Published in final edited form as:
*J Am Chem Soc*. 2011 September 21; 133(37): 14778–14784. doi:10.1021/ja205653v.

# 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

**Anna E. P. Schibel**, **Aaron M. Fleming**, **Qian Jin**, **Na An**, **Jin Liu**, **Charles P. Blakemore**, **Henry S. White***, and **Cynthia J. Burrows***
Department of Chemistry, University of Utah, 315 S. 1400 East, RM 2020, Salt Lake City, UT 84112-0850, USA

## Abstract

Translocation measurements of intact DNA strands with the ion channel α-hemolysin (α-HL) are limited to single-stranded DNA (ssDNA) experiments as the dimensions of the channel prevent double-stranded DNA (dsDNA) translocation; however, if a short oligodeoxynucleotide is used to interrogate a longer ssDNA strand, it is possible to unzip the duplex region when it is captured in the α-HL vestibule, allowing the longer strand to translocate through the α-HL channel. This unzipping process has a characteristic duration based on the stability of the duplex. Here, ion channel recordings are used to detect the presence and relative location of the oxidized damage site 8-oxo-7,8-dihydroguanine (OG) in a sequence-specific manner. OG engages in base pairing to C or A with unique stabilities relative to native base Watson-Crick pairings, and this phenomenon is used here to engineer probe sequences (10–15 mers) that when base-paired with a 65 mer sequence of interest, containing either G or OG at a single site, produce characteristic unzipping times that correspond well with the duplex melting temperature ($T_m$). Unzipping times also depend on the direction from which the duplex enters the vestibule if the stabilities of leading base pairs at the ends of the duplex are significantly different. It is shown here that the presence of a single DNA lesion can be distinguished from an undamaged sequence and that the relative location of the damage site can be determined based on the duration of duplex unzipping.

## Introduction

Detection of chemical damage to genomic and mitochondrial DNA remains an important and challenging task.[1] In an ideal case, one would determine the precise chemistries and locations of all modifications occurring on individual DNA strands in healthy vs. diseased cells and be able to monitor such changes as a function of oxidative, alkylative, or micronutrient stress.[2,3] In the case of the key base oxidation product 8-oxo-7,8-dihydroguanine (OG), the most sensitive and accurate methods of quantification involve either the comet assay to produce strand breaks at each lesion or digestion followed by LC-MS analysis.[4] Shortcomings of these procedures include the loss of sequence information surrounding the lesion and the difficulty in distinguishing between multiple lesions on one strand relative to the average of a set of strands. We therefore sought a single-molecule method that would overcome these limitations.[5]

white@chem.utah.edu; burrows@chem.utah.edu.

Supporting Information Available. Sample $T_m$ traces, example *i-t* traces for duplex unzipping, and event duration histograms and fittings. This information is available free of charge via the Internet at http://pubs.acs.org.

As a step towards these goals, we report the ability of short oligodeoxynucleotides to interrogate a target strand of DNA in a sequence specific manner for the presence of OG vs. G. Our strategy is based on measuring the characteristic time for unzipping the duplex formed by the target and a short oligodeoxynucleotide probe in an α-hemolysin (α-HL) ion channel. Other laboratories have shown that the time required for electrophoretically-driven translocation of single-stranded DNA (ssDNA) oligomers through the α-HL nanopore is greatly increased when a terminal hairpin is present which must unfold before threading through the narrow constriction zone of α-HL.[6–11] Alternatively, addition of a small complementary DNA probe sequence also increases τ by providing a quasi-stable duplex which must unzip before translocation of the strand can be completed.[12–16] We chose the latter approach in designing a detection method for OG vs. G in a specific target sequence (Figure 1). We demonstrate that by monitoring the unzipping time constant τ, a specific target sequence containing G at a position of interest can be differentiated from a duplex containing OG at the same position. This capability results from the difference in duplex stability imparted by the presence of either G or OG and is reflected in τ. In concept, this method could be used to detect multiple DNA lesions within a given DNA target strand as each lesion will uniquely influence the duplex stability and thus τ. In the studies presented here, single OG lesions were examined to demonstrate the ability to detect single nucleotide damage and to determine the relative location of the lesion based on 3′ or 5′ unzipping.

A unique feature of OG is its ability to base pair to either C or A with nearly equal stability, compared to the parent base G, which forms a highly stable base pair only with C. NMR and x-ray crystallographic studies provide the base pair structures shown in Figure 2.[17–22] The Watson-Crick OG:C base pair, in which both bases are in the normal *anti* conformation about the glycosidic bond, is slightly destabilized compared to the G:C base pair because of steric interactions between the C8 oxo group and the C4′ oxygen of the same nucleotide; typical duplex melting temperatures are ~2 °C lower for the OG:C pair.[23] The adverse steric interaction may be relieved by rotating OG to a *syn* conformation which presents the Hoogsteen face of OG for pairing with A via two complementary H bonds rather than three. The OG:A base pair usually has a $T_m$ lowered by another 1–2 °C compared to OG:C. In contrast, the presence of a G:A mismatch is significantly more destabilizing to a short duplex ($\Delta T_m$ = 6–8 °C) because two purines in *anti* conformations require a wider backbone spacing (Figure 2, red arrow).[24] Alternative structures for the G:A mismatch may be stabilized in tandem mismatches that provide greater π stacking or in other conformations that require low pH;[22,25–27] however, these conditions are not present in the single X:Y base (mis)pairs of the current study.

## Results and Discussion

In the present work, we took advantage of the difference in stability of G vs. OG opposite C or A to design short complementary probes that would slow the translocation of ssDNA through the α-HL ion channel. Experiments were conducted initially with a set of 10 mer (10C or 10A) and 15 mer (15C and 15A) probes complementary to the central region of a 65 mer (denoted as G or OG) resulting in the following eight duplexes: G:10C or A, OG:10C or A, G:15C or A and OG:15C or A, as shown in Table 1. (Sequences were selected to minimize secondary structure.) $T_m$ studies conducted in the same buffer solution (10 mM PBS, pH 7.4, 1 M KCl, 1 mM EDTA) as used for electrical measurements of the nanopore showed the general trends expected for stability of a single X:Y mispair compared to a native G:C pair. The order of stability was found to be G:C > OG:C > OG:A > G:A for both 10 mer and 15 mer probes annealed to the 65 mer target strand. The 10 mer probes were more sensitive to the presence of a single-base mismatch; the $T_m$ for a G:A mismatch was 8 °C lower than the parent G:C in the 10 mer but only 5.5 °C lower in the 15 mer duplex.

