UNZIPPING KINETICS OF DUPLEX DNA CONTAINING LESIONS IN A NANOPORE AND ITS APPLICATION IN MONITORING AN ENZYME REACTION

by

Qian Jin

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

August 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Qian Jin**

has been approved by the following supervisory committee members:

Henry S. White	, Chair	**Mar. 28, 2013** Date Approved
Cynthia J. Burrows	, Member	**Mar. 28, 2013** Date Approved
Marc D. Porter	, Member	**Mar. 28, 2013** Date Approved
Jennifer S. Shumaker-Parry	, Member	**Mar. 28, 2013** Date Approved
Jordan M. Gerton	, Member	**Mar. 28, 2013** Date Approved

and by **Henry S. White**, Chair of the Department of **Chemistry**

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

The dissertation presents biophysical studies of duplex DNA during unzipping in a protein ion channel and its application in the fast readout of enzyme activity. Chapter 1 reviews a variety of approaches to enhance the resolution of nucleotide characterization using nanopore methodology, providing insights into future experimental design. This chapter also overviews the work carried out at University of Utah concerning the characterization of DNA damage using nanopore methods.

Chapter 2 examines the unzipping kinetics of lesion-containing duplexes in the protein channel α-hemolysin (α-HL). The voltage-driven unzipping of individual duplex DNA molecules was investigated by pulling the tail of duplex into the α-HL and monitoring the temporal blockage to the ion flow through the channel. The unzipping occurs as a first-order reaction or sequential first-order reactions as determined from the time duration histograms of the blockages.

Chapter 3 extends the unzipping studies of duplex to a larger pool of lesions. As a more destabilizing lesion is introduced into the duplex, the process of strand dissociation grows faster and evolves from a first-order reaction to two sequential first-order reactions. The kinetic stability of duplexes was interpreted in terms of the lesion-induced distortion of duplex backbone as the major factor

and the number of hydrogen bonds in the modified base pair as the minor factor.

Chapter 4 provides an application of duplex unzipping in the measurement of enzyme activity. Specifically, the conversion of uracil to an abasic site by uracil DNA glycosylase (UDG) was monitored using ion-channel recordings based on the difference in blockage current. The single-nucleotide discrimination during UDG digestion was performed in the duplex context at the latch region of the α-HL. The data suggest that the protein latch is a new sensing zone in α-HL, specifically useful for duplex analysis.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank first my Ph.D. advisor, Professor Henry S. White, who taught me how to become a professional and what research is all about. In the past five years, he has always pointed me in the right direction with his tremendous wisdom and patience. I am also very grateful for his generosity in sharing his knowledge and experience, which makes it truly fruitful and enjoyable to learn from him.

I have complete admiration for my graduate committee, Prof. Marc D. Porter, Prof. Jennifer S. Shumaker-Parry, Prof. Jordan M. Gerton, and most especially, Prof. Cynthia J. Burrows, who also mentored me in the collaboration with her unfailing intelligence and sensitivity. Special thanks to insightful Dr. Aaron Fleming, who contributed numerous research ideas to me. I am so fortunate to have his help with DNA design and synthesis. Without him, my Ph.D. would have taken much longer to complete.

Every single person that I have worked with has been phenomenal. Dr. James Muller is always supportive. Dr. Geoffrey Barral and Dr. Eric N. Ervin at Electronic Biosciences are ready to help at any time. My life in Utah would be much less fun without my former and current coworkers: Dr. Anna E. P. Schibel, Dr. Deric A. Holden, Dr. Wenjie Lan, Dr. Jin Liu, Mr. Jiewen Xiong,

Mr. Long Luo, Dr. Robert Johnson, Mr. Rukshan Perera, Ms. Cherie Tan, and Dr. John Watkins. I owe thanks to Dr. Na An, Ms. Yun Ding, Ms. Ania Wolna, and Mr. Lidong He for our pleasant collaborations.

And finally, I am grateful to my parents, who constantly assure me the road ahead is hopeful and safe. Without them, I would truly be lost.

CHAPTER 1

INTRODUCTION

In the past two decades, a nanoscale pore with a size comparable to nucleic acids has been investigated as a promising tool to read a DNA sequence when a strand translocates through it. The proposed method of nanopore sequencing is simple without labels or amplification: a DNA strand is driven electrophoretically into a nanopore as the amplitude of ion current flowing through the nanopore is recorded.[1] Ideally, the ion current should have four levels, each corresponding to one of the nucleotides (A, T, C, or G).[2] The potential of using a nanopore as the next-generation sequencing platform has attracted enormous attention in both academia and industry, resulting in a large number of studies on DNA transport through nanopores. The $1000 goal of mapping the human genome has not yet been achieved using nanopores; however, nanopore research has grown tremendously, providing a new single-molecule approach to investigate nucleic acids.

1.1 Increasing the Resolution of Nanopore Ion-Channel Recordings

The nanopore ion-channel recordings are based on the Coulter-counter technique: two electrodes are placed across the nanopore membrane that

separates the electrolyte reservoir (Figure 1.1). The analyte molecule is electrophoretically driven into the nanopore channel, causing a transient decrease (pulse) in ion current flowing through the pore. As the current is recorded as a function of time, the amplitude and duration of the pulses are used to characterize analyte molecules. Specifically, for a homogeneous single-stranded DNA (ssDNA), the base identity can be determined from the pulse amplitude while the DNA length is related to the pulse duration. It has been found that DNA and RNA homopolymers, such as poly(A), poly(C), and poly(U), are distinguishable based on the pulse amplitude.[1] For a ssDNA between 12 and 100 nucleotides long, the dwell time of the pulse is proportional to the strand length if no unraveling process is involved.[2]

A simple-minded expectation on DNA translocation though a nanopore is that the pulse should include different levels, each distinct to a specific nucleotide; however, the real current signal collected does not have the resolution at the single nucleotide level. In order to increase the signal-to-noise ratio for nanopore ion-channel recordings, three aspects have been manipulated: (1) the nanopore structure, (2) the DNA structure, and (3) the experimental conditions. These aspects are discussed below.

1.1.1 Biological Nanopores vs. Solid-State Nanopores

Nanopore researchers work on either biological nanopores or solid-state nanopores, each having pros and cons. Solid-state nanopores are commonly made with silica nitride or silicon by electron-beam lithography.[3] The modern

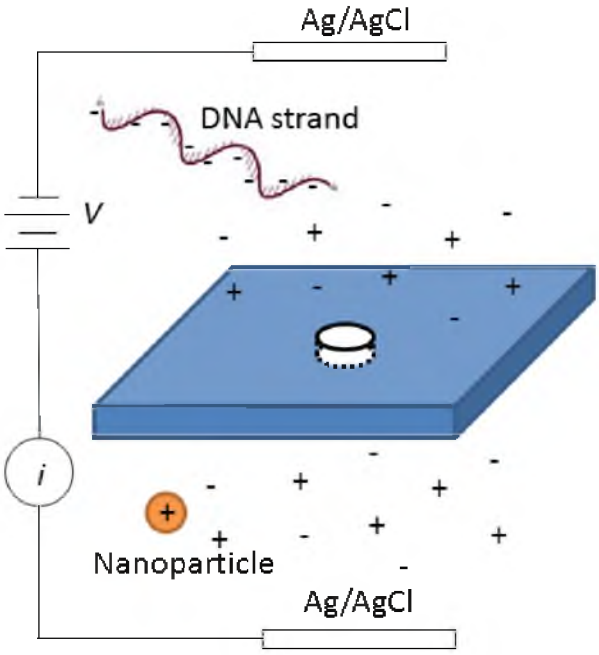


Figure 1.1. Schematics of nanopore ion-channel recordings. Ag/AgCl electrodes are placed one on each side of the nanopore membrane. Driven by electrophoresis, charged molecules, such as DNA strands and nanoparticles, translocate through the nanopore, causing a decrease in ion flux through the pore.

nano-fabrication technology allows solid-state nanopores to have a variety of advantages, such as tunable pore sizes and shapes, durability and robustness to a wide range of experimental conditions, the ability to form high-throughput nanopore arrays, as well as the versatility to integrate with electronics.[4]

The resolution of solid-state nanopores, however, is not as good as that of biological nanopores made with self-assembled proteins inserted in a lipid bilayer.[5] The most studied trans-membrane biological nanopore, α-hemolysin (α-HL), is a heptameric protein channel that comprises a vestibule connected to a narrower β-barrel.[6] The central constriction where the vestibule meets β-barrel is 1.4 nm in diameter, allowing only ssDNA to pass. Bigger DNA structures, such as hairpins and dsDNA, need to unzip to translocate through α–HL. The reproducibility of nanopore size and shape for biological pores has not been achieved by the solid-state pores. In addition, mutagenesis allows tweaking of the identity of amino acid components, enabling the sharpening of the protein sensing zone and thus enhancing the measurement sensitivity.[7] The drawbacks of biological nanopores involve the difficulties of handling fragile lipid bilayers and maintaining a single protein channel.

