HHS Public Access

Author manuscript

ACS Nano. Author manuscript; available in PMC 2016 November 24.

Published in final edited form as:

ACS Nano. 2015 November 24; 9(11): 11325–11332. doi:10.1021/acsnano.5b05055.

Differentiation of G:C vs. A:T and G:C vs. G:mC Base:

Pairs in the Latch Zone of Alpha-Hemolysin

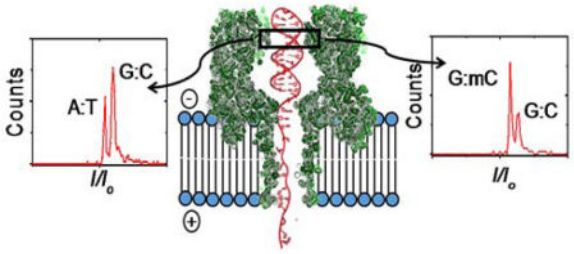
Yun Ding, **Aaron M. Fleming**, **Henry S. White***, and **Cynthia J. Burrows***

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, UT 84112-0850

Abstract

The alpha-hemolysin (α-HL) nanopore can detect DNA strands under an electrophoretic force via many regions of the channel. Our laboratories previously demonstrated that trapping duplex DNA in the vestibule of wild-type α-HL under force could distinguish the presence of an abasic site compared to a G:C base pair positioned in the latch zone at the top of the vestibule. Herein, a series of duplexes were probed in the latch zone to establish if this region can detect more subtle features of base pairs beyond the complete absence of a base. The results of these studies demonstrate the most sensitive region of the latch can readily discriminate duplexes in which one G:C base pair is replaced by an A:T. Additional experiments determined that while neither 8-oxo-7,8-dihydroguanine nor 7-deazaguanine opposite C could be differentiated from a G:C base pair, in contrast, the epigenetic marker 5-methylcytosine, when present in both strands of the duplex, yielded new blocking currents when compared to strands with unmodified cytosine. The results are discussed with respect to experimental design for utilization of the latch zone of α-HL to probe specific regions of genomic samples.

Graphical abstract



Keywords

alpha-hemolysin; latch zone; duplex DNA; base pair detection; epigenetic marker detection

*To whom correspondence should be addressed: Telephone: (801) 585-7290 or (801) 585-6256, ; Email: burrows@chem.utah.edu or ; Email: white@chem.utah.edu.

Supporting Information

Detailed methods, current vs. time traces, and current blocking histograms. This material is available free of charge via the Internet at http://pubs.acs.org.

The alpha-hemolysin (α-HL) protein nanopore has been applied for detection and quantification of solutes.[1–3] To achieve this ability, a single, well-defined protein channel is embedded into a lipid bilayer suspended across a solid support, such as a glass nanopore membrane (GNM).[4] This allows placement of electrodes on the *cis* and *trans* side of the protein, in which an applied current electrophoretically drives charged molecules toward the nanopore (Figure 1). When the molecules enter the channel, the changes in current are characteristic of the molecular interactions between the protein and the analyte; further, the duration of these events can also be diagnostic of the interactions.[5] Of particular interest are studies that analyze DNA with α-HL.[6] DNA can interact with wild-type or engineered α-HL in four general regions: (1) the β-barrel has been utilized for discrimination of nucleotides in single-stranded DNA (ssDNA) with enormous potential for DNA sequencing applications;[3,7–13] (2) the central constriction zone of wild-type α-HL can detect large adducts appended to ssDNA as it translocates through the nanopore, and this region in engineered α-HL can also discriminate nucleotide identity;[14–18] (3) the size-limiting properties of the vestibule of wild-type and engineered α-HL can be harnessed to interrogate folded DNA structures such as hairpins,[19,20] G-quadruplexes,[21,22] i-motifs,[23] and folded RNA structures;[24,25] and, (4) recently, the latch zone at the top of the vestibule of wild-type α-HL was shown to identify single abasic site lesions in double-stranded DNA (dsDNA, Figure 1) in single-molecule events.[26–28] Of these sensing zones, the most recently identified is the latch region, and its full potential for monitoring dsDNA is yet to be determined.

Work in our laboratories identified the latch zone as a sensitive region providing baseline-resolved blocking currents for discriminating a G:C base pair from a G opposite an abasic site (AP).[26–28] In practice, this was achieved by temporarily capturing dsDNA in the vestibule using an appended single-stranded tail on one of the strands to pull the duplex into the vestibule (Figure 1). The single-stranded tail (d = 1.0 nm) can penetrate the central constriction (d = 1.4 nm) and thread into the narrow β-barrel, while the wider dsDNA (d = 2.0 nm) is temporarily trapped in the vestibule.[29,30] The trapping of dsDNA allows ample time to monitor the blocking current prior to voltage-induced unzipping of the duplex, in which the strand with a tail is driven to the *trans* side of the channel.[31–34] Mapping studies identified the latch sensing zone to be a seven base pair detection window that spanned from the sixth to the thirteenth base pair of the duplex above the central constriction.[26,28] The most sensitive site was found at the tenth base pair above the central constriction. Our subsequent studies aimed to optimize the conditions for differentiating a G:C vs. G:AP base pair that found the ideal temperature, salt concentration, and electrolyte cation for detection.[27,28] Probe-based strategies for identification of sequences or modifications to sequences (e.g., 5-methylcytosine or base-pair mismatches) utilizing nanopores have been approached by other laboratories.[33,35–37] Herein, further interrogation of the latch zone to uniquely identify other sequence variations to dsDNA apart from a G:C vs. G:AP were explored. These new studies demonstrate a sequence dependency in the location of the latch sensing zone relative to our previous work,[26,28] and the ability to call a single G:C vs. A:T base pair in a precise location. A partial list of base modifications was also explored to identify others that could be detected in the latch zone. The most interesting of these was

discovery that the latch zone can distinguish a cytosine from a 5-methylcytosine when present in both strands of dsDNA.

Results and Discussion

In these studies, a set of fishhook terminal hairpin duplexes with a 30-mer poly-2′-deoxycytidine tail on the 3′ end were used as models for dsDNA. The hairpins comprised a 12-mer duplex stem with a 5′-GTTA-3′ tetraloop. In accord with the optimized temperature and electrolyte concentrations found in our previous work to achieve maximal separation while maintaining a workable event frequency,[28] the experiments were conducted at 22 ± 1 °C with 1.00 M KCl or other electrolytes.