Electrical measurements were employed to determine the duplex stability as a function of the unzipping time constant τ. To perform these measurements, a voltage was applied across an α-HL ion channel reconstituted into a lipid bilayer suspended across the orifice of a glass nanopore membrane (GNM).[5,28–33] The annealed duplex DNA was then electrophoretically driven into the α-HL channel where it unzipped for translocation (Figure 1). The analysis of duplex unzipping kinetics was performed by plotting histograms of $\log_{10}$(events) vs. event duration time. From the slope of these histograms, the translocation time constant τ was determined for each duplex by assuming that unzipping follows a first-order kinetics law. Figure 3 shows an example of fitting for G:10C (or A) and OG:10C (or A) for +80 mV applied voltage (*trans* vs. *cis*), and the resulting values of τ. Histograms used for the determination of τ for the other DNA duplexes, and at different applied voltages, can be found in the Supporting Information. Probes of 25 nucleotides in length (5′-AAAAAAACATCTGA**(C** or **A)**GGCTCAAAAA) were also examined, but due to the very stable nature of the resulting duplex, an adequate sample of events could not be collected for voltages <140 mV; thus, 25 mer probes were not studied further. Measurements of the translocation of 65 mers G and OG in the presence of C or A-containing complementary probes indicated a correlation between the values of τ and $T_m$. Figure 4 shows the time constant for unzipping as a function of the melting temperature for 10 mer and 15 mer duplex sequences. As seen from this data, the unzipping time decreases in the order G:C > OG:C > OG:A > G:A, and correlates well with duplex melting temperatures, for both 10 and 15 mer probes. The duplexes generated with the 10 mer probes exhibited unzipping times that were distinct from one another at +80 mV, but at higher voltages (+100 mV and +120 mV) unzipping proceeded at similar rates, decreasing the measurement sensitivity, probably due to the inherent instability of 10 mer duplexes. For the 15 mer duplexes, use of a lower voltage (+80 mV) resulted in event durations that were very long (> 10 s for G:15C), making it difficult to collect an adequate population of events to distinguish between the different duplexes, and at higher voltages again the unzipping times show less sensitivity to the duplex stability. From this initial study it was determined that the probe design must balance the sensitivity, stability and the longer τ values of the 15 mer that limit event frequency. This process of probe refinement is described below.

Although the 10 mer probes displayed $T_m$ values that were more sensitive to the presence of a single X:Y mispair, the τ values for the 15 mer probes were generally in a better event duration range for discriminating between G and OG. Duplexes with $T_m$ values >50 °C provided sufficiently long translocation times that could be readily discerned from translocations of the 65 mer (~280 μs, see Supporting Information). In order to refine the probe design, we sought methods to enlarge the difference in event duration between OG:A and G:A mispairs. Previous studies have shown that locating the mismatch further away from the center of the probe led to a greater distinction in $T_m$ values between correctly paired and mismatched duplexes.[34] Thus, three new sets of 12 mer probes were tested in which the X:Y base pair of interest was located in the middle or near either the 3′ or 5′ end of the probe, and care was taken to maintain the same number of G:C and A:T base pairs in each of the two sets so that a meaningful comparison could be made. As a further refinement, the G:C base pair immediately 5′ to the G/OG site was converted to a C:I base pair for all 12 mer probes (Table 2). Inosine (I) is the nucleoside form of the hypoxanthine base which lacks the 2-amino group of G and forms a weaker base pair with C. We predicted this substitution might destabilize the adjacent G:A mispair to a greater extent than the OG:A pair; G:A mispairs are known to be sensitive to the surrounding sequence, particularly on the 5′ side of the G site.[26] The $T_m$ data for the resulting 12 mer probes annealed to 65 mer target strands are shown in Tables 2A–C, in which the probe is shifted in either the 5′ or 3′ direction or centered on the X:Y pair of interest.

The data in Tables 2A–C confirm that the displacement of the probe sequences with respect to the potential mismatched base pair can influence the thermal stability of the duplexes. Importantly, shifting the probe sequence toward the 3′ direction relative to the target strand (Y in position 10) led to a greater difference between the relatively stable OG:A pairing ($\Delta T_m$ = −4.6 °C with respect to G:C, Table 2C) and the unstable pairing of G:A ($\Delta T_m$ = −8.3 °C) compared to other probe arrangements. When the probe was shifted in the opposite direction (position 3), the $\Delta T_m$ values were smaller (Table 2A).

Electrical measurement of the 12 mers annealed to G or OG-containing 65 mers unzipping and translocating through the α-HL ion channel were conducted at 80 mV to maximize the apparent differences in unzipping times due to the duplex stability. The unzipping time, τ, was determined in the same manner as described above for the 10 mer and 15 mer duplexes; the unzipping events were plotted as histograms of $\log_{10}$(events) vs. event duration (individual duplex fittings and example *i*-*t* traces are available in the Supporting Information). A plot of τ as a function of the shift in probe position is presented in Figure 5 for each of the 12 mer duplexes shown in Table 2.

We again found that the unzipping times corresponded well to the melting temperatures; for each position (3, 6, and 10), the τ value reflected the expected duplex stability as shown in the melting temperatures (Table 2): G:C > OG:C > OG:A > G:A. Additionally, the largest difference in τ between the duplexes occurs for position 10 as predicted by the $\Delta T_m$ discussed above, demonstrating that the unzipping time τ is a sufficient measure of duplex DNA stability to allow the presence of a single DNA damage site to be detected.

However, these positional studies highlighted another feature of DNA translocation studies in the ion channel. DNA can enter the α-HL vestibule to unzip the duplex from either the 5′ or 3′ orientation. For the generation of the asymmetrical 12 mer probes (position 3 and position 10), the duplex unzipping times are expected to depend on whether the strand enters from the terminus closer to or farther away from the mismatched base pair, as suggested in recent studies of the kinetics of hairpin unzipping.[35] The melting temperature, a thermodynamic phenomenon, will not necessarily reflect the kinetics of 5′ or 3′ initiated duplex dissociation.