Our laboratory at the University of Utah has developed a glass nanopore membrane (GNM) as the support structure for lipid bilayers (Figure 1.2). Compared with the commonly used Teflon solid support (orifice diameter 10-100 μm), the GNM aperture diameter is one to two orders of magnitude smaller.[8] A consequence is that the smaller lipid bilayer supported by GNM is more resistive to environmental disturbance. Since charging the lipid bilayer is the major source

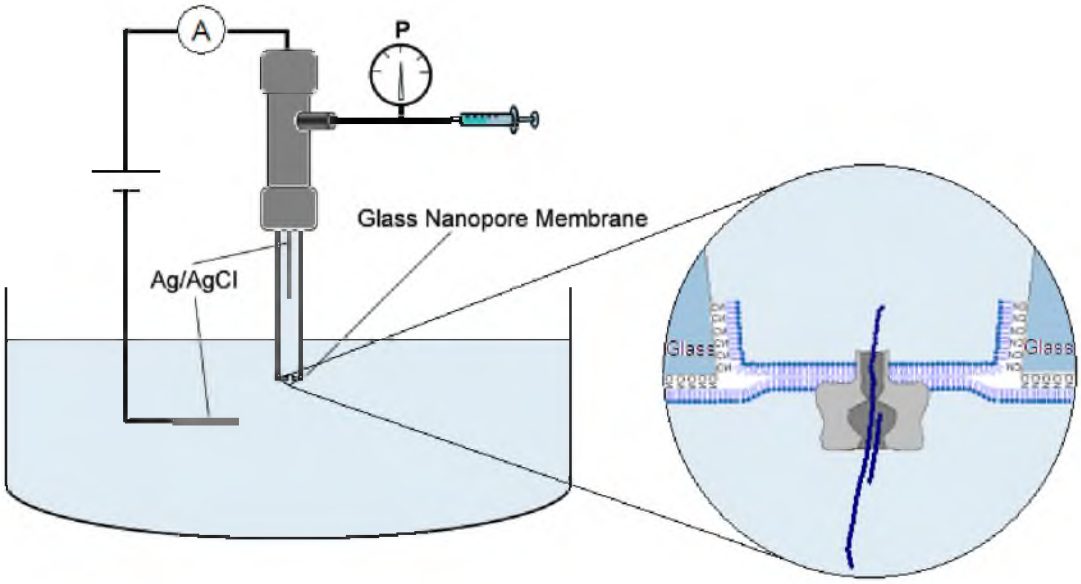


Figure 1.2. Schematics of experimental setup for ion-channel recordings using a GNM. A Ag/AgCl electrode was placed inside the GNM capillary (the *trans* compartment), which is connected to a gas tight syringe and a pressure gauge. Another Ag/AgCl electrode was placed outside the GNM capillary (the *cis* compartment) in a 300 μL cell. Both the *cis* and *trans* compartments of GNM were filled with 1 M KCl, 25 mM tris•HCl, and 1 mM EDTA (pH 7.9). An α-HL channel was formed in the lipid bilayer across the orifice of GNM and is used as a stochastic sensor for DNA characterization.

of system noise in ion-channel recordings, GNM contributes to a lower noise as a result of shrinking the area of the bilayer capacitor.[9] Stability studies performed in our laboratory have shown that the bilayer supported by GNM can last two weeks, and ion-channel recordings with α-HL can be achieved for 24 h without further optimization.[10]

1.1.2 Building Structural Obstacles to Slow Down DNA

At 120 mV in 1 M KCl, each nucleotide of a strand takes less than 10 μs to translocate through α-HL, during which time there are only 8 electrons in the circuit contributing to the electrical signal.[1] The fast translocation speed of a strand makes it very difficult to achieve single nucleotide resolution. To slow down DNA, the structure of ssDNA has been modified to incorporate obstacles in the forms of (1) streptavidin-biotin anchors, (2) hairpins, (3) complementary strands, (4) DNA-binding proteins, and (5) polymerase ratchets. [11-14]

1.1.2.1 Streptavidin-Biotin Anchors

Each streptavidin molecule can bind up to four strands of biotinylated ssDNAs. Once a biotinylated strand is captured by the pore, the streptavidin, whose size is bigger than the α-HL entrance, prevents the strand from further threading. As a result, the DNA strand is immobilized into α-HL for an infinite time until the polarity is switched to eject the streptavidin-biotin-DNA complex. Though this strategy does not enable the readout of a full strand, the long observation time during immobilization significantly enhances the resolution. Different DNA

nucleotides located in the otherwise same DNA strands can be discriminated in the immobilized form if the point of interest falls into the recognition site of α-HL.[15] The recognition site of α-HL will be discussed in Chapter 4.

1.1.2.2 Hairpins and Complementary Strands

Since the narrowest part of α-HL is smaller than double-stranded DNA, hairpins and duplexes need to unzip in order to translocate through the pore.[16] The unzipping process, which involves breaking hydrogen bonds between base pairs, requires a time that is one to three orders of magnitude longer than simple translocation. The duration of an unzipping event is generally determined by the sequence, specifically the ratio of G:C base pairs compared to A:T and the number of base pairs. The strand dissociation of duplexes will be investigated in Chapters 2 and 3.

1.1.2.3 DNA-Binding Proteins

DNA-binding proteins have an affinity, either general or specific, for ssDNA or dsDNA. Nanopore experiments with poly(A) binding proteins have shown that the translocation time of the bound strand is two orders of magnitude longer than the unbound free stand.[17] In these measurements, the removal of binding proteins is required to ensure a complete translocation.

1.1.2.4 Polymerase Ratchets

Polymerase is a progressive enzyme that moves along the DNA strand one nucleotide at a time while implementing replication. The progression speed of polymerases in bulk is usually milliseconds per nucleotide step. By coupling the polymerase with a partially double-stranded DNA in α-HL, the strand can be ratcheted slowly along the channel at the speed equivalent to the enzyme sliding rate in bulk. Using phi29 DNA polymerase, DNA strands can thread through the pore in a slow-moving fashion at 25-400 milliseconds per nucleotide, a speed reasonable enough to obtain single-nucleotide resolution.[18] This provides a promising strategy for a slow and even-speed readout on each nucleotide of the DNA strand.

1.1.3 Varying Experimental Conditions

Several experimental conditions of nanopore ion-channel recordings have been studied to enhance the S/N ratio of the electrical readout of the nucleotide sequence. These include (1) filter frequency, (2) electrolyte concentration, (3) viscosity, and (4) temperature.

1.1.3.1 Filter Frequency

The temporal resolution of ion-channel recordings is limited by the filter frequency (or bandwidth) of the measurement. If a 10 kHz low-pass filter is used in a nanopore experiment, events longer than 100 μs should be able to be recorded without distortion. For a 10-mer poly(A), estimating a translocation

speed of 3 μs/nt, a 10 kHz filter is too low to distinguish nucleotides. Although a higher filter frequency can always be used to record fast events, the electrical noise increases correspondingly with higher bandwidth (the noise is proportional to the square root of bandwidth). The tradeoff between filter and noise requires careful adjustment of the bandwidth to meet criteria of specific experiments. In addition, the sampling rate needs to be at least two times of filter frequency according to Nyquist–Shannon sampling theorem.

1.1.3.2 Electrolyte Concentration

Although the open channel current grows in proportion of the salt concentration, the DNA blockage current increases slowly as the salt is more concentrated. [19, 20] The consequence is that the percentage blockade (DNA blockage current/open channel current) decreases as the salt concentration increases. As described in Chapters 2 and 3, the percentage blockade of dsDNA at 1 M KCl is as low as 10%. In contrast, the nanopore experiments in Chapter 4 have been performed at 150 mM KCl, under which condition the percentage blockade of dsDNA is ~40%.

Regardless of the salt concentration in bulk, the salt environment in the nanopore remains relatively constant based on both simulation and experimental results obtained in our laboratories. The time durations of ssDNA translocation and dsDNA unzipping remain the same at various salt concentrations. Further, it has been found that when solutions containing different salt concentrations are

placed on two sides of the nanopore, the local electrical field on one side is enhanced, resulting in a higher rate of DNA capture.[21]

1.1.3.3 Solvent Viscosity and Temperature

There are both advantages and disadvantages to varying the solvent viscosity and temperature. DNA translocation can be slowed down by increasing the solvent viscosity; however, the amplitude resolution is comprised and the overall S/N ratio may not necessarily increase.[22] On the other hand, decreasing the temperature reduces DNA translocation velocity. The difference in translocation time between DNA homopolymers at 15°C is 10 times larger than that at 40°C.[23] However, a lower temperature reduces the event rate and the amplitude resolution, and therefore is not always favored.

1.2 Probing Modified DNA Containing Lesions in a Nanopore

In real cells, DNA nucleotides undergo continuous modifications into lesions either spontaneously or by exposure to DNA damaging agents.[24] The damage occurs at a surprisingly high rate of tens of thousands of times per cell per day. Though the majority of lesions are corrected by repair enzymes, damages beyond the repair capability may lead to cell death and further to cancer, cardiovascular and aging-related diseases.[25-27]

Two principal methods of DNA damage detection commonly being used today are the comet assay and the digestion/mass spectrometry analysis. In the comet assay, fluorescent labeled broken DNA strands are released from lysed cells,

and then migrate under electrophoresis on the gel to a distance related to the strand length, forming a comet tail.[28] The comet assay has been a major DNA damage detection method with an extremely high sensitivity. The second widely used DNA detection method, the digestion/mass spectrometry analysis, digests DNA isolated from cells into individual nucleosides using phosphodiesterase and phosphatase. Those individual nucleosides are characterized by LC-MS.[29] The digestion/MS method introduces high levels of artifacts and needs to overcome the difficulties of digesting bulky base adducts. Both the comet assay and the digestion/MS method share the drawbacks of not being able to determine the lesion location or to detect multiple lesions on one DNA strand. Neither of these two methods is a single-molecule method.

During the past five years, the White group and Burrows group at University of Utah have been collaborating to study lesions and lesion-related reactions at the single-molecule level using nanopores. Below, the collaboration work is reviewed briefly.

1.2.1 Unzipping Kinetics for Duplexes that Contain Lesions

In Chapters 2 and 3, force-induced unzipping of duplex DNA has been investigated for the purpose of understanding the enzymatically mediated process of strand separation that universally occurs during DNA replication and transcription. As an emerging method of single-molecule force manipulation, the nanopore-based technique provides a fast and linker-free approach that overcomes many of the drawbacks of optical tweezers and atomic force

microscopy (AFM).[30] In this dissertation, the nanopore method has been used to study the process of strand dissociation for lesion-containing duplexes, which are performed by pulling the DNA overhang through the size-constrained α-HL channel under an electrical field in ion-channel recordings. In Chapter 2, we found that the unzipping process follows either a model of a first-order reaction (Type I) or a model of two sequential first-order reactions (Type II), depending on the extent of the destabilizing effect imparted by the lesion.[31] As discussed in Chapter 3, extensive nanopore studies on duplexes containing a different damaged base pair display an evolution of unzipping kinetics from the Type I model to the Type II model, with the duplexes containing the more damaged base pair tending to unzip faster in a two-step fashion.

In addition to studying the physical models of unzipping, we also demonstrate the application of nanopore force experiment in the measurement of uracil-DNA glycosylase (UDG) enzymatic reaction on dsDNA (Chapter 4). The commonly used approach to measure the UDG activity involves quenching aliquots of the reaction solution at a series of time intervals followed by gel electrophoresis profiling in parallel lanes.[32-34] This method is complicated by lengthy preparation procedures of radioactive labeling as well as long gel development time. In Chapter 4, we monitor the conversion of uracil to an abasic site using the methods of ion-channel recordings. Compared with the gel electrophoresis method to characterize the enzyme activity, the nanopore-based enzymatic study has developed a new technique of monitoring biological processes in a real-time, time-efficient, and linker-free way.