The first hairpin duplex studied, which was also used as an internal standard for other experiments, had six base pairs of G:C on the tail side and six base pairs of A:T on the tetraloop side (Figure 2A, hp-1). The use of an internal standard allowed analysis of sequence variations between different protein channels that inherently have natural variation of 5% in the open channel current;[4] further, the use of an internal standard also provided the best possible way to determine if the deep blocking currents changed for each sequence studied (Figure 2B). The hp-1 standard was mixed in a 1:2 ratio with a strand that introduced three G:C base pairs at positions 7, 8, and 9 (Figure 2A, hp-2) and blocking currents were recorded. For all studies reported, blocking current histograms were comprised of >200 events. The blocking currents (I) were normalized by the open channel current (I_o) yielding plots of I/I_o (Figure 2C). These plots allow comparisons between samples measured with different protein pores. Initial plots revealed two populations of events with a difference in blocking current of 1.0 ± 0.2 pA, corresponding to $\Delta I/I_o = 0.010 \pm 0.002$ (Figure 2D; the current values are reported in the text instead of the normalized values for clarity). Because the two samples were always mixed with twofold more analyte strand than standard, the histogram peak areas could readily be used to identify the hp-1 signals. In the next study, the hp-1 standard was compared to hp-3 (Figure 2A) that placed G:C base pairs at positions 10, 11, and 12 relative to the standard. The histogram of the normalized blocking currents displays two populations that differed in current by 0.5 ± 0.1 pA (Figure 2D). The last sequence comparison was conducted with the hp-1 standard and a sequence that had all G:C base pairs (hp-4, Figure 2A). Histograms of the blocking currents for this study were found to be separated by 1.4 ± 0.3 pA (Figure 2D). These studies identified that a change of A:T vs. G:C blocks between positions 7 and 12 in the duplex significantly affects the blocking current level.

Interestingly, when the six base-pair A:T track from positions 7 - 12 was replaced with blocks of three G:C base pairs, the recorded current was always greater (*i.e.*, G:C was less blocking to the current). Specifically, a current increase of ~0.5 pA (positions 7 - 9) or ~1.0 pA (positions 10 - 12) was observed for the replacement of three A:T base pairs with G:C base pairs, while replacing all six A:T base pairs with G:C led to a ~1.4 pA increase (Figure 2D). Structurally, A:T tracks in DNA adopt a conformation with a narrower minor groove and wider major groove compared to classical B-form duplexes leading to a slightly wider duplex.[30] The wider duplex at A:T tracks is expected to block the current more than the duplex section observed when G:C tracks are present in the latch zone. This observation

adds further support for the latch zone of the α-HL nanopore as a detector of sequence variations in dsDNA. This led us to question if single base pair variations in this region could be detected.

Our previous studies for analyzing damaged dsDNA in the latch,[26,28] and the results with native dsDNA above, pointed us to positions 8 through 10 as likely sites that would be most sensitive for detecting alterations in the duplex sequence. A series of duplexes in which A:T base pairs were switched for G:C base pairs at positions 8 (hp-5), 9 (hp-6), or 10 (hp-7, Figure 3A) were studied in comparison to the standard hp-1. The results of these studies found positions 8 and 9 to show the greatest difference in blocking current with a value of 0.5 ± 0.1 pA between the individual A:T and G:C base pairs (Figures 3B and C). The present results demonstrate the latch zone can differentiate a single G:C vs. A:T base pair. When damaged duplexes were studied in the latch zone, a G:C vs. G:AP was most discriminated at position 10.[26,28] Further, these prior studies were conducted in a different sequence context that was comprised of G:C and A:T base pairs between the central constriction and the latch zone unlike the present studies that only had G:C base pairs in this region. We rationalize this difference in the base pair position that was detected by the latch zone in the following way. Tracks of poly (G:C) adopt secondary structures different from the classical B-form helix that can induce a slight elongation of the strand,[40] thereby elongating the duplex in the vestibule. This subtle difference in structure is reflected in the most sensitive site being at a different position than in our previous work.

Next, to demonstrate that the ability to detect native base pairs in dsDNA in the latch zone of α-HL is a general phenomenon for all dsDNAs and not those derived from hairpins, additional studies were pursued; accordingly, a control system was studied that maintained the same sequence context as the hairpin while studying it as a template-probe complementary pair without a connecting tetraloop. This study found a nearly identical current separation between the G:C vs. A:T base pairs at position 9 (0.5 ± 0.1 pA) to that observed with the fishhook hairpins (0.5 ± 0.1 pA, Figure S3); curve shapes indicate that the resolution was in fact slightly better with the two-strand duplex compared to the hairpin. This verifies that the ability to resolve the blocking currents for G:C and A:T base pairs in DNA duplexes is not limited to fishhook hairpin duplexes.

Next, studies were conducted to determine if G:C vs. C:G base pairs gave different current signals. These studies were conducted in two-stranded duplexes and yielded identical current levels (Figure S12). Thus, the latch zone of α-HL nanopore cannot differentiate different base pair orientations (i.e., G:C vs. C:G), as might be expected based on the 7-fold symmetry of interior of the latch. As further tests of the latch zone sensing capabilities, studies were then conducted with duplexes bearing single-stranded tails comprised of 5′-$(CAT)_{10}$-3′ which yielded the same current difference between A:T and G:C-containing duplexes, as previously observed (Figure S13), showing that the composition of the single-stranded tail has little or no effect on the current level discrimination. The tails were moved from the 3′ end to the 5′ end of the hairpin duplexes, and the base pair discrimination abilities were similar (5′ entry = 0.3 ± 0.1 pA difference and 3′ entry = 0.5 ± 0.1 pA difference, Figure S14). Lastly, further changes in the sequence context around position 9 led to the same 0.5 ± 0.1 pA current difference originally noted (Figure S15). These studies

conclude that the latch zone is a robust platform for detection of these base pairs in many different contexts.

