, **86**, 3435–3438.
13. Nakane,J., Akeson,M. and Marziali,A. (2002) Evaluation of nanopores as candidates for electronic analyte detection. *Electrophoresis*, **23**, 2592–2601.
14. Vercoutere,W., Winters-Hilt,S., Olsen,H., Deamer,D., Haussler,D. and Akeson,M. (2001) Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat. Biotechnol.*, **19**, 248–252.
15. Wang,H. and Branton,D. (2001) Nanopores with a spark for single-molecule detection. *Nat. Biotechnol.*, **19**, 622–623.
16. Kasianowicz,J.J., Brandin,E., Branton,D. and Deamer,D.W. (1996) Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl Acad. Sci. USA*, **93**, 13770–13773.
17. Akeson,M., Branton,D., Kasianowicz,J.J., Brandin,E. and Deamer,D.W. (1999) Microsecond time-scale discrimination among polycytidylic acid,

polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.*, **77**, 3227–3233.
18. Henrickson,S.E., Misakian,M., Robertson,B. and Kasianowicz,J.J. (2000) Driven DNA transport into an asymmetric nanometer-scale pore. *Phys. Rev. Lett.*, **85**, 3057–3060.
19. Howorka,S., Cheley,S. and Bayley,H. (2001) Sequence-specific detection of individual DNA strands using engineered nanopores. *Nat. Biotechnol.*, **19**, 636–639.
20. Akeson,M., Deamer,D.W., Vercoutere,W., Braslau,R. and Olsen,H. (2002) Use of a nanoscale pore to read short segments within single polynucleotide molecules. In Kasianowicz,J.J., Kellermeyer,M.S.Z. and Deamer,D.W. (eds), *NATO Advanced Research Workshop on Polymer Structure and Transport in Confined Spaces*. Kluwer Academic, Bikal, Hungary, Vol. 87, p. 390.
21. Deamer,D.W. and Branton,D. (2002) Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.*, **35**, 817–825.
22. Voss,D. (1999) Gene Express. *New Scientist*, **164**, 40–43.
23. Winters-Hilt,S., Vercoutere,W., DeGuzman,V., Deamer,D., Akeson,M. and Haussler,D. (2003) Highly accurate classification of Watson–Crick base-pairs on termini of single DNA molecules. *Biophys. J.*, in press.
24. SantaLucia,J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
25. Colquhoun,D. and Sigworth,F.J. (1983) Fitting and statistical analysis of single-channel recordings. In Sakmann,B. and Neher,E. (eds), *Single-Channel Recording*. Plenum Press, New York, NY, pp. 191–263.
26. Colquhoun,D. and Hawkes,A.G. (1983) The principles of stochastic interpretation of ion-channel mechanisms. In Sakmann,B. and Neher,E. (eds), *Single-Channel Recording*. Plenum Press, New York, NY, pp. 135–175.
27. Song,L., Hobaugh,M.R., Shustak,C., Cheley,S., Bayley,H. and Gouaux,J.E. (1996) Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. *Science*, **274**, 1859–1866.
28. Michael,D. (1999) Chem-Site (3.01 edn). Pyramid Learning LLC, Hudson, OH.
29. Delcourt,S.G. and Blake,R.D. (1991) Stacking energies in DNA. *J. Biol. Chem.*, **266**, 15160–15169.
30. Doktycz,M.J., Goldstein,R.F., Paner,T.M., Gallo,F.J. and Benight,A.S. (1992) Studies of DNA dumbbells. 1. Melting curves of 17 DNA dumbbells with different duplex stem sequences linked by T4 endloops — evaluation of the nearest-neighbor stacking interactions in DNA. *Biopolymers*, **32**, 849–864.
31. Breslauer,K.J., Frank,R., Blöcker,H. and Marky,L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl Acad. Sci. USA*, **83**, 3746–3750.
32. Gotoh,O. and Tgashira,Y. (1981) Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers*, **20**, 1033–1042.
33. Santalucia,J., Allawi,H.T. and Seneviratne,A. (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, **35**, 3555–3562.
34. Vologodskii,A.V., Amirikyan,B.R., Lyubchenko,Y.L. and Frank-Kamenetskii,M.D. (1984) Allowance for heterogeneous stacking in the DNA helix–coil transition theory. *J. Biomol. Struct. Dyn.*, **2**, 131–148.
35. Sugimoto,N., Nakano,S., Yoneyama,M. and Honda,K. (1996) Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res.*, **24**, 4501–4505.
36. Kool,E.T. (2001) Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 1–22.
37. Guckian,K.M. and Kool,E.T. (1998) Highly precise shape mimicry by a difluorotoluene deoxynucleoside, a replication-competent substitute for thymidine. *Angew. Chem. Int. Ed. Engl.*, **36**, 2825–2828.
38. Morales,J.C. and Kool,E.T. (2000) Importance of terminal base pair hydrogen-bonding in 3′-end proofreading by the Klenow fragment of DNA polymerase I. *Biochemistry*, **39**, 2626–2632.
39. Moran,S., Ren,R.X.F. and Kool,E.T. (1997) A thymidine triphosphate shape analog lacking Watson–Crick pairing ability is replicated with high sequence selectivity. *Proc. Natl Acad. Sci. USA*, **94**, 10506–10511.
40. Bommarito,S., Peyret,N. and SantaLucia,J. (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.*, **28**, 1929–1934.
41. Guckian,K.M., Schweitzer,B.A., Ren,R.X.F., Sheils,C.J., Tahmassebi,D.C. and Kool,E.T. (2000) Factors contributing to aromatic stacking in water: evaluation in the context of DNA. *J. Am. Chem. Soc.*, **122**, 2213–2222.