1.2.2 Tethering and Translocation Experiments with DNA Adducts

The nanopore approach has been used to detect DNA lesions. By chemically modifying damaged bases into adducts, detectable signals were attained when the lesion-containing DNA strands translocate through or stay immobilized within the channel.

The tethering concept has been applied to examine a single lesion, either an abasic site or oxidized guanine in the α-HL.[35,36] To amplify the signal for the lesion, the lesion site was modified with adducts of various sizes, shapes, and rigidity through chemical reactions. By tethering the modified DNA to streptavidin via a biotin linkage and capturing the DNA inside an α-HL, we found that adducts, in both homo- and heteropolymer ssDNA sequences, produced current blockage levels at 100 mV that differ by as much as ~8 pA from those of the native bases.

Remarkably, using a crown-ether adduct as a marker, the nanopore can detect a single abasic site or two abasic sites per strand as the lesion-containing strand translocates through the pore.[37] The crown-ether adducted DNA was slowed down by the bulky addition by one to two orders of magnitude, substantially enhancing the temporal resolution of measurement.

In addition to studying the unzipping kinetics and the electrical signatures of strands that contain oxidized guanine and abasic sites, secondary structures in telomeric DNA, such as guanine-quadruplex and i-motif, have been examined when unraveling in the nanopore. The electrical signature generated during unfolding the strand can be used to detect guanine-quadruplex and i-motif, the presence of which are related to 85% of all cancers.[38] Further, these secondary

structures can be constructed as obstacles to DNA translocation, causing the strand to proceed in a slower motion in the channel.

1.3 References

1. Akeson, M.; Branton, D.; Kasianowicz, J. J.; Brandin, E.; Deamer, D. W. *Biophys. J* **1999**, *77*, 3227.

2. Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. *Proc. Nat. Acad. Sci.* **2000**, *97*, 1079.

3. Storm, A. J.; Chen, J. H.; Ling, X. S.; Zandbergen, H. W.; Dekker, C. *Nature Materials* **2003**, *2*, 537.

4. Healy, K.; Schiedt, B.; Morrison, A. P. *Nanomedicine* **2007**, *2*, 875.

5. Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Schloss, J. A. *Nat. biotechnol.* **2008**, *26*, 1146.

6. Gouaux, J. E.; Braha, O.; Hobaugh, M. R.; Song, L.; Cheley, S.; Shustak, C.; Bayley, H. *Proc. Nat. Acad. Sci.* **1994**, *91*, 12828.

7. Bayley, H.; Jayasinghe, L. *Mol. Membr. Biol.* **2004**, *21*, 209.

8. Zhang, B.; Galusha, J.; Shiozawa, P. G.; Wang, G.; Bergren, A. J.; Jones, R. M.; White, R. J.; Ervin, E. N.; Cauley, C. C.; White, H. S. *Anal Chem.* **2007**, *79*, 4778.

9. Wonderlin, W. F.; Finkel, A.; French, R. J. *Biophys. J.* **1990**, *58*, 289-297.

10. White, R. J.; Ervin, E. N.; Yang, T.; Chen, X.; Daniel, S.; Cremer, P. S.; White, H. S. *J. Am. Chem. Soc.* **2007**, *129*, 11766.

11. Purnell, R. F.; Mehta, K. K.; Schmidt, J. J. *Nano Lett.* **2008**, *8*, 3029.

12. Vercoutere, W.; Winters-Hilt, S.; Olsen, H.; Deamer, D.; Haussler, D.; Akeson, M. *Nat. biotechnol.* **2001**, *19*, 248.

13. Sauer-Budge, A. F.; Nyamwanda, J. A.; Lubensky, D. K.; Branton, D. *Phys. Rev. Lett.* **2003**, *90*, 238101.

14. Lieberman, K. R.; Cherf, G. M.; Doody, M. J.; Olasagasti, F.; Kologji, Y.; Akeson, M. *J. Am. Chem. Soc.* **2010**, *132*, 17961.

15. Stoddart, D.; Heron, A. J.; Mikhailova, E.; Maglia, G.; Bayley, H. *Proc. Nat. Acad. Sci.* **2009**, *106*, 7702.

16. Mathé, J.; Visram, H.; Viasnoff, V.; Rabin, Y.; Meller, A. *Biophys. J.* **2004**, 87, 3205.

17. Lin, J.; Fabian, M.; Sonenberg, N.; Meller, A. *Biophys. J.* **2012**, *102*, 1427.

18. Cherf, G. M.; Lieberman, K. R.; Rashid, H.; Lam, C. E.; Karplus, K.; Akeson, M. *Nat. biotechnol.* **2011**, *30*, 344.

19. Smeets, R. M. M.; Keyser, U. F.; Krapf, D.; Wu, M.-Y.; Dekker, N. H.; Dekker, C. *Nano Lett.* **2006**, *6*, 89.

20. Bonthuis, D. J.; Zhang, J.; Hornblower, B.; Mathé, J.; Shklovskii, B. I.; Meller, A. *Phys. Rev. E.* **2006**, *97*, 128104.

21. Wanunu, M.; Morrison, W.; Rabin, Y.; Grosberg, A. Y.; Meller, A. *Nat. Nanotechnol.* **2009**, *5*, 160.

22. Kawano, R.; Schibel, A. E. P.; Cauley, C.; White, H. S. *Langmuir* **2009**, *25*, 1233.

23. Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. *Proc. Natl. Acad. Sci.* **2000**, *97*, 1079.

24. Loeb, L. A.; Harris, C. C. *Cancer Res.* **2008,** 68, 6863-6872.

25. Mangal, D.; Vudathala, D.; Park, J.-H.; Lee, S. H.; Penning, T. M.; Blair, I. A. *Chem. Res. Toxicol.* **2009**, *22*, 788.

26. Sohal, R. S.; Weindruch, R. *Science* **1996**, 273, 59.

27. Zawia, N. H.; Lahiri, D. K.; Cardozo-Pelaez, F. *Free Radical Biology and Medicine* **2009**, *46*, 1241.

28. Azqueta, A.; Shaposhnikov, S.; Collins, A. R. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **2009,** 674, 101.

29. Yu, H.; Venkatarangan, L.; Wishnok, J. S.; Tannenbaum, S. R. *Chem. Res. Toxicol.* **2005,** 18, 1849-1857.

30. Albrecht, C.; Blank, K.; Lalic-Mülthaler, M.; Hirler, S.; Mai, T.; Gilbert, I.; Schiffmann, S.; Bayer, T.; Clausen-Schaumann, H.; Gaub, H. E. *Science* **2003**, *301*, 367.

31. Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. *J. Am. Chem. Soc.* **2012**, *134*, 11006.

32. Souza-Pinto, N. C.; Harris, C. C.; Bohr, V. A. Oncogene 2004, 23, 6559.
33. Neddermann, P.; Jiricny, J. Proc. Natl. Acad. Sci. U.S.A. 1994, 91, 1642.

34. Blaisdell, P.; Warner, H. J. Biol. Chem. 1983, 258, 1603.

35. Schibel, A. E. P.; Fleming, A. M.; Jin, Q.; An, N.; Liu, J.; Blakemore, C. P.; White, H. S.; Burrows, C. J. **2011**, *133*, 14778-14784.

36. An, N.; White, H. S.; Burrows, C. J. *Chem. Commun.* **2012**, *48*, 11410.

37. An, N.; Fleming, A. M.; White, H. S.; Burrows, C. J. *Proc. Nat. Acad. Sci.* **2012**, *109*, 11504.

38. Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2007**, *129*, 10220-10228.

CHAPTER 2

UNZIPPING KINETICS OF DUPLEX DNA CONTAINING OXIDIZED LESIONS IN A NANOPORE

2.1 Introduction

The unzipping of double-stranded DNA (dsDNA) duplexes occurs from the forces exerted by enzymes during DNA replication, transcription, and translocation. Single-molecule manipulation methods, such as laser tweezers and atomic force microscopy (AFM), have been used to study the process of strand separation by exerting a mechanical force on dsDNAs through molecular linkers.[1,2] These techniques, though powerful, are complicated by the time-consuming process of data collection and the lengthy preparation procedure to couple the analyte molecules to the force probe. In the past decade, the nanopore method has come forth as a time-efficient and linker-free approach that allows the study of single DNA molecules, thus avoiding measurements of an ensemble of molecules that provide average values of the population.[3-12] In these measurements, the conductance of a biological or synthetic nanopore in an aqueous electrolyte is transiently reduced as the DNA passes through the nanopore. For a target single-stranded DNA (ssDNA) hybridized to a short probe

strand, the α-hemolysin (α-HL) nanopore allows the threading of the ssDNA overhang, but its constriction zone prevents translocation of the dsDNA segment.[13-15] Together with the electrical field applied across the nanopore, the narrow aperture contributes to a localized denaturing force on the DNA duplex, causing strand dissociation to occur. The time, *t,* required to open the duplex within the nanopore provides useful information on unzipping kinetics under a controlled force. Previously, the nanopore method has been applied to a variety of DNA kinetic studies, such as intramolecular unfolding of DNA hairpins, intermolecular dissociation of two complementary DNA strands, dissociation between a DNA strand and a protein, as well as interaction between DNA aptamers and targets.[16-32]

In this chapter, we present a nanopore-based study of the unzipping kinetics of DNA duplexes that contain single-lesion sites, a system of interest that provides insight into the kinetic stability of damaged DNA duplexes. As shown in Figure 2.1, the lesions of focus include the oxidatively-damaged guanine (G) product, 8-oxo-7,8-dihydroguanine (OG), and its further oxidized products the hydantoins guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp).[33,34] Studies concerning DNA base oxidation products are of particular interest due to their mutagenic potential, which is thought to be a leading contributor to age-associated diseases, such as cardiovascular diseases, cancer, and Alzheimer's disease.[35-38]