To examine 3′ vs. 5′ entrance into the α-HL channel, multi-dimensional plots were generated to examine the current blockage level as a function of event duration and event population density. Figure 6 illustrates how the unzipping events are dispersed based on the current blockage level and translocation time. Generally, the duplexes formed with an asymmetrical probe sequence (positions 3 and 10) displayed two well-resolved populations of events separated by both the current blockage level and the event duration, whereas the symmetrical sequences (position 6 probes), did not show any separate event populations based on the event duration. Symmetrical sequences did show some separation of populations based on the current blockage level, however. We interpret the multiple populations as being due to 3′ vs. 5′ entrance into the α-HL channel. It has been shown that for immobilized homopolymer sequences that 3′ and 5′ entrances into the α-HL channel have different current blockage levels, and these may change significantly based on the nucleotide identity within the channel.[36] Additionally, as shown above and by previous reports, duplex sequences of different stabilities have different unzipping event durations, with the less stable duplex exhibiting the faster time constant.[10,11,15,37]

Because the 65 mer sequence is a heterosequence embedded within a poly-dT background sequence at each end, and based on the orientation of the molecule upon entry into the α-HL channel and the probe sequence, a variable amount of the heterosequence will be within the α-HL channel. Thus, assignment of entry via the 3′ or 5′ end was not made based on the

current blockage level, but rather on the event duration. We assign 3′ vs. 5′ entrance of the 65 mer into the α-HL based on the expected stability of the duplex by looking at the first few terminal base pairs. (Note: all 3′/5′ designations are made referring to the orientation of the 65 mer sequence because it enters first.) For example, the position 3 G:C duplex is expected to be more stable for 3′ entrance relative to 5′ entrance as there are two G:C base pairs and one A:T base pair on the 3′ end of the duplex, but three A:T base pairs on the 5′ end (Table 2). Thus the 3′ entrance is expected to produce a longer unzipping time due to the higher G:C base pair content. Examining the position 3 G:C duplex density plot in Figure 6, the deeper blocking events have longer event durations than those of more shallow blocking, indicating that the deeper block corresponds to 3′ entrance whereas the more shallow blocking level with shorter event durations indicates 5′ entrance. The opposite is expected to be true for the position 10 G:C duplex, where the higher terminal G:C base pair content is located at the 5′ end of the duplex, which is expected to produce longer event durations relative to the 3′. From the density plot in Figure 6, the deeper blockage assigned as 3′ entrance possesses shorter event durations relative to the shallower 5′ blockage, consistent with the expected duplex stability of the position 10 G:C duplex. This convention of the deeper blockage corresponding to 3′ entrance and shallower blockages being 5′ entrance was applied to all the duplexes that could easily be separated into two distinct populations. The data for each duplex were separated into two populations, and each population was fit with a single exponential decay to determine the unzipping time constants, $\tau_{3'}$ and $\tau_{5'}$, for 3′ entrance or 5′ entrance, respectively. Histograms and data analysis are presented in the Supporting Information. The position 6 OG:C and OG:A duplexes could not be resolved into two populations, and the time constant was assigned to be τ.

Figure 7 compares the time constants for 3′ vs. 5′ entrance as assigned by the expected event duration from the orientation specific duplex stability. This $\tau_{3'}$ and $\tau_{5'}$ assignment agrees remarkably well with the predicted unzipping times for 3′ and 5′ entry. For the symmetrical probe sequence where the data can be resolved into two populations (G:C and G:A, position 6), there is little difference between $\tau_{3'}$ and $\tau_{5'}$. This is expected as both duplexes possess the same amount of terminal G:C base pair content. It is not understood why two populations are not resolvable for OG:C and OG:A in position 6. Further, for position 3 duplexes, as discussed above, the 5′ entrance is expected to reflect a more stable duplex relative to 3′ entrance. This is observed for position 3 duplexes G:C, OG:C, and OG:A, with $\tau_{3'}>\tau_{5'}$, but also for G:A, which is not well understood as the G:A mismatch should destabilize the duplex for 3′ entrance. Finally, for position 10 duplexes, the 5′ entrance is expected to be more stable than the 3′ entrance ($\tau_{3'}<\tau_{5'}$), except for the G:A duplex where the mismatch destabilizes the 5′ end of the duplex relative to the 3′. This prediction is reflected in $\tau_{3'}>\tau_{5'}$.

It has been previously demonstrated that the entry of ssDNA into the α-HL pore produced two event populations as a result of either 3′ entry or 5′ entry of the DNA strand,[36] separated in current blockage level and event duration.[38] Additionally, it has been reported that 5′ entry is less favorable relative to 3′ entry due to the tilt of the DNA bases toward the 5′ terminus, making 5′ entry, and subsequent translocation, occur at a lower rate relative to 3′ entry.[12,38,39] Based on this, it was not anticipated that distinct populations would appear as shown in Figure 6, separated by both the event duration and current blockage level, with an applied bias of 80 mV, as an 80 mV voltage bias is near the minimum electrophoretic force required to initiate DNA interaction with the α-HL channel.[40] Other duplex unzipping experiments have noted multiple event populations distinguishable by the event duration for a single duplex DNA sample,[11,14,15,37,41] requiring a more complicated analysis than presented above. Unlike the previous duplex unzipping experiments in which the current blockage level was not emphasized in the analysis, and only the temporal separation was directly applied, the analysis presented here uses the current blockage to distinguish populations prior to analyzing the temporal separation. This allowed a simplified mechanism

to be applied via single-exponential fits that generally correspond well to the relative 3′ vs. 5′ duplex stability. It is unknown if the duplex region influences the DNA entry into the α-HL channel in a manner different from ssDNA, allowing equal 3′ and 5′ entry rates.

Overall, the data presented in Figure 7 reveal that the relative position of a DNA damage site can be determined by how it influences the stability in an orientation-specific manner. Deviations present in the data from the expected duplex stability are attributed to possible interactions other than hydrogen bond strength between the base pairs and are not fully understood. However, further refinement of the probe design may lead to improved sensitivity to DNA damage, for example by chemical modification within the probe to further stabilize or destabilize the resultant duplex. This approach should be applicable to most sequence contexts surrounding OG; a summary of $T_m$ data for four additional sequences containing these base pairs shows the same order of stabilities, often with higher $\Delta T_m$ values (see Supporting Information).

## Conclusions

These studies demonstrate that a single oxidized damage site can be detected in a sequence specific location by annealing a probe to the sequence surrounding the damage site and observing the rate of duplex unzipping required for translocation through the α-HL ion channel. The presence of a single OG lesion will influence the duplex stability in a specific manner depending on whether OG base pairs with C or A, both of which produce event duration times distinct from the native G:C and G:A base pairs. Additionally, if the probe is designed so that the DNA damage site occurs on the 3′ or 5′ end of the probe, as opposed to the middle, the duplex stability will vary based on 5′ or 3′ entry into the α-HL channel and will be reflected in the current blockage level as a function of the event duration, producing two distinct populations of events. This ability to discern the relative location of the DNA damage site is a first step towards location-specific DNA damage detection. Further tailoring of the probe sequence with additional chemical modifications might be used to enhance the selectivity for the DNA damage site, providing a powerful tool in DNA damage detection and characterization.