Inspired by the ability to differentiate a G:C vs. A:T base pair in a duplex, a number of other base modifications were placed at position 9 to determine if the latch zone of α-HL could distinguish between them. Position 9 was chosen because it gave the best discrimination between G:C vs. A:T base pairs. Our interest in the oxidation of the guanine heterocycle[41] led us to first place 8-oxo-7,8-dihydroguanine (OG) at position 9 (hp-8, Figure 4A). When the OG:C duplex (hp-8) was compared to the standard with a G:C at the same position (hp-6), histograms of the normalized blocking current showed a single peak. This result demonstrates that the latch zone is not capable of differentiating an OG:C vs. G:C base pair in this duplex system (Figure 4A). Next, we placed 7-deazaguanine (7) in the duplex system base paired with C (hp-9, Figure 4A). When a 1:2 mixture of hp-6 and hp-9 was studied by the nanopore a single peak was again observed in the current histogram (Figure 4A). Again, this result determined the latch zone was incapable of discriminating these two very similar base pairs. These initial results, in which modifications to the G base did not lead to an observable current difference, suggest that the 1–2 atom structural differences that OG and 7-deazaguanine cause to the major groove of duplex DNA are not great enough to impact the blocking currents in the latch zone of the wild-type α-HL nanopore.

Methylation of the cytosine heterocycle at C5 to yield 5-methylcytosine (mC) yields an epigenetic marker utilized in biology to regulate gene transcription.[42–44] These methyl groups are installed at 5′-CpG-3′ sequences in the genome; therefore, we designed a duplex placing a 5′-CpG-3′ sequence at the most sensitive positions 8 and 9. In this study, the internal standard was the strand without the methylation modification (hp-10, Figure 4B), and the first test had a single mC at position 8 (hp-11, Figure 4B) representing a hemimethylated strand, a rare occurrence in the genome. Analysis of a mixture of hp-10 and hp-11 gave only one peak in the current histogram profile (Figure 4B). Next, mC was placed at the C sites at positions 8 and 9 (hp-12, Figure 4B). Interestingly, analysis of a mixture of hp-10 and hp-12 led to observation of two peaks in the current histogram that were separated by 0.2 ± 0.1 pA, and they were not baseline resolved to a point that would allow calling a sequence with >95% confidence (Figure 4B). This final result determined that the latch zone can discriminate bismethylated strands from the parent duplex. Protein nanopores have been utilized for detecting or sequencing epigenetic marks on DNA,[35,45–49] and the present results demonstrate an additional nanopore method to achieve detection of mC in specific sequences using wild-type α-HL. Studies with more than one OG or 7-deazaguanine were not pursued as for the study with mC because there is insufficient evidence that these modifications would occur in tandem like mC.

The base modifications probed in the latch zone all induced changes to the major groove of the duplex. When one modification was present, such as OG, 7, or mC, the current level remained the same as the parent duplex without the modification (Figures 4A and B). On the other hand, when two mC modifications were placed in the duplex, current differentiation with the C-containing parent duplex was observed. The observation with mC suggests an additive affect for the modifications to impact the blocking current level. When comparisons between duplex structures with C or mC were studied, a significant difference in base pair

geometry or helical parameters between them was not observed;[50] in contrast, mC induces a change in the ordered waters found in the major groove.[51] Whether the difference in observed current results from the methyl groups altering the duplex hydration, or if the difference results from a steric effect altering the current, cannot be determined from these initial results.

Herein, the latch zone of wild-type α-HL was utilized for sensing base pairs in dsDNA while they were tethered in the channel by a single-stranded tail prior to unzipping of the duplex. The results exhibit the ability to differentiate a G:C vs. A:T base pair, as well as to detect the presence of the mC epigenetic marker in a DNA duplex. When the G:C vs. A:T base pairs were placed at position 9 in the duplex, a 0.5 ± 0.1 pA separation in the histograms of blocking currents was observed (Figure 3). Protein engineering of the α-HL monomers in the β-barrel region has allowed great progress to be made for sequencing the individual DNA nucleotides in ssDNA.[7,52] Progress toward mutations in the latch zone has not been reported, and doing so is anticipated to allow greater resolution of G:C vs. A:T base pairs in dsDNA when positioned correctly in this region. Utilizing this as an actual DNA sequencing technology is hard to envision because the present method only examines one site in the duplex; however, probing methods to determine specific DNA sequences are still actively being pursued to overcome one limitation of full genome sequencing.[53,54]

Next-generation DNA sequencing is generally applied to genome samples taken from a population of cells that in the case of cancerous tissue is likely to be heterogeneous at the sites of critical single nucleotide polymorphisms (SNPs).[55] High sequence coverage (up to 1000x) is required to identify low levels of SNPs by current sequencing methods.[56] In contrast, probing methods can target specific sequences to interrogate them for low level mutations. Typical probes for SNP analysis are labeled with reporter groups (e.g., fluorophores)[57] that monitor fluorescence changes for only one of the possible alleles in the sample. In contrast, the nanopore can analyze samples label free and interrogates all alleles in solution via counting of the current levels. When looking at a population of duplexes that differ in a single base pair in the middle, the entry rate of each sequence will be dependent on the tail identity that will be identical for each sequence variant. Therefore, counting the populations of G:C vs. A:T will report on the presence and quantity of these types of SNPs in solution. The nanopore experiment presently described could be applied for the determination and quantification of these mutations via a PCR-based method to amplify specific regions of the genome followed by nanopore analysis (see outline in Figure S16) Once protein engineering has provided protein nanopores that can baseline resolve an A:T vs. G:C in dsDNA, single-molecule profiling can be achieved allowing the call of a low-level mutation via a single event (i.e., single DNA strand). Additionally, progress toward multiplexing nanopores is being carried out, as exemplified by the progress being made at Oxford Nanopores,[7] and once this capability is perfected, latch zone analysis of many mutation sites in the genome could be monitored concurrently. Lastly, the present results demonstrate that the latch zone of the α-HL nanopore can differentiate methylated vs. unmethylated CpGs in dsDNA (Figure 4B). Therefore, the ability to measure mC levels could provide a boon for future biomedical studies that address the heterogeneous changes in genomic methylation levels as a function of cancer initiation, progression, and response to treatment.[58]

Conclusion

Protein nanopores, such as α-HL, detect DNA solutes through multiple recognition zones that are specific for different aspects of DNA, such as sequence and secondary structure. The present work demonstrated that the latch zone of α-HL nanopore can differentiate the canonical DNA base pairs G:C and A:T (Figure 2). Additional studies found the latch zone to identify the presence of the epigenetic marker mC (Figure 4). These results illustrate the ability of this region to detect sequence variations and modifications to DNA. These sequence variations and modifications are characteristic changes that occur to the genome in many diseases, are part of the developmental process, and regulate many cellular functions. Our future studies will apply this methodology for detecting and quantifying modifications of critical interest.