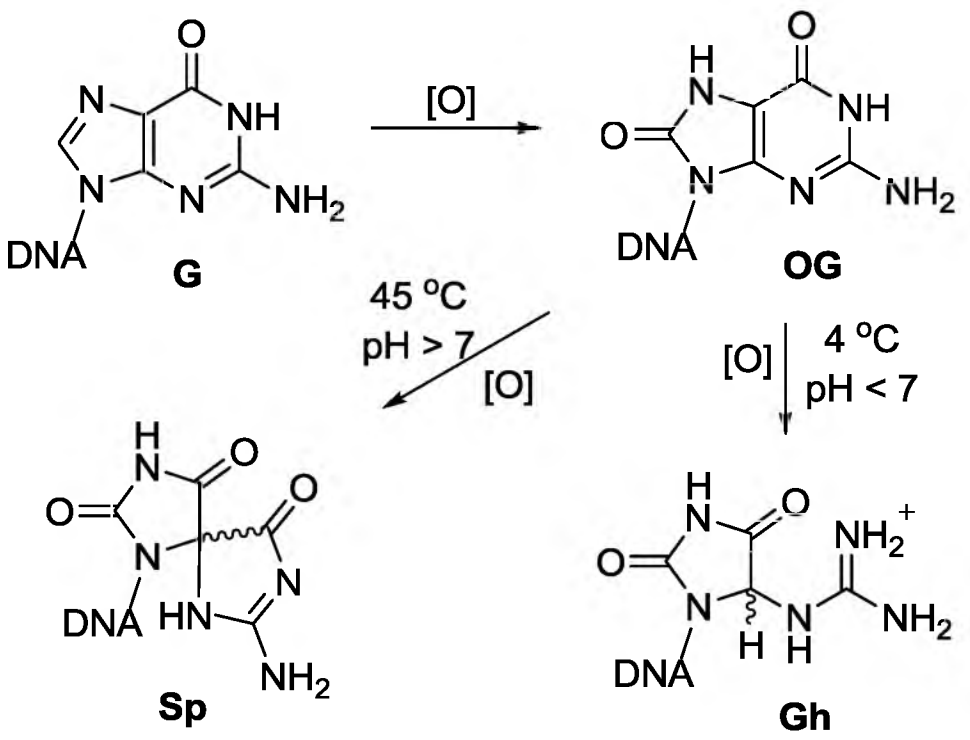


Figure 2.1. Oxidation of G yields OG, the further oxidation of which leads to the hydantoins Gh and Sp. The reaction direction from OG to Gh or Sp depends on pH, temperature, and the surrounding base stack.

2.2 Experimental Section

2.2.1 DNA Preparation and Purification Procedures

DNA was synthesized from commercially available phosphoramidites (Glen Research, Sterling, VA) by the DNA-Peptide Core Facility at the University of Utah. After synthesis, each DNA was cleaved from the synthetic column and deprotected according to the manufacturer's protocols, followed by purification using an 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 NaP_i, 1 M NaCl pH 7 in 10% CH_3CN/90% ddH_2O, B = 10% CH_3CN/90% ddH_2O, flow rate = 1 mL/min). The Gh- and Sp-containing DNAs were synthesized following previously established protocols, and purified by HPLC (see HPLC traces in SI).[39]

2.2.2 Thermal Denaturation Studies

All thermal denaturation experiments were conducted with the truncated 23-mer strand, 5'- TT TTG GAG CTG XTG GCG TAG GTT, in which X =G, OG, Sp, or Gh. By removing the poly-dT tails, the hyperchromic shift for the transition from double-stranded to single-stranded DNA was more clearly observed. First, the dsDNA was prepared by mixing the two complementary strands in a 1:1 ratio at a final concentration of 10 μM in 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH 7.4), followed by heating the sample to 90 °C, then slowly cooling to room temperature over 3 h. Next, the samples were diluted to 1 μM dsDNA concentration in 1 M KCl, then loaded into T_m analysis cuvettes following the manufacturer's protocol

and placed into a UV/vis spectrophotometer equipped with a temperature-regulated heat block. Samples were thermally equilibrated at 20 °C for 20 min followed by heating to 75 °C at a rate of 0.5 °C/min. As the samples were heated, absorbance readings at 260 nm were taken twice every minute. The background corrected data were plotted and the melting temperature (T_m) was determined using a two-point average analysis.

2.2.3 Chemicals and Materials for Nanopore Measurement

A 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH 7.4) solution was the buffer electrolyte used after being filtered by a sterile 0.22 μm Millipore vacuum filter. Wild-type α-hemolysin (lyophilized power, monomer, List Biological Laboratories) was dissolved in water at 1 mg/mL and stored at -80 °C. The phospholipid, 1,2-diphytanoyl-*sn*-glycero-3-phospho-choline (DPhPC), was purchased in a powder form from Avanti Polar Lipids and stored at -20 °C. DPhPC was dissolved in decane at 10 mg/mL before use. Glass nanopore membranes (GNMs) were used as the support structure for the lipid bilayer.[40 - 42] Before use, GNMs were chemically modified with a 2% (v/v) (3-cyano-propyl)dimethylchlorosilane in acetonitrile to produce a hydrophobic surface. All DNA oligomers were prepared as described above. The duplex DNA samples were formed by mixing target and probe strands at a 1:5 mole ratio (target vs. probe) in 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH 7.4), followed by heating in a 90 °C water bath for 5 min and then cooling slowly to room temperature over 3 h. The 1:5 mole ratio of target to probe was used in order to shift the duplex-ssDNA equilibrium to the duplex form.

2.2.4 Current-Time Recordings

Current-time recordings were performed at 23.5 ± 1 °C using a custom-built high-impedance, low-noise amplifier and data acquisition system (Electronic Bio Sciences, San Diego, CA). The schematics of the experimental setup are shown in Figure 2.2. Electrolyte containing 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH = 7.4) was used to fill the experimental cell and the GNM capillary. The inside of the GNM capillary was connected to a pressure gauge and a 10 mL gastight syringe (Hamilton). A voltage was applied between two Ag/AgCl electrodes positioned inside and outside of the GNM capillary. Formation of a lipid bilayer was accomplished by depositing the DPhPC/decane solution across the GNM orifice; successful bilayer formation was indicated by a resistance increase from ~10 MΩ (corresponding to the open GNM) to ~100 GΩ.[43] A positive pressure of 20 to 40 mmHg was then applied to the GNM capillary, allowing the lipid bilayer to be functional for the protein channel reconstitution.[44-46] 0.2 μL of α-hemolysin solution (monomer, 1 mg/mL, prepared as described above) was added to the experimental cell (300 μL). After protein reconstitution into the lipid bilayer, the duplex DNA sample was added to the experimental cell at 5 μM. A voltage (-120 or -140 mV, *cis* vs. *trans*) was applied across the GNM orifice and was referenced to the Ag/AgCl electrode placed inside of the GNM capillary. A minimum of 500 duplex unzipping events were collected for each sample. The current-time traces were filtered at 100 kHz and sampled at 500 kHz.

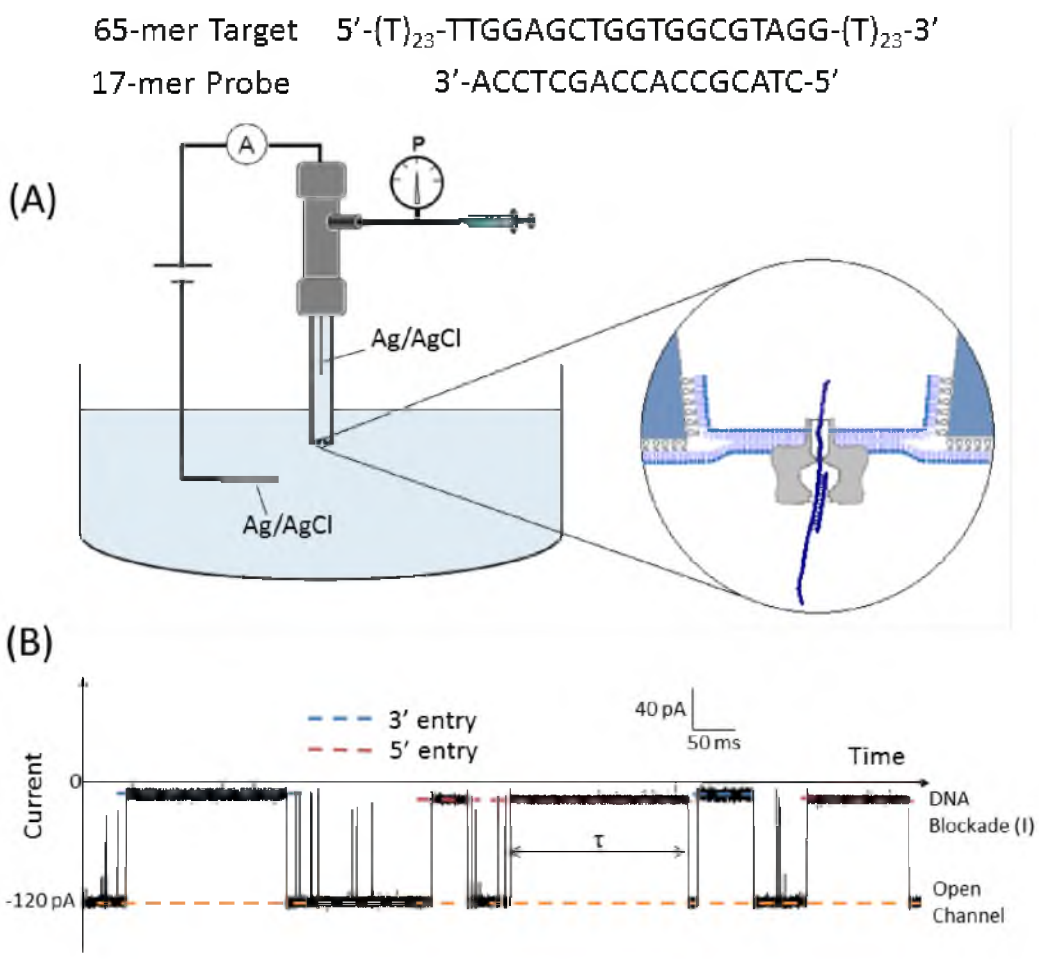


Figure 2.2. Nanopore ion-channel recordings. (A) Schematic illustration of the experimental setup. (B) Example current-time (*I-t*) trace obtained for the G:C duplex (sequence shown below) at -120 mV (*cis* vs. *trans*).

2.2.5 Data Analysis

Blockades that lasted longer than 0.5 ms and reduced the channel conductance to -40 to 0 pA were analyzed as DNA unzipping events, while the shorter blockades (< 0.5 ms) were identified as translocation events of unbound strands (see Appendix). Events were extracted using QuB (version 1.5.0.31). Histograms of unzipping durations were plotted using OriginPro (version 8.5.1). Entry direction was determined by results in Figures 2.3 and 2.4. Histograms corresponding to 5' or 3' entry were fit using eq 2.1 for G- and OG-containing duplexes or eq 2.2 for Gh- and Sp-containing duplexes to obtain the kinetic rate constants. Density plots were generated using data analysis programs provided by Electronic Bio Sciences, San Diego, CA.