## Experimental Section

### DNA preparation and purification procedures

The oligodeoxynucleotides were synthesized from commercially available phosphoramidites (Glen Research, Sterling, VA) by the DNA-Peptide Core Facility at the University of Utah. After synthesis, each oligodeoxynucleotide was cleaved from the synthetic column and deprotected according to the manufacturer's protocols, followed by purification using a semi-preparation ion-exchange HPLC column with a linear gradient of 25% to 100% B over 30 min while monitoring absorbance at 260 nm (A = 20 mM Tris, 1 M NaCl pH 7 in 10% $CH_3CN$/90% $ddH_2O$, B = 10% $CH_3CN$/90% $ddH_2O$, flow rate = 3 mL/min). The purities of the oligodeoxynucleotides 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. The identities for the oligodeoxynucleotides containing the modified bases OG and I were confirmed by ESI-MS.

### $T_m$ analysis on dsDNA samples

The dsDNA samples were formed by placing each strand in buffer (10 mM PBS, pH 7.4, 1 M KCl) at a 1 μM concentration, followed by heating the sample at 90 °C for 5 min, then allowing the samples to cool slowly to room temperature over 3 h. Next, the samples were loaded into $T_m$ analysis cuvettes following the manufacturer's protocol (Beckman DU 650) and placed into a UV/vis spectrophotometer equipped with a temperature-regulated heat

block. Samples were thermally equilibrated at 25 °C for 20 min followed by heating to 70 °C at a rate of 0.5 °C/min. As the samples were heated, absorbance readings at 260 nm were taken every 0.5 min. The background corrected data were plotted and the $T_m$ values were determined using two-point average analysis. Since refinement to determine the best possible probes occurred after conducting the translocation experiments, the $T_m$ studies for the 10 mer, 12 mer, and 15 mers were not collected under the same experimental conditions as the probes containing the modified base I; thus the data do not appear to be self consistent, and this inconsistency in $T_m$ analysis has been previously described.[42] Because of this experimental error, only data collected under identical conditions were directly compared to generate $\Delta T_m$ values.

## Chemicals and materials for nanopore analysis

All aqueous solutions were prepared with ultrapure water obtained from a Barnstead E-pure water purifier, with a resistance > 18 MΩ. KCl, $K_2HPO_4$, $KH_2PO_4$, EDTA, and HCl were used as received. A 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH 7.4) buffered electrolyte solution was prepared and filtered using a sterile 0.22 μm Millipore vacuum filter before use, and was used for all unzipping experiments. Wild-type α-hemolysin (referred to as α-HL above) was obtained as a monomer as a lyophilized powder from Sigma-Aldrich. The α-HL was frozen in ultrapure water at a concentration of 1 mg per mL for long term storage in a −80 °C freezer, and upon use was diluted with buffered electrolyte and added to the experimental cell. The phospholipid 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) was purchased as a powder from Avanti Polar Lipids and stored in a −20 °C freezer. Before use, the DPhPC powder was dispersed in decane for a concentration of 10 mg DPhPC per mL decane. Glass nanopore membranes (GNM) were fabricated as previously described,[43] and used as a solid support for a suspended bilayer for ion channel reconstitution. Before use, GNMs were chemically modified via silanization with 2% (v:v) 3-cyanopropyldimethylchlorosilane in acetonitrile.[28] Upon use, GNMs were rinsed inside and out with acetonitrile, ethanol, and water, before being filled with buffered electrolyte. Ag/AgCl electrodes were prepared from 0.25 mm diameter silver wire being soaked in bleach. All DNA oligomers were obtained as described above, and were annealed by mixing the 65 mer and probe at a 1:1 ratio, placing DNA mixture in a 90 °C water bath, and allowing sample to cool slowly to room temperature. When not in use, the annealed DNA was stored at 4 °C.

## Ion Channel Recordings

Current-time (*i*-*t*) measurements were performed for the unzipping experiments with 10 mer and 15 mer probes using a Dagan Corporation CHEM-CLAMP (Voltammeter and Amperometer Voltage Clamp Amplifier) and a Pine Instrument Company RDE4 Analog Bipotentiostat, interfaced with a PC. An in-house written LabVIEW 802 (National Instruments) program was used to record the *i*-*t* traces. Data for the 12 mer duplexes were collected using a custom built high-impedance, low noise amplifier and data acquisition system (Electronic Bio Sciences, San Diego CA).

The GNM was rinsed with ethanol and ultra pure water prior to use, then filled with buffered electrolyte. An Ag/AgCl electrode was positioned inside the GNM and the back of the GNM was sealed using a Dagan Corporation pipette holder, which was attached to a pressure gauge and 10 mL gas-tight syringe (Hamilton). A second Ag/AgCl electrode was positioned within the experimental cell; the same buffered electrolyte within the GNM was used to fill the cell and submerge the GNM orifice, α-HL was then added to the cell (external to the GNM). A voltage was applied across the GNM orifice and the resultant current was measured as a function of time.

A suspended bilayer was generated by depositing a 10 mg DPhPC per mL decane lipid solution across the GNM orifice, which produced a drop in conductance as a voltage was applied across the GNM orifice; an open GNM orifice has a resistance of ~10 MΩ and the presence of a bilayer increases the resistance to ~100 GΩ. A pressure was then applied to the back of the GNM to thin the lipid solution into a functional bilayer for protein channel reconstitution to occur.[28] After protein channel insertion, the annealed DNA was added to the experimental cell to a final concentration of 5 μM, and a voltage was applied to drive the DNA through the channel (*cis* vs *trans*, which is equivalent to external vs internal solution). A minimum of 500 translocation events were collected for each sample, expect the 15 mer samples due the stability of the duplex as discussed above. Data were collected with a 20 kHz filter and sampled at 100 kHz.