Experimental Section

Materials and oligodeoxynucleotide (ODN) preparation

KCl, RbCl, ethanol (200 proof), wild-type α-HL monomer, the phospholipid 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC), EDTA, and Na_2HPO_4 were purchased from commercial suppliers and used without further purification. All of the nanopore experiments were performed in 1.00 M KCl, 10 mM KP_i at pH 7.4 and 22 ± 1 °C unless otherwise noted.

The ODNs were synthesized from commercially available phosphoramidites (Glen Research, Sterling, VA) by the DNA-Peptide Core Facility at the University of Utah, followed by purification using a semi-preparative ion-exchange HPLC column with a linear gradient of 20% to 100% B over 30 min while monitoring absorbance at 260 nm (A = 10% CH_3CN/90% ddH_2O; B = 20 mM Tris, 1 M NaCl, pH = 8 in 10% CH_3CN/90% ddH_2O; flow rate = 3 mL/min). The purities of the ODNs were determined by analytical ion-exchange HPLC, running the previously mentioned buffers and method, with the exception that the flow rate was 1 mL/min.

After purification, all ODNs were annealed by incubating them at 90 ºC in analysis buffer electrolyte solution for 5 min and then rapidly cooling on ice. The duplex DNA was annealed by incubating at 90 ºC in analysis buffer electrolyte solution for 5 min and then cooling slowly in the water bath to room temperature. The prepared samples were stored in a -20 ºC freezer before they were used in other experiments.

Ion current recordings

The glass nanopore membrane (GNM) (radius 800 nm) was constructed as previously established.[4,59] Diphytanoylphosphatidylcholine (DPhPC) bilayers spanning across the orifice of the GNM were prepared as previously described.[4] A proper bilayer was determined by a resistance of ~200 GΩ, a value consistent with literature sources.[4] The protein α-HL was diluted to 1 mg/mL in ultra-pure water and the DPhPC was dissolved in decane to a concentration of 10 mg/mL, both of which were stored at 80 °C. A pipette holder with a pressure gauge and a 10-mL gas-tight syringe were used to attach the GNM to the direct current (DC) system. Two Ag/AgCl electrodes were positioned inside and outside of

the GNM to apply a voltage. A plastic pipette tip was used to paint the DPhPC solution (1 μL, 10 mg/mL) on the GNM surface. After the addition of monomer α-HL (0.2 μL, 1 mg/mL), pressure was applied to form a suspended bilayer, followed by reconstitution of a single α-HL ion channel in the bilayer.[60] All experiments were performed at 22 ± 1 ºC and 100 mV (*trans* vs. *cis*). Next, the internal control was added to the *cis* side of the chamber with a final concentration of 5 μM. After collecting > 200 events, another sample was added to the same protein channel at a concentration of 10 μM to allow the comparison of the current levels between two ODNs. For each experiment, data were collected from three individual proteins channels and > 200 events were collected for each protein channel with a 10 kHz low pass filter and a 50 kHz data acquisition rate.

Data processing and analysis

Current levels and blockage durations for events in the *i-t* traces were extracted using QUB 1.5.0.31 and plotted using Origin 9.1. Only events with deep blockage current (I) that were less than 20% of the open channel current (I_o) were considered to be unzipping events, similar to our previous work with fishhook hairpin unzipping in the α-HL nanopore.[20] The error bars associated with the difference in blocking currents between the analyte and internal standard strand were determined from three individual protein channels. For the current histograms, either 100 or 150 bins of 0.1 or0.05 pA widths were used in each plot.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by a grant provided by the National Institutes of Health (R01 GM093099). The authors greatly appreciate colleagues at Electronic BioSciences (San Diego, CA) for supplying us with the instrument and software utilized for recording the current vs. time traces, and Dr. Qian Jin for conducting the initial latch zone studies that facilitated the present work. The oligonucleotides were provided by the DNA/Peptide core facility at the University of Utah that is supported in part by the NCI Cancer Support Grant (P30 CA042014).

References

1. Wanunu M. Nanopores. A Journey Towards DNA sequencing. Phys Life Rev. 2012; 9:125–158. [PubMed: 22658507]
2. Reiner JE, Balijepalli A, Robertson JW, Campbell J, Suehle J, Kasianowicz JJ. Disease Detection and Management *via* Single Nanopore-based Sensors. Chem Rev. 2012; 112:6431–6451. [PubMed: 23157510]
3. Bayley H. Nanopore Sequencing: from Imagination to Reality. Clin Chem. 2015; 61:25–31. [PubMed: 25477535]
4. White RJ, Ervin EN, Yang T, Chen X, Daniel S, Cremer PS, White HS. Single Ion-channel Recordings Using Glass Nanopore Membranes. J Am Chem Soc. 2007; 129:11766–11775. [PubMed: 17784758]
5. Kasianowicz JJ, Robertson JW, Chan ER, Reiner JE, Stanford VM. Nanoscopic Porous Sensors. Annu Rev Anal Chem. 2008; 1:737–766.
6. Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Ventra MD, Garaj S, Hibbs A, Huang X, et al. The Potential and Challenges of Nanopore Sequencing. Nat Biotechnol. 2008; 26:1146–1153. [PubMed: 18846088]