2.3 Results and Discussion

The 65-mer target strand used for study has a heterosequence, specifically, the sequence surrounding codon 12 of the Kras gene in which the highlighted G is key to a cancer-related mutation.[47]

This sequence is embedded in a poly-dT background. After hybridization with a complementary 17-mer probe, a duplex is formed with a double-stranded segment placed between two poly-dT overhangs. The voltage-driven unzipping of single molecules of DNA duplexes was studied by pulling the overhang segment of the molecule into the α-HL channel in an electrical field. A bias of -120 mV (*cis* vs. *trans*) was applied across an α-HL channel that was reconstituted into a lipid bilayer suspended across the orifice of a glass nanopore membrane (GNM). The

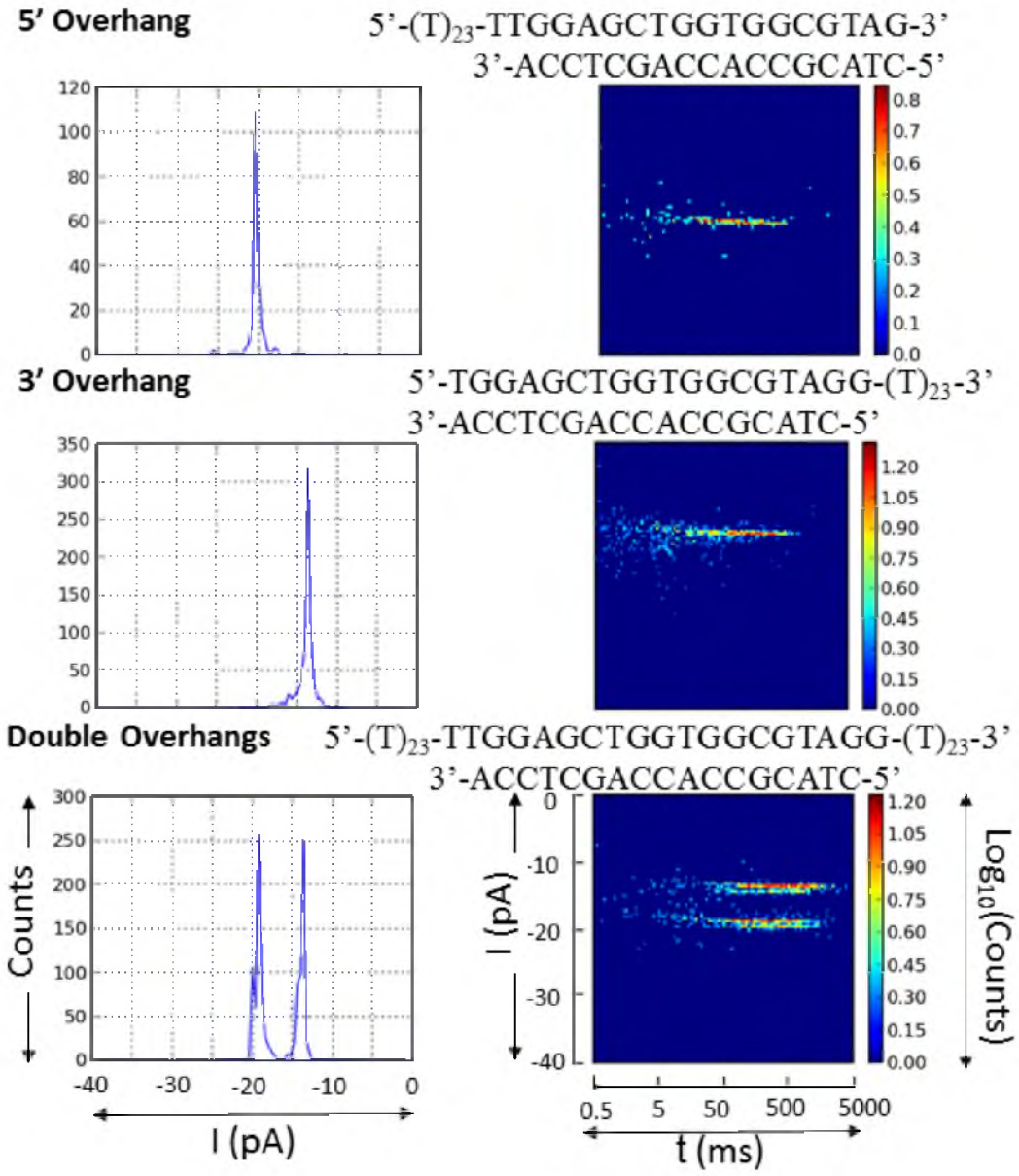


Figure 2.3. Histograms of blockage current (I) for the duplexes with either one overhang or two overhangs and plots of event population density for I as a function of unzipping time (t). For the double-overhang duplex with the sequence shown above, the more blocking population was due to 3' entry while the less blocking population was due to 5' entry.

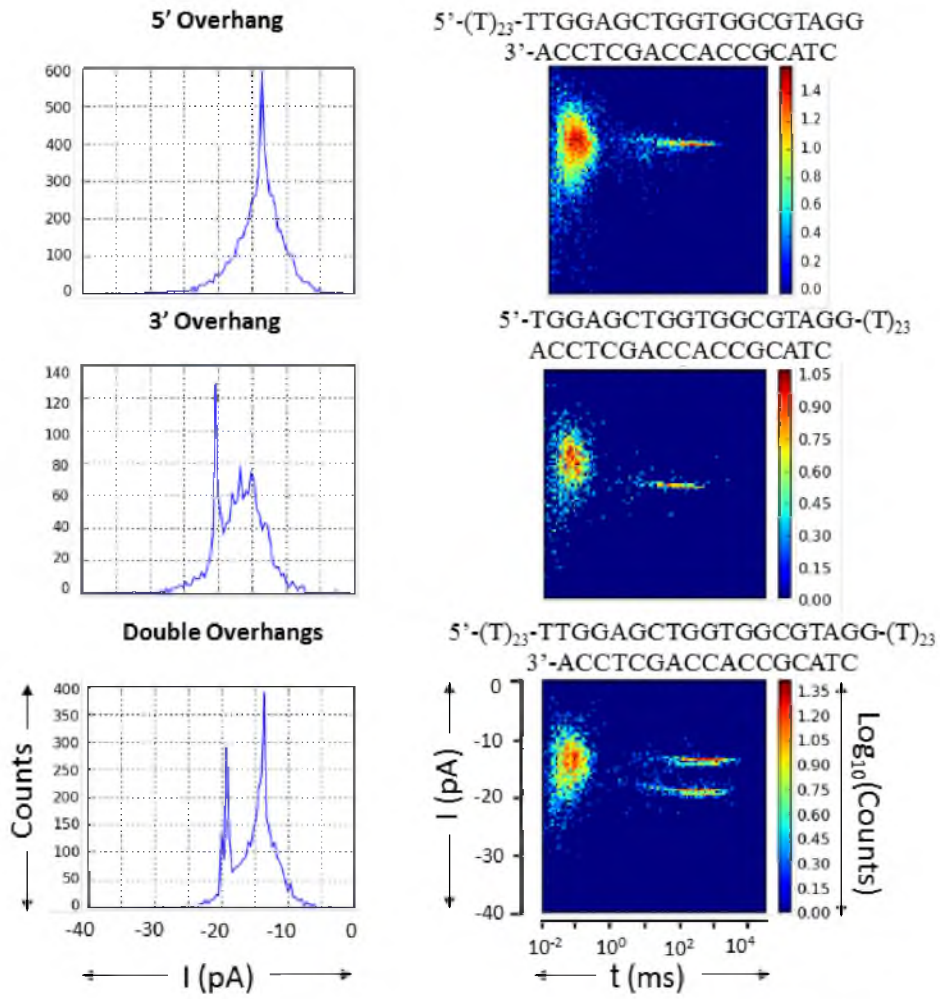


Figure 2.4. Histograms of blockage current (I) for the target-probe mixture with a single-sided overhang or double-sided overhangs, and event population density plots of I as a function of t. Due to the equilibrium between the complementary ssDNA and the duplex, solutions examined contain both ssDNAs and the duplex. In order to shift the equilibrium towards the duplex, a 5-fold excess of the short probe stands versus target strands was added to the solution. At -120 mV, the excess ssDNA probe strands generated a translocation population at times shorter than 0.5 ms, while the duplex underwent an unzipping process that was typically two or three orders of magnitude longer.

DNA duplex driven into the channel unzips into two separated single strands; the unzipping kinetic analysis was based on the unzipping time (*t)* of the duplex.

2.3.1 Directionality Studies on Duplex Unzipping

The unzipping process can be initiated from either the 3' or 5' end of the target, depending on which overhang enters into the channel first. (Note: 3' and 5' entry in this context refers to the termini of the 65-mer target strand.) We found that the two unzipping orientations have distinct current blockage levels and different unzipping durations.[48] In order to assign the two distinct current blockage levels to entry from the 3' or 5' direction, we initially performed measurements on unzipping of complementary duplexes that contain either one overhang or two overhangs in different experiments. The duplex segment remained the same for all of these duplexes, while the overhang was present either on the 3' end, or on the 5' end, or on both ends.

Figure 2.3 illustrates how the current blockage level and unzipping duration depend on 3' or 5' entry of the overhang. For the 3'-overhang duplex, a single stripe-like population is observed in the *I* vs. *t* plot. Similarly, unzipping of the 5'-overhang duplex displayed a single unzipping population in the *I* vs. *t* plot. The current blockage level of 5' unzipping (-20 pA) was shallower than that of 3' unzipping (-14 pA) and the duration of the former was slightly shorter than the later. In contrast, the double-overhang duplex generated two well-resolved *I* vs. *t* populations with different blockage currents (-19 pA and -14 pA), consistent with the blockade current level observed when employing either the 3' or the 5' single-

overhang duplex in separate experiments. We therefore assigned the more blocking population of the double-overhang duplex as being due to 3' entry and the less blocking population being due to 5' entry. The conclusion that 3' entry induces a deeper blockade than 5' entry for the double-overhang duplex is in agreement with the directionality study for ssDNAs with terminal hairpins, which also reported that 3'-threading ssDNA blocks the α-HL channel more than 5'-threading strands.[49,50] Interestingly, for poly-dT immobilized in the α-HL using a biotin-streptavidin terminus, a more blocking current is associated with 5' entry and a less blocking current associated with 3' entry.[51] In addition, essentially equal entry rates via 3' and 5' entry were observed for the double-overhang duplex. This is in contrast with reports for the immobilized homopolymer DNA and ssDNAs with terminal hairpins, which display a biased entry rate that depends on the sequence.