### Data Analysis

Only events that were >1 ms in duration and produced ≥ 75% blocking to the open channel current were analyzed. Histograms of event duration were plotted and fit as a single exponential decay for either each population as a whole to determine τ, or as two populations to determine $\tau_{3'}$ and $\tau_{5'}$. When the data were fit as whole populations to determine τ, the fit excluded the first bin to avoid weighting the fit toward faster event durations due to unduplexed DNA. Events where extracted using QuB (version 1.5.0.31) and fit using OriginPro (ver 8). Density plots were generated using data analysis programs provided by Electronic Bio Sciences, San Diego.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by a seed grant from the University of Utah and a research grant from the National Institutes of Health (R01 GM093099). Instruments and software donated by Electronic Bio Sciences, San Diego, are gratefully acknowledged.

## References

1. Cadet J, Poulsen H. Free Radical Biol Med. 2010; 48:1457–1459. [PubMed: 20227488]
2. Delaney JC, Essigmann JM. Chem Res Toxicol. 2008; 21:232–252. [PubMed: 18072751]
3. Cadet J, Douki T, Ravanat JL. Free Radical Biol Med. 2010; 49:9–21. [PubMed: 20363317]
4. Azqueta A, Shaposhnikov S, Collins AR. Mutat Res, Genet Toxicol Environ Mutagen. 2009; 674:101–108.
5. Schibel AEP, An N, Jin Q, Fleming AM, Burrows CJ, White HS. J Am Chem Soc. 2010; 51:17992–17995. [PubMed: 21138270]
6. Vercoutere W, Winters-Hilt S, Olsen H, Deamer D, Haussler D, Akeson M. Nat Biotechnol. 2001; 19:248–252. [PubMed: 11231558]
7. Vercoutere WA, Winters-Hilt S, DeGuzman VS, Deamer D, Ridino SE, Rodgers JT, Olsen HE, Marziali A, Akeson M. Nucleic Acids Res. 2003; 31:1311–1318. [PubMed: 12582251]
8. Winters-Hilt S, Vercoutere W, DeGuzman VS, Deamer D, Akeson M, Haussler D. Biophys J. 2003; 84:967–976. [PubMed: 12547778]
9. Dudko OK, Mathe J, Szabo A, Meller A, Hummer G. Biophys J. 2007; 92:4188–4195. [PubMed: 17384066]
10. Mathé J, Visram H, Viasnoff V, Rabin Y, Meller A. Biophys J. 2004; 87:3205–3212. [PubMed: 15347593]
11. McNally B, Wanunu M, Meller A. Nano Lett. 2008; 8:3418–3422. [PubMed: 18759490]

12. Muzard J, Martinho M, Mathé J, Bockelmann U, Viasnoff V. Biophys J. 2010; 98:2170–2178. [PubMed: 20483325]
13. Liu A, Zhao Q, Krishantha DMM, Guan X. J Phys Chem Lett. 2011; 2:1372–1376. [PubMed: 21709813]
14. Sauer-Budge AF, Nyamwanda JA, Lubensky DK, Branton D. Phys Rev Lett. 2003; 90:238101-1–238101-4. [PubMed: 12857290]
15. Sutherland TC, Dinsmore MJ, Kraatz HB, Lee JS. Biochem Cell Biol. 2004; 82:407–412. [PubMed: 15181475]
16. Viasnoff V, Chiaruttini N, Bockelmann U. Eur Biophys J. 2009; 38:263–269. [PubMed: 18836709]
17. McAuley-Hecht KE, Leonard GA, Gibson NJ, Thomson JB, Watson WP, Hunter WN, Brown T. Biochemistry. 1994; 33:10266–10270. [PubMed: 8068665]
18. Lipscomb LA, Peek ME, Morningstar ML, Verghis SM, Miller EM, Rich A, Essigmann JM, Williams LD. Proc Natl Acad Sci U S A. 1995; 92:719–723. [PubMed: 7846041]
19. Maskos K, Gunn BM, LeBlanc DA, Morden KM. Biochemistry. 1993; 32:3583–3595. [PubMed: 8385483]
20. Carbonnaux C, Van der Marel GA, Van Boom JH, Guschlbauer W, Fazakerley GV. Biochemistry. 1991; 30:5449–5458. [PubMed: 2036413]
21. Nikonowicz EP, Gorenstein DG. Biochemistry. 1990; 29:8845–8858. [PubMed: 2271561]
22. Patel DJ, Kozlowski SA, Ikuta S, Itakura K. Biochemistry. 1984; 23:3207–3217. [PubMed: 6466638]
23. Plum GE, Grollman AP, Johnson F, Breslauer KJ. Biochemistry. 1995; 34:16148–16160. [PubMed: 8519772]
24. Berashevich J, Chakraborty T. J Chem Phys. 2009; 130:015101. [PubMed: 19140634]
25. Allawi HT, SantaLucia J. Biochemistry. 1998; 37:2170–2179. [PubMed: 9485363]
26. Cheng JW, Chou SH, Reid BR. J Mol Biol. 1992; 228:1037–1041. [PubMed: 1474575]
27. Gao X, Patel DJ. J Am Chem Soc. 1988; 110:5178–5182.
28. White RJ, Ervin EN, Yang T, Chen X, Daniel S, Cremer PS, White HS. J Am Chem Soc. 2007; 129:11766–11775. [PubMed: 17784758]
29. Kawano R, Schibel AEP, Cauley C, White HS. Langmuir. 2008; 25:2850–2855.
30. Ervin EN, Kawano R, White RJ, White HS. Anal Chem. 2008; 80:2069–2076. [PubMed: 18293946]
31. Ervin EN, Kawano R, White RJ, White HS. Anal Chem. 2008; 81:533–537. [PubMed: 19140775]
32. Lathrop DK, Ervin EN, Barrall GA, Keehan MG, Kawano R, Krupka MA, White HS, Hibbs AH. J Am Chem Soc. 2010; 132:17992–17995. [PubMed: 21138270]
33. Chen Q, Liu J, SchibelAEPWhite HS, Wu C. Macromolecules. 2010; 43:10594–10599.
34. Urakawa H, El Fantroussi S, Smidt H, Smoot JC, Tribou EH, Kelly JJ, Noble PA, Stahl DA. Appl Environ Microbiol. 2003; 69:2848–2856. [PubMed: 12732557]
35. Renner S, Bessonov A, Gerland U, Simmel FC. J Phys: Condens Matter. 2010; 22:454119–454127. [PubMed: 21339606]
36. Purnell RF, Mehta KK, Schmidt JJ. Nano Lett. 2008; 8:3029–3034. [PubMed: 18698831]
37. Howorka S, Cheley S, Bayley H. Nat Biotechnol. 2001; 19:636–639. [PubMed: 11433274]
38. Mathé J, Aksimentlev A, Nelson DR, Schulten K, Meller A. Proc Natl Acad Sci U S A. 2005; 102:12377–1282. [PubMed: 16113083]
39. Akeson M, Branton D, Kasianowicz JJ, Brandin E, Deamer DW. Biophys J. 1999; 77:3227–3233. [PubMed: 10585944]
40. Henrickson SE, Misakian M, Robertson B, Kasianowicz JJ. Phys Rev Lett. 2000; 85:3057–3060. [PubMed: 11006002]
41. Howorka S, Movileanu L, Braha O, Bayley H. Proc Natl Acad Sci U S A. 2001; 98:12996–13001. [PubMed: 11606775]
42. Mergny JL, Lacroix L. Oligonucleotides. 2003; 13:515–537. [PubMed: 15025917]