7. Jain M, Fiddes IT, Miga KH, Olsen HE, Paten B, Akeson M. Improved Data Analysis for the MinION Nanopore Sequencer. Nat Methods. 2015; 12:351–356. [PubMed: 25686389]
8. Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. Single-nucleotide Discrimination in Immobilized DNA Oligonucleotides with a Biological Nanopore. Proc Natl Acad Sci US A. 2009; 106:7702–7707.
9. Purnell R, Schmidt J. Discrimination of Single Base Substitutions in a DNA Strand Immobilized in a Biological Nanopore. ACS Nano. 2009; 3:2533–2538. [PubMed: 19694456]
10. Ayub M, Bayley H. Individual RNA Base Recognition in Immobilized Oligonucleotides Using a Protein Nanopore. Nano Lett. 2012; 12:5637–5643. [PubMed: 23043363]
11. Rincon-Restrepo M, Mikhailova E, Bayley H, Maglia G. Controlled Translocation of Individual DNA molecules Through Protein Nanopores with Engineered Molecular Brakes. Nano Lett. 2011; 11:746–750. [PubMed: 21222450]
12. Purnell R, Mehta K, Schmidt J. Nucleotide Identification and Orientation Discrimination of DNA Homopolymers Immobilized in a Protein Nanopore. Nano Lett. 2008; 8:3029–3034. [PubMed: 18698831]
13. Ayub M, Hardwick SW, Luisi BF, Bayley H. Nanopore-based Identification of Individual Nucleotides for Direct RNA Sequencing. Nano Lett. 2013; 13:6144–6150. [PubMed: 24171554]
14. Stoddart D, Heron AJ, Klingelhoefer J, Mikhailova E, Maglia G, Bayley H. Nucleobase Recognition in ssDNA at the Central Constriction of the Alpha-Hemolysin Pore. Nano Lett. 2010; 10:3633–3637. [PubMed: 20704324]
15. An N, Fleming AM, White HS, Burrows CJ. Nanopore Detection of 8-Oxoguanine in the Human Telomere Repeat Sequence. ACS Nano. 2015; 9:4296–4307. [PubMed: 25768204]
16. Mitchell N, Howorka S. Chemical Tags Facilitate the Sensing of Individual DNA Strands with Nanopores. Angew Chem Int Ed Engl. 2008; 47:5565–5568. [PubMed: 18553329]
17. Zhang X, Wang Y, Fricke BL, Gu LQ. Programming Nanopore Ion Flow for Encoded Multiplex microRNA Detection. ACS Nano. 2014; 8:3444–3450. [PubMed: 24654890]
18. Ayub M, Stoddart D, Bayley H. Nucleobase Recognition by Truncated Alpha-Hemolysin Pores. ACS Nano. 2015; 9:7895–7903. [PubMed: 26114210]
19. Vercoutere W, Winters-Hilt S, Olsen H, Deamer D, Haussler D, Akeson M. Rapid Discrimination Among Individual DNA Hairpin Molecules at Single-nucleotide Resolution Using an Ion Channel. Nat Biotechnol. 2001; 19:248–252. [PubMed: 11231558]
20. Ding Y, Fleming AM, White HS, Burrows CJ. Internal vs Fishhook Hairpin DNA: Unzipping Locations and Mechanisms in the Alpha-Hemolysin Nanopore. J Phys Chem B. 2014; 118:12873–12882. [PubMed: 25333648]
21. An N, Fleming AM, Middleton EG, Burrows CJ. Single-molecule Investigation of G-Quadruplex Folds of the Human Telomere Sequence in a Protein Nanocavity. Proc Natl Acad Sci U S A. 2014; 111:14325–14331. [PubMed: 25225404]
22. Shim JW, Tan Q, Gu LQ. Single-molecule Detection of Folding and Unfolding of the G-Quadruplex Aptamer in a Nanopore Nanocavity. Nucleic Acids Res. 2009; 37:972–982. [PubMed: 19112078]
23. Ding Y, Fleming AM, He L, Burrows CJ. Unfolding Kinetics of the Human Telomere i-Motif Under a 10 pN Force Imposed by the Alpha-Hemolysin Nanopore Identify Transient Folded-state Lifetimes at Physiological pH. J Am Chem Soc. 2015; 137:9053–9060. [PubMed: 26110559]
24. Japrung D, Henricus M, Li Q, Maglia G, Bayley H. Urea Facilitates the Translocation of Single-stranded DNA and RNA Through the Alpha-Hemolysin Nanopore. Biophys J. 2009; 98:1856–1863. [PubMed: 20441749]
25. Smith AM, Abu-Shumays R, Akeson M, Bernick DL. Capture, Unfolding, and Detection of Individual tRNA Molecules Using a Nanopore Device. Front Bioeng Biotechnol. 2015; 3:91–102. [PubMed: 26157798]
26. Jin Q, Fleming AM, Johnson RP, Ding Y, Burrows CJ, White HS. Base-excision Repair Activity of Uracil-DNA Glycosylase Monitored Using the Latch Zone of Alpha-Hemolysin. J Am Chem Soc. 2013; 135:19347–19353. [PubMed: 24295110]