2.3.2 Unzipping Kinetics for Duplexes Containing Lesions

To investigate the unzipping kinetics of damaged duplexes, we chose the guanine in the middle of the 65-mer target strand (highlighted in red in the sequence below) as our point of interest for lesion insertion. A 17-mer probe was allowed to hybridize with a set of 65-mer target strands that differ by one nucleotide (X = G, OG, Gh, or Sp). The 17-mer probe has a cytosine (C) placed opposite to X. The resultant four duplexes were denoted as G:C, OG:C, Gh:C and Sp:C. Electrical measurements were performed on these four duplexes at -120 mV (*cis* vs. *trans*) to investigate the kinetics of unzipping undamaged and

damaged duplexes. Each duplex generated two unzipping populations with distinct current blockage levels; the same 3' versus 5' assignment previously determined was used to identify the entry direction. Histograms of unzipping duration for each orientation are plotted in Figure 2.5.

For the G:C and OG:C duplexes, histograms of unzipping duration for both 3' and 5' entry display a single exponential decay, indicating the unzipping process follows a first-order process (Figure 2.6, Type I). However, the peak shape of the histograms for Gh:C and Sp:C indicate that the unzipping of Gh:C and Sp:C follows a different kinetic model. The peak shape of the duration histogram is consistent with unzipping occurring by two sequential first-order reactions (Figure 2.6, Type II). Therefore, we propose a three-state model for the unzipping of Gh:C and Sp:C with the intermediate state corresponding to the duplex unzipped up to the damaged spot X (X = Gh or Sp). Branton and coworkers reported that unzipping of mismatched duplexes in α-HL also generates a multistep model and suggested the intermediate occurs when the duplex unzips to the mismatched site.[3] The presence of an intermediate state was also found when two mismatched DNA strands were forced apart by the optical force clamp.[1]

Additionally, we discovered that whether the insertion of a lesion into the duplex results in the Type I model or Type II model is strongly correlated to the extent to which the lesion destabilizes the duplex. OG is able to base pair with C, though the duplex is slightly destabilized by repulsive interactions introduced by the 8-oxo group.[52] As shown in Table 2.1, this small destabilizing effect of OG caused a T_m decrease by ~1 °C relative to the G:C duplex and a 3 to 4 fold

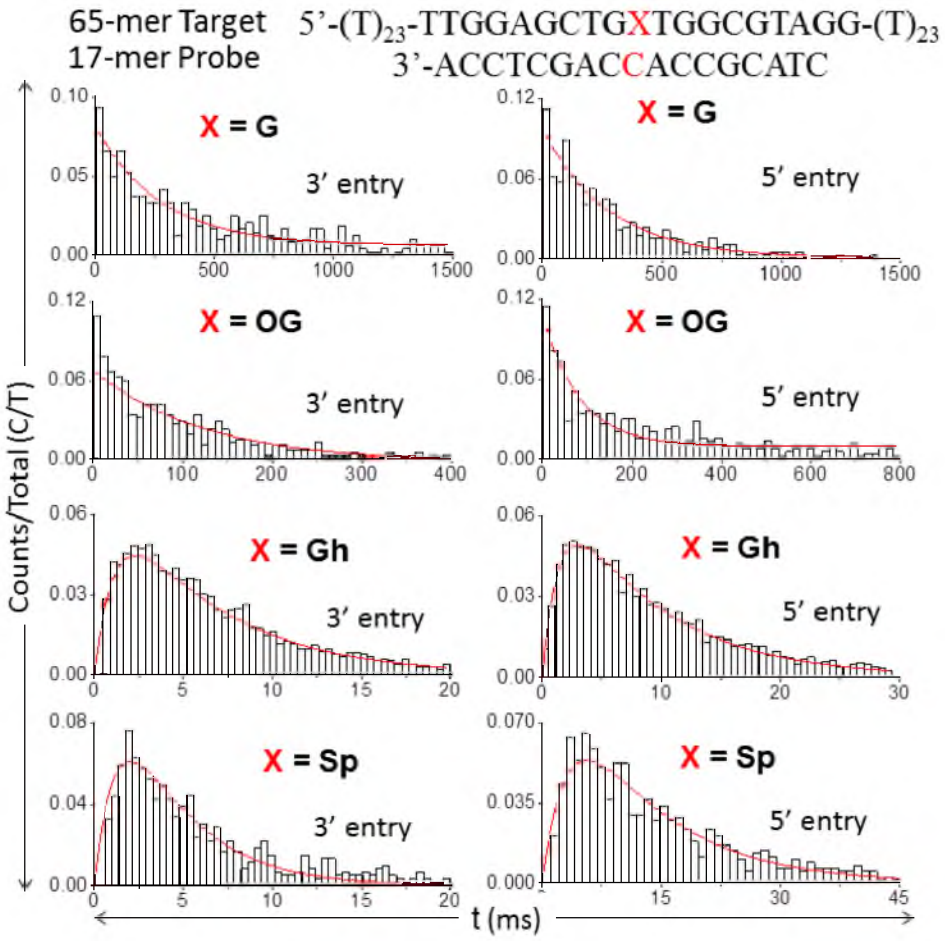


Figure 2.5. Histograms of unzipping duration (*t*) for 3' and 5' entry at -120 mV (*cis* vs. *trans*) for the duplex formed by the 65-mer target (where X = G, OG, Gh or Sp) and the 17-mer probe. The rate constants for unzipping of each damaged duplex were obtained based on the fit (red curve) to the histograms using the Type I (X = G, OG) or Type II (X = Gh, Sp) model.

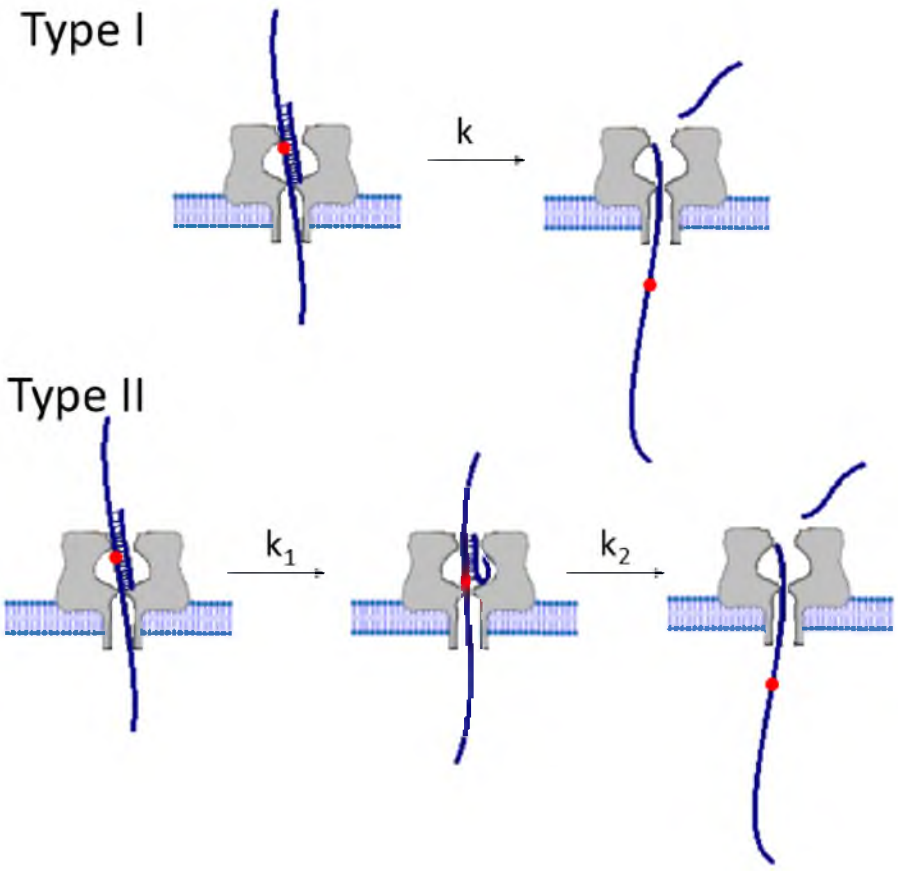


Figure 2.6. Duplex unzipping models. Unzipping of G:C and the slightly destabilized OG:C follows the Type I model, while unzipping of the highly destabilized Gh:C and Sp:C follows the Type II model. The red spot in the strand indicates the lesion site.

Table 2.1. The kinetic models and corresponding rate constants for unzipping the 65-mer:17-mer duplex at -120 mV.

5'-$(T)_{23}$-TTGGAGCTGXTGGCGTAGG-$(T)_{23}$-3'

3'-ACCTCGACCACCGCATC-5'

Sequence	T_m (°C)	Unzipping Model	Rate Constants (s^{-1})	
			3' entry	5' entry
X = G	70.3 ± 0.4	Type I	3.1 ± 0.3	3.5 ± 0.2
X = OG	69.3 ± 0.5	Type I	13 ± 1	9 ± 1
X = Gh	55.9 ± 0.5	Type II [a]	890 ± 50, 180 ± 10	720 ± 50, 110 ± 10
X = Sp	57.1 ± 0.5	Type II [a]	730 ± 60, 290 ± 20	320 ± 50, 90 ± 10

[a] For the Type II model, the two rate constants are listed without assignment to the first or second step.

increase in unzipping rate. The small decrease in stability is apparently insufficient to generate a pronounced minimum in the unzipping energy profile.[53] On the other hand, Gh and Sp cannot form stable hydrogen bonds with C, which significantly decreases the duplex stability. [54,39] The presence of Gh or Sp caused the T_m to decrease by ~13 °C and the unzipping rate to increase by one to two orders of magnitude relative to G:C. Since Gh and Sp are highly destabilizing lesions, we propose that they are able to produce an energy well in the unzipping energy landscape that is deep enough to generate an intermediate state.

We assume the unzipping of G:C and OG:C to follow the Type I model, and therefore, fit their *t* histograms using the kinetic equation for a first-order reaction,

$$C/T = \text{ce}^{-t}\Delta t \tag{2.1}$$

where *C/T* is event counts in a time increment Δt centered at time *t* divided by the total counts (see SI), *k* is the rate constant for the unzipping process, and *t* is the unzipping duration.