43. Zhang B, Galusha J, Shiozawa PG, Wang G, Bergren AJ, Jones RM, White RJ, Ervin EN, Cauley C, White HS. Anal Chem. 2007; 79:4778–4787. [PubMed: 17550232]

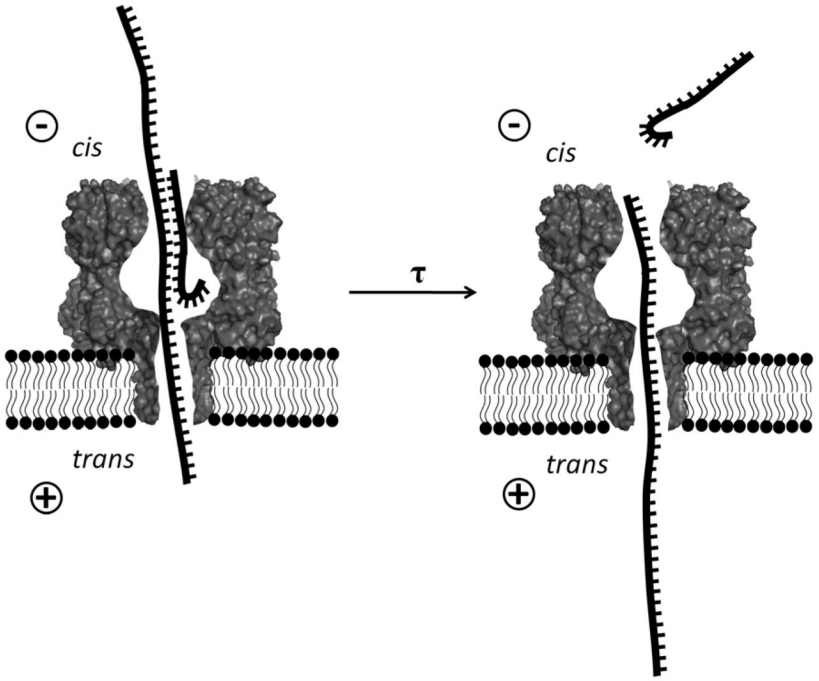


**Figure 1.**
Unzipping of hybridized DNA as it translocates through an α-HL ion channel. A target sequence is duplexed to a short probe sequence; the short probe dissociates upon translocation of the target sequence through the ion channel.

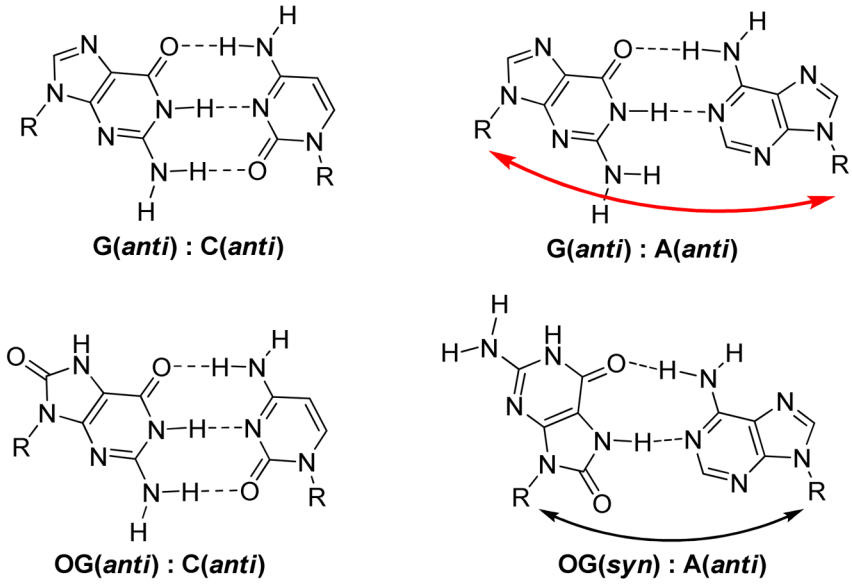


**Figure 2.**

Base pairing scheme of G vs. OG opposite C or A. The sugar-phosphate backbone distance (arrows) of a G(*anti*):A(*anti*) mismatch (red arrow) is wider than that of a OG(*syn*):A(*anti*) pair or a Watson-Crick G:C pair, accounting for its lower stability in the duplex.

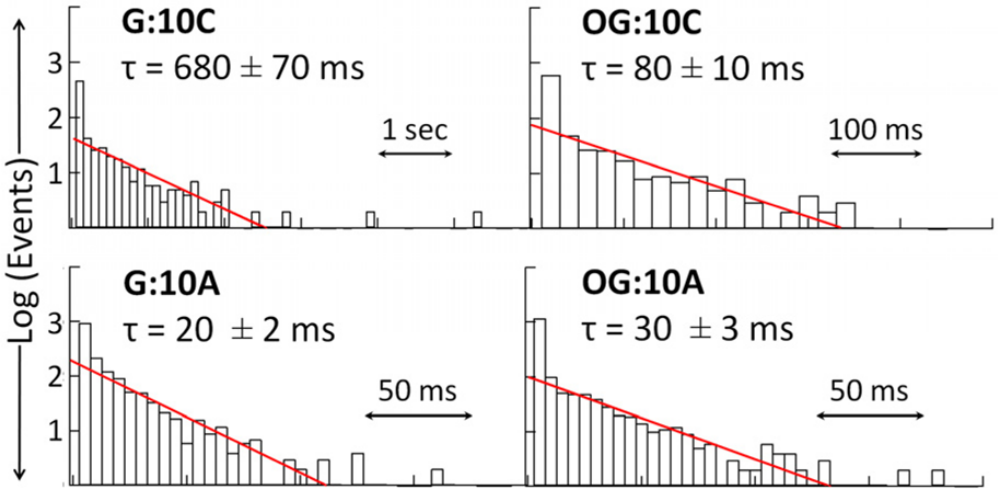


**Figure 3.**

Plot of $\log_{10}$(Events) vs. event duration for unzipping of duplexes formed with a 65 mer containing G or OG at position 33 and either C or A centered opposite in the 10 mer. The slope of the straight line is used to compute the unzipping time constant, τ, at 80 mV, *trans* vs. *cis*.