27. Johnson RP, Fleming AM, Burrows CJ, White HS. Effect of an Electrolyte Cation on Detecting DNA Damage with the Latch Constriction of Alpha-Hemolysin. J Phys Chem Lett. 2014; 5:3781–3786. [PubMed: 25400876]
28. Johnson RP, Fleming AM, Jin Q, Burrows CJ, White HS. Temperature and Electrolyte Optimization of the α,-Hemolysin Latch Sensing Zone for Detection of Base Modification in Double-stranded DNA. Biophys J. 2014; 107:924–931. [PubMed: 25140427]
29. Song L, Hobaugh M, Shustak C, Cheley S, Bayley H, Gouaux J. Structure of *Straphylococcal* α,-Hemolysin, a Heptameric Transmembrane pore. Science. 1996; 274:1859–1866. [PubMed: 8943190]
30. Drew HR, Wing RM, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE. Structure of a B-DNA Dodecamer: Conformation and Dynamics. Proc Natl Acad Sci US A. 1981; 78:2179–2183.
31. Schibel A, Fleming A, Jin Q, An N, Liu J, Blakemore C, White H, Burrows C. Sequence-specific Single-molecule Analysis of 8-oxo-7,8-Dihydroguanine Lesions in DNA Based on Unzipping Kinetics of Complementary Probes in Ion Channel Recordings. J Am Chem Soc. 2011; 133:14778–14784. [PubMed: 21875081]
32. Wang Y, Tian K, Hunter LL, Ritzo B, Gu LQ. Probing Molecular Pathways for DNA Orientational Trapping, Unzipping and Translocation in Nanopores by Using a Tunable Overhang Sensor. Nanoscale. 2014; 6:11372–11379. [PubMed: 25144935]
33. Sauer-Budge AF, Nyamwanda JA, Lubensky DK, Branton D. Unzipping Kinetics of Double-stranded DNA in a Nanopore. Phys Rev Lett. 2003; 90:23801.
34. Jin Q, Fleming AM, Burrows CJ, White HS. Unzipping Kinetics of Duplex DNA Containing Oxidized Lesions in an α,-Hemolysin Nanopore. J Am Chem Soc. 2012; 134:11006–11011. [PubMed: 22690806]
35. Wang Y, Luan BQ, Yang Z, Zhang X, Ritzo B, Gates K, Gu LQ. Single Molecule Investigation of Ag^{+} Interactions with Single Cytosine-, Methylcytosine- and Hydroxymethylcytosine-cytosine Mismatches in a Nanopore. Sci Rep. 2014; 4:5883. [PubMed: 25103463]
36. Balagurusamy VS, Weinger P, Ling XS. Detection of DNA Hybridizations Using Solid-state Nanopores. Nanotechnology. 2010; 21:335102. [PubMed: 20657045]
37. Nakane J, Wiggin M, Marziali A. A Nanosensor for Transmembrane Capture and Identification of Single Nucleic Acid Molecules. Biophys J. 2004; 87:615–621. [PubMed: 15240494]
38. Ulyanov NB, Bauer WR, James TL. High-resolution NMR Structure of an AT-rich DNA Sequence. J Biomol NMR. 2002; 22:265–280. [PubMed: 11991355]
39. Acosta-Reyes FJ, Subirana JA, Pous J, Sanchez-Giraldo R, Condom N, Baldini R, Malinina L, Campos JL. Polymorphic Crystal Structures of an All-AT DNA Dodecamer. Biopolymers. 2015; 103:123–133. [PubMed: 25257185]
40. Timsit Y, Westhof E, Fuchs RP, Moras D. Unusual Helical Packing in Crystals of DNA Bearing a Mutation Hot Spot. Nature. 1989; 341:459–462. [PubMed: 2797169]
41. Fleming AM, Burrows CJ. G-Quadruplex Folds of the Human Telomere Sequence Alter the Site Reactivity and Reaction Pathway of Guanine Oxidation Compared to Duplex DNA. Chem Res Toxicol. 2013; 26:593–607. [PubMed: 23438298]
42. Zheng G, Fu Y, He C. Nucleic Acid Oxidation in DNA Damage Repair and Epigenetics. Chem Rev. 2014; 114:4602–4620. [PubMed: 24580634]
43. Booth MJ, Raiber EA, Balasubramanian S. Chemical Methods for Decoding Cytosine Modifications in DNA. Chem Rev. 2015; 115:2240–2254. [PubMed: 25094039]
44. Wagner M, Steinbacher J, Kraus TF, Michalakis S, Hackner B, Pfaffeneder T, Perera A, Muller M, Giese A, Kretzschmar HA, et al. Age-dependent Levels of 5-Methyl-, 5-Hydroxymethyl-, and 5-Formylcytosine in Human and Mouse Brain Tissues. Angew Chem Int Ed Engl. 2015; 54:12511–12514. [PubMed: 26137924]
45. Wescoe ZL, Schreiber J, Akeson M. Nanopores Discriminate among Five C5-cytosine Variants in DNA. J Am Chem Soc. 2014; 136:16582–16587. [PubMed: 25347819]
46. Zeng T, Liu L, Li T, Li Y, Gao J, Zhao Y, Wu HC. Detection of 5-Methylcytosine and 5-Hydroxymethylcytosine in DNA *via* Host-guest Interactions inside Alpha-Hemolysin Nanopores. Chem Sci. 2015; 6:5628–5634.

47. Clarke J, Wu H, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous Base Identification for Single-molecule Nanopore DNA Sequencing. Nat Nanotechnol. 2009; 4:265–270. [PubMed: 19350039]

48. Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM, Bartlett JJ, Pavlenok M, Gundlach JH. Detection and Mapping of 5-Methylcytosine and 5-Hydroxymethylcytosine with Nanopore MspA. Proc Natl Acad Sci U S A. 2013; 110:18904–18909. [PubMed: 24167255]

49. Wallace EV, Stoddart D, Heron AJ, Mikhailova E, Maglia G, Donohoe TJ, Bayley H. Identification of Epigenetic DNA Modifications with a Protein Nanopore. Chem Commun. 2010; 46:8195–8197.

50. Renciuk D, Blacque O, Vorlickova M, Spingler B. Crystal Structures of B-DNA Dodecamer Containing the Epigenetic Modifications 5-Hydroxymethylcytosine or 5-Methylcytosine. Nucleic Acids Res. 2013; 41:9891–9900. [PubMed: 23963698]

51. Mayer-Jung C, Moras D, Timsit Y. Hydration and Recognition of Methylated CpG Steps in DNA. EMBO J. 1998; 17:2709–2718. [PubMed: 9564052]

52. Stoddart D, Maglia G, Mikhailova E, Heron A, Bayley H. Multiple Base-recognition Sites in a Biological Nanopore: Two Heads are Better than One. Angew Chem Int Ed Engl. 2010; 49:556–559. [PubMed: 20014084]

53. Wang T, Guan W, Lin J, Boutaoui N, Canino G, Luo J, Celedon JC, Chen W. A Systematic Study of Normalization Methods for Infinium 450K Methylation Data Using Whole-genome Bisulfite Sequencing Data. Epigenetics. 2015; 10:662–669. [PubMed: 26036609]

54. Xuan F, Fan TW, Hsing IM. Electrochemical Interrogation of Kinetically-controlled Dendritic DNA/PNA Assembly for Immobilization-free and Enzyme-free Nucleic Acids Sensing. ACS Nano. 2015; 9:5027–5033. [PubMed: 25872652]

55. Burrell RA, McGranahan N, Bartek J, Swanton C. The Causes and Consequences of Genetic Heterogeneity in Cancer Evolution. Nature. 2013; 501:338–345. [PubMed: 24048066]

56. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing Depth and Coverage: Key Considerations in Genomic Analyses. Nat Rev Genet. 2014; 15:121–132. [PubMed: 24434847]

57. Mir KU, Southern EM. Sequence Variation in Genes and Genomic DNA: Methods for Large-scale Analysis. Annu Rev Genomics Hum Genet. 2000; 1:329–360. [PubMed: 11701633]

58. Jones PA, Baylin SB. The Epigenomics of Cancer. Cell. 2007; 128:683–692. [PubMed: 17320506]

59. Zhang B, Galusha J, Shiozawa P, Wang G, Bergren A, Jones R, White R, Ervin E, Cauley C, White H. Bench-top Method for Fabricating Glass-sealed Nanodisk Electrodes, Glass Nanpore Electrodes, and Glass Nanopore Membranes of Controlled Size. Anal Chem. 2007; 79:4778–4787. [PubMed: 17550232]

60. Schibel AE, Heider EC, Harris JM, White HS. Fluorescence Microscopy of the Pressure-dependent Structure of Lipid Bilayers Suspended across Conical Nanopores. J Am Chem Soc. 2011; 133:7810–7815. [PubMed: 21542629]