For Gh:C and Sp:C, we assume their unzipping to follow the Type II model. Their *t* histograms were fit using the kinetic equation for two sequential first-order reactions,

$$C/T = \frac{k_1 k_2}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t})\Delta \tag{2.2}$$

where k_1 and k_2 correspond to the rate constants for the first and second unzipping steps (see SI).

Table 2.1 lists the rate constants for the unzipping of G:C, OG:C, Gh:C and Sp:C for 3' and 5' entry. For Gh:C and Sp:C, two values of rate constants were obtained from the Type II fit. However, these two values are interchangeable in eq 2.2.

The centrally placed X:C in the 65-mer:17-mer duplex was located between two duplex subsections, each comprised of eight base pairs. In the Type II model, the unzipping process involves opening of the first 8-bp subsection that is close to the entry side, followed by opening of the second 8-bp subsection on the other side of X:C. The subsection on either side of X has the same number of hydrogen bonds, and the thermal stability for each of the two subsections based on the T_m was predicted to be the same by an mfold model. The free energy of the 8 bp-subsection on the 5' side of X, calculated using NuPack, is slightly higher than that on the 3' side ($\Delta G_{5'\text{-side}}$ = -13.48 kcal/mol, $\Delta G_{3'\text{-side}}$ = -13.71 kcal/mol at 23.5 ^{o}C). However, it is not the overall free energy but the activation energy of strand separation that directly relates to unzipping kinetics;[55] thus, the determination of k_1 and k_2 cannot be made based on the predicted stability of two subsections. Additional experiments were required to assign the values of the rate constants to each individual step of the unzipping mechanism.

To make the duplex subsections distinctly asymmetrical in terms of stability, we designed two 13-mer probes, each being truncated by four bases either from the 3' end or from the 5' end. These are denoted as 13-mer 3'-truncated probe or

13-mer 5'-truncated probe. (Note: the 3' and 5' designations here refer to the probe.) Unzipping experiments were carried out for truncated duplexes when X = Gh at -140 mV and the Type II model was applied to obtain two rate constants. A potential of -140 mV, *cis* vs. *trans*, was used here to increase the event frequency for the convenience of data collection. Because opening of the 8-bp subsection on one side of Gh should take considerably longer than opening of the 4-bp subsection on the other side, the smaller rate constant obtained from the fit was assigned to the dissociation step of the 8-bp subsection. For example, as entry of the 13-mer 5'-truncated probe into the channel from the 5' terminus, two rate constants, 790 ± 70 s^{-1} and 1000 ± 170 s^{-1}, were obtained from the Type II fit using eq 2.2. The rate constant of 790 ± 70 s^{-1} should be k_1, associated with opening of the longer 8-bp subsection close to 5' terminus (Table 2.2, blue). Assuming it takes approximately the same time to open the 8-bp subsection in the 65-mer:17-mer duplex as in the 65-mer:13-mer duplex via 5' entry (both highlighted as blue in Table 2.2), k_1 for the 65-mer:17-mer duplex via 5'-unzipping should be very close to 790 ± 70 s^{-1}. Out of the two k values (790 ± 40 s^{-1} and 280 ± 20 s^{-1}) obtained from the 5'- unzipping of the 65-mer:17-mer duplex, 790 ± 40 s^{-1} was therefore assigned as k_1. The other k value (280 ± 20 s^{-1}) was assigned to be k_2 because it is very similar to the k_2 for the same 8-bp subsection (Table 2.2, pink) of the duplex containing 13-mer 3'-truncated probe at 5' entry, 190 ± 20 s^{-1}. This method to resolve k_1 and k_2 was successfully applied to determine the rate constants at 3' entry as well. The assignments of k_1 and k_2 for unzipping of the 65-mer:17-mer duplex are listed in Table 2.2 along with their

Table 2.2. Unzippping rate constants (*k*) at -140mV for the Gh duplexes formed with 17-mer probe, 13-mer 5'-truncated probe, or 13-mer 3'-truncated probe[a].

17-mer Probe

5'-$(T)_{23}$-TTGGAGCTGGhTGGCGTAGG-$(T)_{23}$-3'
3'-ACCTCGAC C ACCGCATC-5'

13-mer 5'-Truncated Probe

5'-$(T)_{23}$-TTGGAGCTGGhTGGCGTAGG-$(T)_{23}$-3'
3'-ACCTCGAC C ACCG-5'

13-mer 3'-Truncated Probe

5'-$(T)_{23}$-TTGGAGCTGGhTGGCGTAGG-$(T)_{23}$-3'
3'-CGAC C ACCGCATC-5'

Probe	Rate Constants (s^{-1})			
	3' entry		5' entry	
	k_1	k_2	k_1	k_2
17-mer Probe	550 ± 30	930 ± 50	790± 40	280 ± 30
13-mer 5'-Truncated Probe	1690 ± 130	880 ± 50	790 ± 70	1000± 170
13-mer 3'-Truncated Probe	490 ± 30	3270 ± 440	1603 ± 200	190 ± 20

[a]The 8-bp duplex subsections of the same sequence and their corresponding *k* are highlighted using the same colors (blue or pink). The *k* values for unzipping of subsections are assumed to be the same regardless if the subsection is in the 65-mer:17-mer duplex or the truncated duplex with the 13-mer.

reference rate constants obtained from the truncated probes. The histograms for duplexes in Table 2.2 are shown in Figure 2.7.

Regardless of the entry direction, it takes less time to open the duplex subsection on the 5' side of Gh:C than the 3' side. This may imply that the destabilizing effect imparted by Gh influences the 3' and 5' sides of the lesion differently. While the local stability around a Gh lesion has not been previously reported, molecular dynamics simulation suggests that there exists a stability variation on the 3' and 5' sides of another hydantoin lesion, Sp.[56]

In addition, which orientation of unzipping is faster appears to be sequence-dependent. For the duplex containing Gh, the *k* for 5' entry is always smaller than that for 3' entry regardless if it is in the truncated or original duplex. The trend of 5' unzipping slower than 3' unzipping was observed for OG:C, Gh:C and Sp:C but not for G:C. This implies that the DNA conformation of OG:C, Gh:C and Sp:C for 5'-threading is more unstable to unzip in the nanopore, while for the matched G:C duplex, the 3'-threading conformation instead facilitates the strand dissociation.

The rate constants k_1 and k_2 do not appear to be strongly correlated. Presumably, if the unzipping process takes place in the α-HL vestibule, the shorter subsection of the 13-mer truncated probe being dissociated first would impose smaller steric hindrance to the opening of the second subsection relative to the 17-mer probe. Therefore, a larger k_1 for opening of the shorter subsection would result in a larger k_2 for opening of the second subsection. Because our data for rate constants have an error of 5 to 17%, we are not able to tell if such a

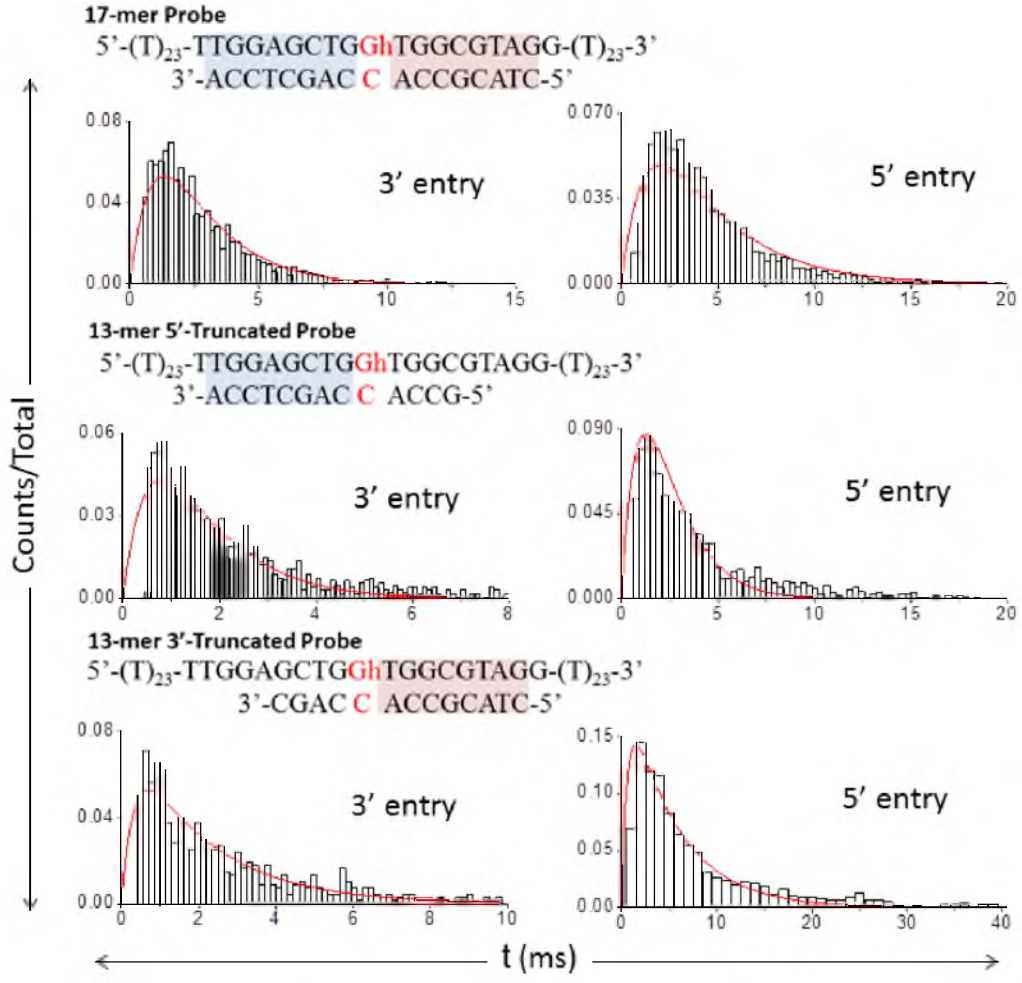


Figure 2.7. Histograms of unzipping time (t) for 3' and 5' entry at -140 mV for Gh:C duplexes formed with (top) the 17-mer probe; (middle) the 13-mer 5'-truncated probe; and (bottom) the 13-mer 3'-truncated probe. The red curves show the fits using the Type II model, eq 2.4. k_1 and k_2 values obtained from the fits are listed in Table 2.2 of the main text. Blockades that lasted longer than 0.5 ms were analyzed as DNA unzipping events and are plotted below, while the shorter blockades (< 0.5 ms) were identified as translocation events of unbound strands.

correlation exists. On the other hand, if the unzipping process takes place outside of the vestibule, this correlation between k_1 and k_2 originating from steric hindrance would be expected to be much less.