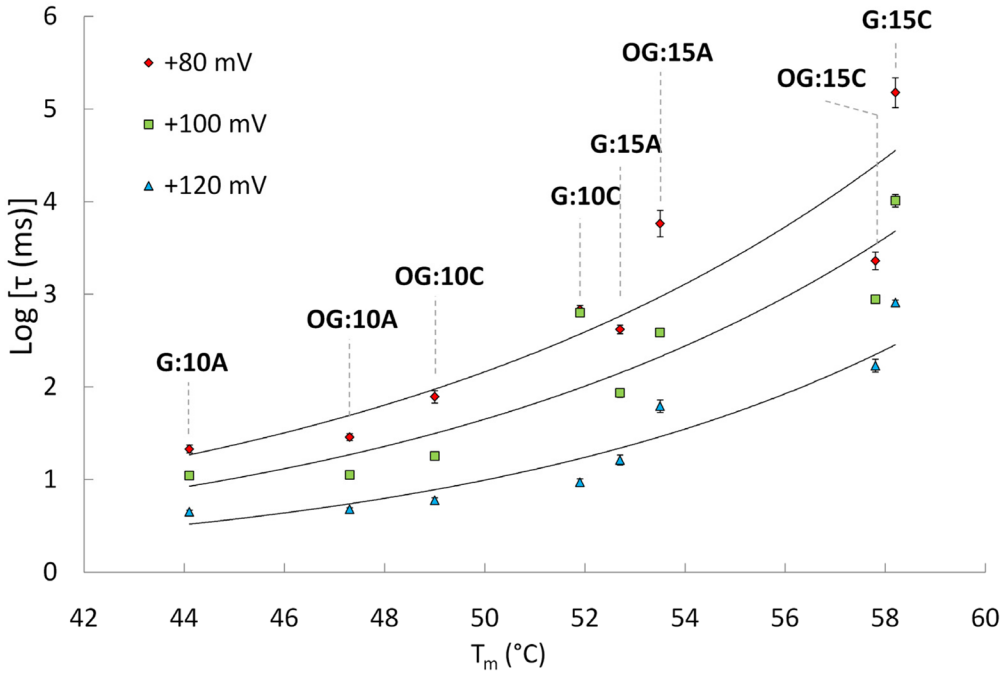


**Figure 4.**

Plot of $\log_{10}[\tau(ms)]$ as a function of the duplex melting temperature for G:10C or A and OG:15C or A, where 10 and 15 refer to the length of the probe. Sequences are given in Table 1. Data are shown for applied biases of 80 mV, 100 mV, and 120 mV, *trans* vs. *cis*.

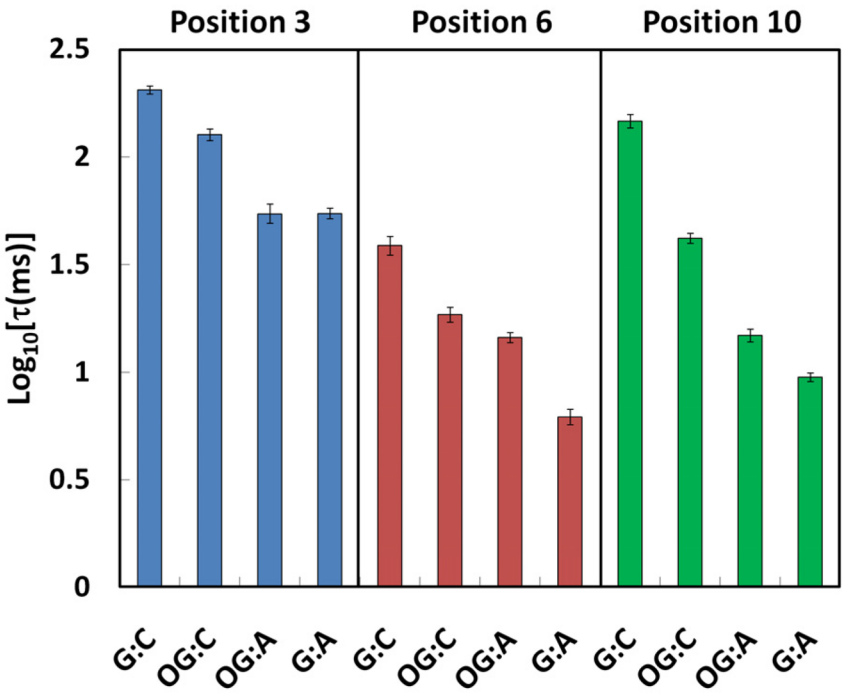


**Figure 5.**
Plot of τ as a function of position for the X:Y site of interest. Y position is denoted relative to the 5′ end of the 12-mer probe sequence. Sequences are shown in Table 2.

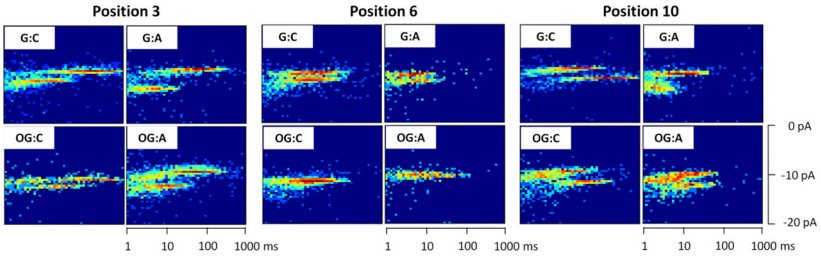


**Figure 6.**

Plots showing the event population density for the current blockage level of 12 mer probes annealed to 65 mer target strands as a function of event duration for an applied voltage of 80 mV (*trans vs. cis*).