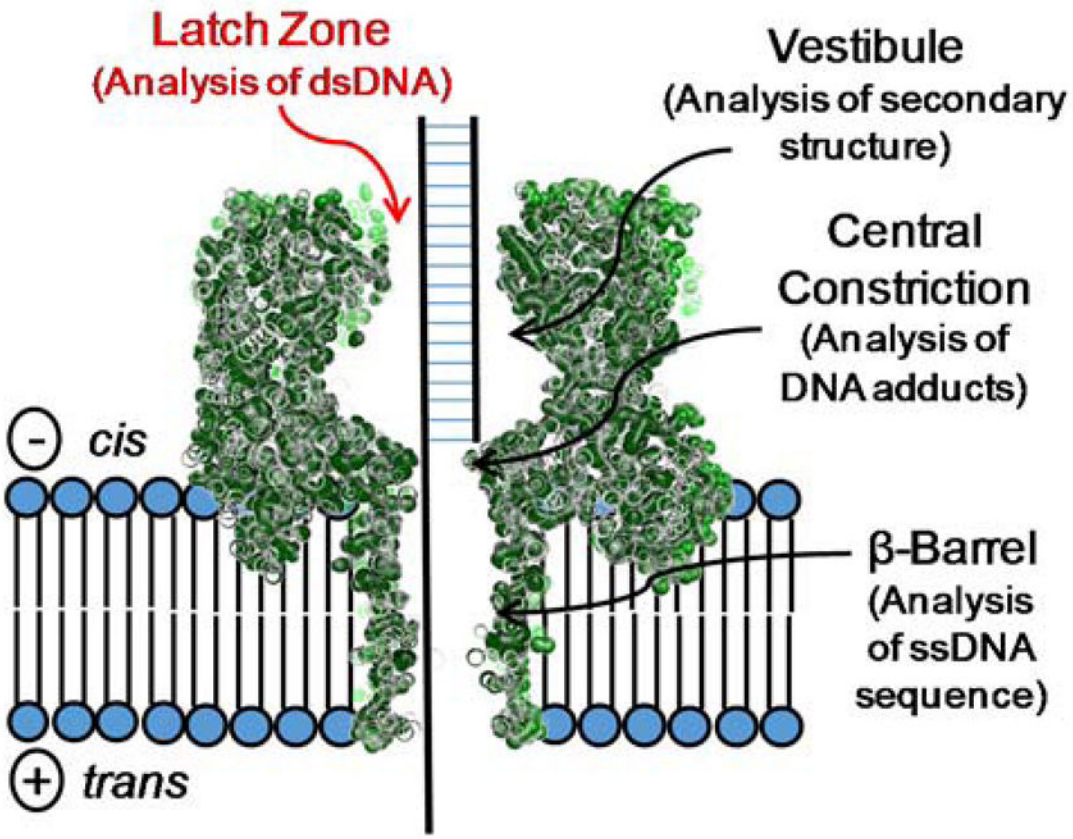
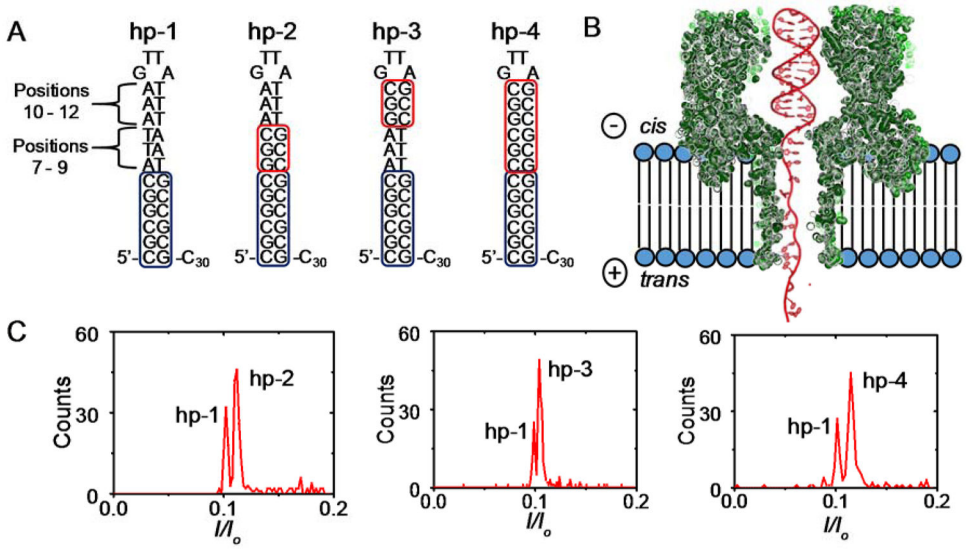


Figure 1.
The DNA-specific sensing regions of the α-HL protein nanopore. The protein space filling model is derived from pdb 7AHL.[29]

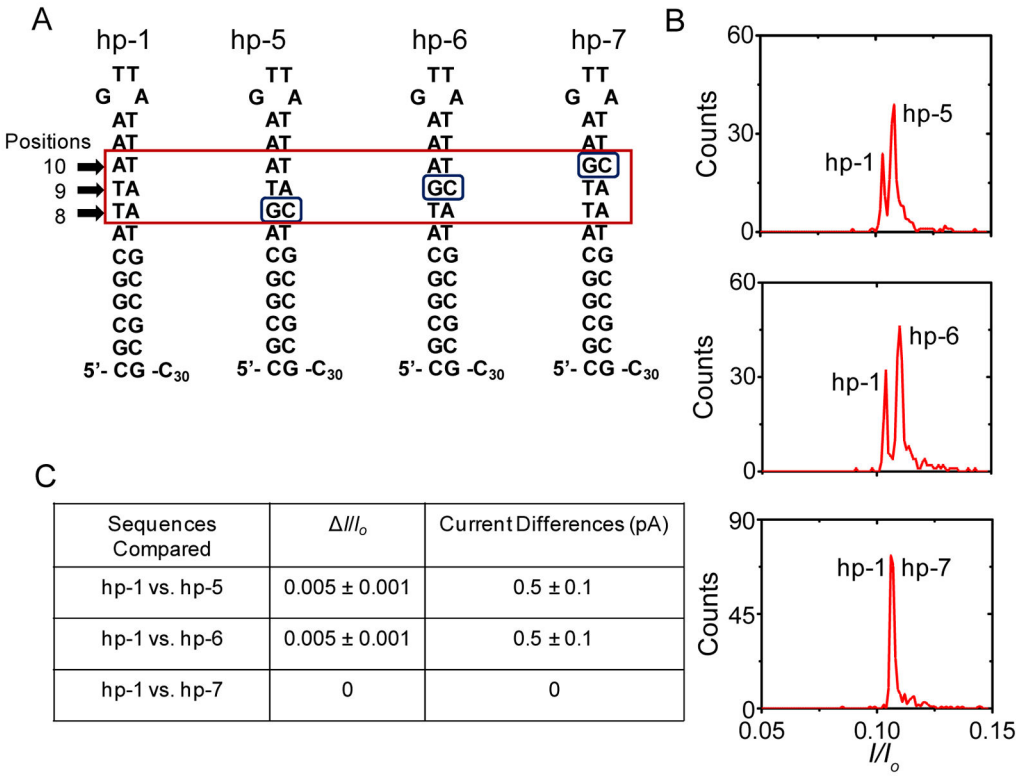


D

Sequences Compared	$\Delta I/I_o$	Current Difference (pA)
hp-1 vs. hp-2	0.010 ± 0.002	1.0 ± 0.2
hp-1 vs. hp-3	0.005 ± 0.001	0.5 ± 0.1
hp-1 vs. hp-4	0.014 ± 0.003	1.4 ± 0.3

Figure 2.

Mapping the latch sensing zone of the α-HL nanopore with blocks of G:C base pairs in a background of A:T base pairs. (A) The hairpin duplex sequences studied. (B) Representative structure for α-HL and the hairpin duplex based on pdb 7AHL,[29] 1JVE,[38] 4HW1.[39] (C) Current blocking histograms for mixtures of analyte and standard duplexes. The hp-1 sequence was used as an internal standard, and it was always mixed with the analyte strand in a ratio of 1:2, respectively; therefore, the smaller peak area always represents hp-1. (D) Table of peak-to-peak $\Delta I/I_o$ and current differences measured between the internal standard (hp-1) and the analyte strand. The error in each value represents the standard deviation of peak-to-peak widths from three individual protein channels. The data were recorded with 10 mM KP_i (pH 7.4), 1.00 M KCl, at 22 ± 1 °C, and a 100 mV (*trans* vs. *cis*) bias. The histograms represent > 200 recorded events.



C

Sequences Compared	$\Delta I/I_o$	Current Differences (pA)
hp-1 vs. hp-5	0.005 ± 0.001	0.5 ± 0.1
hp-1 vs. hp-6	0.005 ± 0.001	0.5 ± 0.1
hp-1 vs. hp-7	0	0

Figure 3.

Current resolution studies of a G:C vs. A:T base pair at positions 8, 9, and 10 in the latch zone of the α-HL protein nanopore. (A) Sequences for the hairpin duplexes studied. (B) Current blocking histograms for mixtures of the hairpin duplexes. The hp-1 sequence was used as an internal standard and it was always mixed with the analyte strand in a ratio of 1:2, respectively; therefore, the smaller peak area always represents hp-1. (C) Table of peak-to-peak $\Delta I/I_O$ and current differences measured between the standard (hp-1) and the analyte. The error in each value represents the standard deviation of peak-to-peak widths from three individual protein channels. The data were recorded with 10 mM KP_i (pH 7.4), 1.00 M KCl, at 22 ± 1 °C, and a 100 mV (*trans* vs. *cis*) bias. The histograms represent > 200 recorded events.

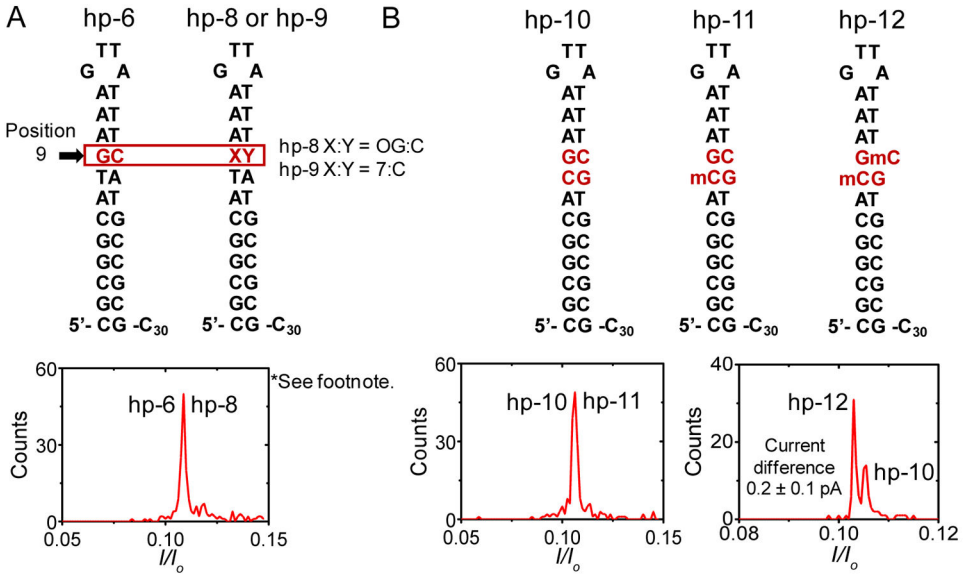


Figure 4.

Additional base pair alterations monitored in the latch zone of α-HL. (A) Comparison of G:C vs. OG:C or G:C vs. 7-deazaguanine (7):C when placed at position 9. The blocking current histograms for both studies identified a single population. These results suggest the latch zone of wild-type α-HL cannot resolve these base pairs. (B) Comparison of the epigenetic markers mC with C-containing duplexes. These blocking current histograms identify that one mC cannot be distinguished from one C, though; two mCs can be differentiated from the parent strand. The data were recorded with 10 mM KP_i (pH 7.4), 1.00 M KCl, at 22 ± 1 °C, and a 100 mV (*trans* vs. *cis*) bias. The histograms represent > 200 recorded events. *Blocking current histograms for hp-6 vs. hp-8 and hp-6 vs. hp-9 were nearly identical; therefore, in the figure data for hp-6 vs. hp-8 is provided and the data for hp-6 vs. hp-9 can be found in Figure S2.