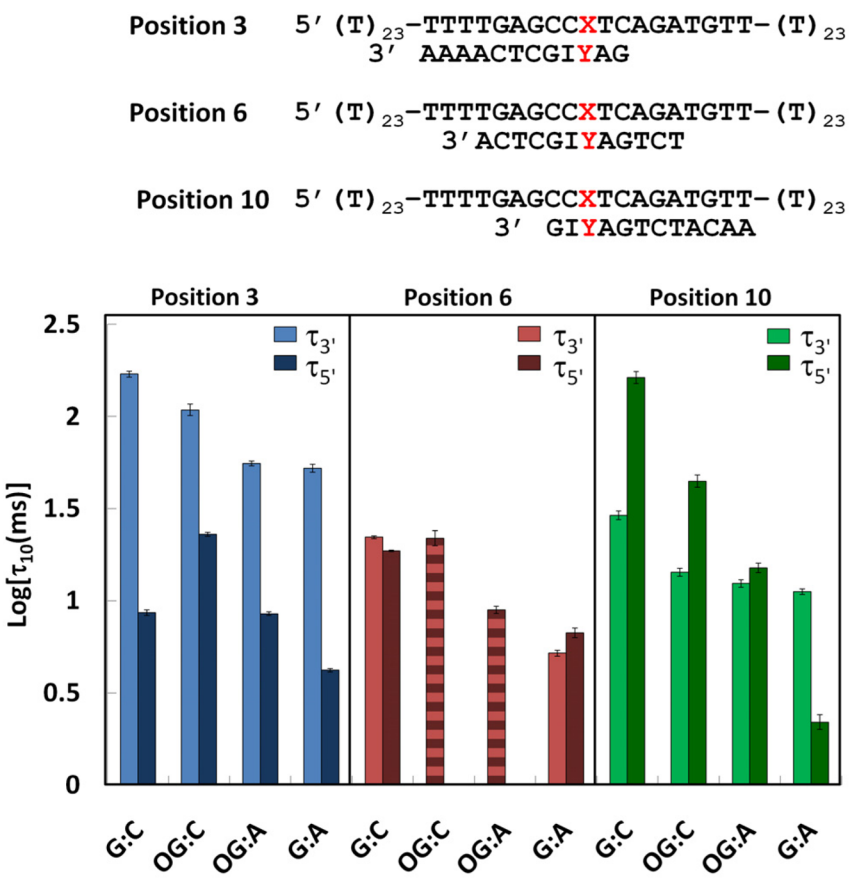


**Figure 7.**

Plot of $\tau_{3'}$ (light bars) and $\tau_{5'}$ (dark bars) as a function of position for the X:Y pair. The assignment of $\tau_{3'}$ and $\tau_{5'}$ is based on the expected event duration times for duplex unzipping; see text for further interpretation of the data. Position 6 OG:C and OG:A duplexes could not be resolved into two distinct populations and were only assigned a single τ (striped bars).

## Table 1

Melting temperatures ($T_m$) for 10 mer and 15 mer complementary probes annealed to a target 65 mer oligodeoxynucleotide.

| **5′-(T)$_{25}$-TTGAGCCXTCAGATG-(T)$_{25}$ 10 mer: 3′-CTCGGYAGTC** | | |
|---|---|---|
| **Sequence** | **$T_m$ (°C)** | **$\Delta T_m$ (°C)** |
| **X**=G, **Y**=C | 51.9 ± 0.6 | -- |
| **X**=G, **Y**=A | 44.1 ± 0.6 | −7.8 ± 0.8 |
| **X**=OG, **Y**=C | 49.0 ± 1.6 | −2.9 ± 1.7 |
| **X**=OG, **Y**=A | 47.3 ± 1.3 | −4.6 ± 1.4 |
| 5′-(T)$_{25}$-TTGAGCC**X**TCAGATG-(T)$_{25}$ 15 mer: 3′-AACTCGG**Y**AGTCTAC | | |
| Sequence | $T_m$ (°C) | $\Delta T_m$ (°C) |
| **X**=G, **Y**=C | 58.2 ± 1.0 | -- |
| **X**=G, **Y**=A | 52.7 ± 0.6 | −5.5 ± 1.2 |
| **X**=OG, **Y**=C | 57.8 ± 0.5 | −0.4 ± 1.1 |
| **X**=OG, **Y**=A | 53.5 ± 1.2 | −4.7 ± 1.6 |

## Table 2

$T_m$ studies of 12 mer complementary probes with various positions of X:Y

| | | |
|---|---|---|
| **A. Position 3**[a]<br>5`-$(T)_{23}$-TTTTGAGCC**X**TCAGATGTT-$(T)_{23}$<br>3`-AAAACTCG**IY**AG[b] | | |
| Sequence | $T_m$ (°C) | $\Delta T_m$ (°C) |
| **X**=G, **Y**=C | 54.2 ± 0.4 | -- |
| **X**=G, **Y**=A | 48.1 ± 1.3 | -6.2 ± 1.4 |
| **X**=OG, **Y**=C | 52.8 ± 0.3 | -1.4 ± 0.5 |
| **X**=OG, **Y**=A | 50.5 ± 0.3 | -3.7 ± 0.5 |
| **B. Position 6**<br>5`-$(T)_{23}$- TT TTGAGCC**X**TCAGATGTT -$(T)_{23}$<br>3`-ACT CG **IY**AGTCT | | |
| Sequence | $T_m$ (°C) | $\Delta T_m$ (°C) |
| **X**=G, **Y**=C | 55.0 ± 0.1 | -- |
| **X**=G, **Y**=A | 50.2 ± 0.3 | -4.8 ± 0.3 |
| **X**=OG, **Y**=C | 53.0 ± 1.3 | -2.0 ± 1.3 |
| **X**=OG, **Y**=A | 51.0 ± 0.1 | -4.0 ± 0.1 |
| **C. Position 10**<br>5`-$(T)_{23}$- TT TTGAGCC**X**TCAGATGTT -$(T)_{23}$<br>3`-G**IY**AGTCTACAA | | |
| Sequence | $T_m$ (°C) | $\Delta T_m$ (°C) |
| **X**=G, **Y**=C | 51.8 ± 0.3 | -- |
| **X**=G, **Y**=A | 43.5 ± 0.7 | -8.3 ± 0.8 |
| **X**=OG, **Y**=C | 48.8 ± 0.4 | -3.0 ± 0.5 |
| **X**=OG, **Y**=A | 47.2 ± 0.3 | -4.6 ± 0.4 |

[a] The probe strand is shifted 5′ (A), centered (B), or shifted 3′ (C) with respect to the interrogated base **X**. The position of **Y** is denoted as position 3, 6 or 10, accordingly.

[b] I = inosine