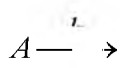
2.4 Conclusions

Our nanopore-based kinetic study has demonstrated that unzipping of duplexes that contain single sites of an oxidized guanine lesion (OG, Gh or Sp) via 3' and 5' entry in the α-HL follows a first-order reaction path or a model of two sequential first-order reactions. Which model should be applied depends on the extent of the destabilizing effect imparted by the lesion. Our work has highlighted the ability of the nanopore to be used as a powerful tool to study the force-induced kinetics on the single-molecule level. In addition to investigating the kinetic model describing unzipping of the damaged duplex, this work has also shown the nanopore-based method as a very useful approach to study DNA local dynamics and interaction at different orientations, which are often hidden by the global behaviors of the DNA molecules.

2.5 Appendix

2.5.1 Kinetic Equations for the Type I Model

For a first-order reaction with a rate constant k,



the reaction rates for *A* and *B* are

$$\frac{d[A]}{dt} = \quad [A]$$
$$\frac{d[B]}{dt} = \quad [A]$$

The solutions for the above two equations are

$$[A] = \quad A_0]e^{\quad t}$$
$$[B] = \quad A_0](1 - \quad ^{\quad t})$$

where $[A_0]$ is the initial concentration of *A*. Taking the derivative of [*B*] and multiplying both sides by the time increment, Δt, yields

$$\frac{\frac{d[B]}{dt} \times \quad}{[A_0]} = \quad e^{\quad t}\Delta \qquad (2.3)$$

The left side of eq 2.3 is equal to the percentage of *B* generated in an increment Δt at time *t*, which corresponds to the counts/total (*C/T*) in eq 2.1 and Figure 2.3.

2.5.2 Kinetic Equations for the Type II Model

For two sequential first-order reactions with rate constants k_1 and k_2

$$A \xrightarrow{k_1} \xrightarrow{k_2}$$

the reaction rates for *A* and *B* are

$$\frac{d[A]}{dt} = -k_1[A]$$
$$\frac{d[B]}{dt} = k_1[A] - k_2[B]$$
$$\frac{d[C]}{dt} = k_2[B]$$

The solutions are

$$[A] = [A_0]e^{-k_1 t}$$
$$[B] = [A_0]\frac{k_1}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t})$$
$$[C] = [A_0]\left(1 + \frac{k_1 e^{-k_2 t} - k_2 e^{-k_1 t}}{k_2 - k_1}\right)$$

where $[A_0]$ is the initial concentration of *A*. Taking the derivative of [*C*] and multiplying both sides by Δt yields

$$\frac{\frac{d[C]}{dt} \times \Delta t}{[A_0]} = \frac{k_1 k_2}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t}) \times \Delta t \qquad (2.4)$$

The left side of eq 2.4 is equal to the percentage of *C* generated in an increment Δt at time *t*, and corresponds to the counts/total (*C/T*) in eq 2.2.

2.6 References

1. Woodside, M. T.; Anthony, P. C.; Behnke-Parks, W. M.; Larizadeh, K.; Herschlag, D.; Block, S. M. *Science* **2006**, *314*, 1001–1004.

2. Cocco, S.; Monasson, R.; Marko, J. F. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8608-8613.

3. Sauer-Budge, A. F.; Nyamwanda, J. A.; Lubensky, D. K.; Branton, D. *Phys. Rev. Lett.* **2003**, *90*, 238101.

4. Muzard, J.; Martinho, M.; Mathé, J.; Bockelmann, U.; Viasnoff, V. *Biophys. J.* **2010**, *98*, 2170-2178.

5. Howorka, S.; Cheley, S.; Bayley, H. *Nat. Biotechnol.* **2001**, *19*, 636.

6. Sutherland, T. C.; Dinsmore, M. J.; Kraatz, H.-B.; Lee, J. S. Biochem. *Cell Biol.* **2004**, *82*, 407-412.

7. Viasnoff, V.; Chiaruttini, N.; Bockelmann, U. *Eur. Biophys. J.* **2009**, *38*, 263-269.

8. Renner, S.; Geltinger, S.; Simmel, F. C. *Small* **2010**, *6*, 190–4.

9. Jetha, N. N.; Wiggin, M.; Marziali, A. *Methods Mol. Biol.* **2009**, *544*, 129.

10. Viasnoff, V.; Chiaruttini, N.; Muzard, J.; Bockelmann, U. *J. Phys.: Condens. Matter.* **2010**, *22*, 454122.

11. Dudko, O. K.; Mathé, J.; Meller, A. *Methods Enzymol.* 2010, 475, 565.

12. Kasianowicz, J. J.; Robertson, J. W. F.; Chan, E. R.; Reiner, J. E.; Stanford, V. M. Annu. Rev. Anal. Chem. 2008 1, 737.

13. Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouax, J. E. *Science* **1996**, *274*, 1859–1865.

14. Akeson, M.; Branton, D.; Kasianowicz, J.; Brandin, E; Deamer, D. *Biophys. J.* **1999**, *77*, 3227–3233.

15. Kasianowicz, J.; Brandin, E.; Branton, D.; Deamer, D. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13770–13773.

16. Rotem, D.; Jayasinghe, L.; Salichou, M.; Bayley, H. *J. Am. Chem.* Soc. **2012**, *134*, 2781.

17. Kawano, R.; Osaki, T.; Sasaki, H.; Takinoue, M.; Yoshizawa, S.; Takeuchi, S. *J. Am. Chem. Soc.* **2011**, *133*, 8474.

18. Jetha, N. N.; Feehan, C.; Wiggin, M.; Tabard-Cossa, V.; Marziali, A. *Biophys. J.* **2011**, *100*, 2974.

19. Comer, J.; Dimitrov, V.; Zhao, Q.; Timp, G.; Aksimentiev, A. *Biophys J.* **2009**, *96*, 593.

20. Gyafas, B.; Abu-Shumays, R.; Wang, H.; Dunbar, W. B. *Biophys. J.* **2011**, *100*, 1509.

21. Winters-Hilt, S.; Morales, E.; Amin, I.; Stoyanov, A. *BMC Bioinformatics* **2007**, *8*, S20.

22. Ying, Y. L.; Wang, H. Y.; Sutherland, T. C.; Long, Y. T. *Small* **2011**, *7*, 87.

23. DeGuzman, V. S.; Lee, C. C.; Deamer, D. W.; Vercoutere, W. A. *Nucleic Acids Res.* **2006**, *34*, 6425.

24. Shim, J. W.; Tan, Q.; Gu, L.-Q. *Nucleic Acides Res.* **2009**, *37*, 972.

25. Mathé, J.; Visram, H.; Viasnoff, V.; Rabin, Y.; Meller, A. *Biophy. J.* **2004**, *87*, 3205-3212.

26. Dudko, O. K.; Mathé, J.; Szabo, A.; Meller, A.; Hummer, G. *Biophys. J.* **2007**, *92*, 4188-4195.

27. Renner, S.; Bessonov, A.; Gerland, U.; Simmel, F. C. *J. Phys.: Condens. Matter.* **2010**, *22*, 454119.

28. McNally, B.; Wanunu, M.; Meller, A. *Nano. Lett.* **2008**, *8*, 3418-3422.

29. Hornblower, B.; Coombs, A.; Whitaker, R. D.; Kolomeiskym, A.; Picone, S. J.; Meller, A.; Akeson, M. *Nature Methods* **2007**, *4*, 315-317.

30. Cockroft, S. L.; Chu. J.; Amorin, M.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2008**, *130*, 818-20.

31. Zhao, Q.; Comer, J.; Dimitrov, V.; Yemenicioglu, S.; Aksimentiev, A.; Timp, G. *Nucleic Acids Res.* **2008**, *36,* 1532–1541

32. Zhao, Q.; Sigalov, G.; Dimitrov, V.; Dorvel, B.; Mirsaidov, U.; Sligar, S.; Aksimentiev, A.; Timp, G. *Nano Lett.* **2007**, *7*, 1680-1685.

33. Luo, W.; Muller, J. G.; Rachlin, E.; Burrows, C. J. *J. Org. Lett.* **2000**, *2*, 613.

34. Ye, Y.; Muller, J. G.; Luo, W.; Mayne, C. L.; Shallop, A. J.; Jones, R. A.; Burrows, C. J. *J. Am. Chem. Soc.* **2003**, *125*, 13926-13927.

35. Schibel, A. E. P.; An. N.; Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. *J. Am. Chem. Soc.* **2010**, *132*, 17992.

36. Mangal, D.; Vudathala, D.; Park, J.-H.; Lee, S. H.; Penning, T. M.; Blair, I. A. *Chem. Res. Toxicol.* **2009**, *22* , 788–797.

37. Lonkar, P.; Dedon, P. C. *Int. J. Cancer* **2011**, *128*, 1999-2009.

38. Mangialasche, F.; Polidori, M. C.; Monastero, R.; Ercolani, S.; Camarda, C.; Cecchetti, R.; Mecocci, P. *Ageing Res. Rev.* **2009**, *8*, 285-305.

39. Kornyushyna, O.; Berges, A. M.; Muller, J. G.; Burrows, C. J. *Biochemistry* **2002**, *41*, 15304-16314.

40. Zhang, B.; Galusha, J.; Shiozawa, P. G.; Wang, G.; Bergren, A. J.; Jones, R. M.; White, R. J.; Ervin, E. N.; Cauley, C.; White, H. S. *Anal. Chem.* **2007**, *79*, 4778–4787.

41. Lan, W.-J.; Holden, D. A.; Zhang, B.; White, H. S. *Anal. Chem.* **2011**, *83*, 3840.

42. Lan, W.-J.; White, H. S. *ACS Nano* **2012**, *6*, 1757.

43. White, R. J.; Ervin, E. N.; Yang, T.; Chen, X.; Daniel, S.; Cremer, P. S.; White, H. S. *J. Am. Chem. Soc.* **2007**, *129*, 11766-11775.

44. Schibel, A. E. P.; Heider, E. C.; Harris, J. M.; White, H. S. *J. Am. Chem. Soc.* **2011**, *133*, 7810–7815.

45. Lan, W.-J.; Holden, D. A.; Liu, J.; White, H. S. *J. Phys. Chem. C.* **2011**, *115*, 18445.

46. Lan, W.-J.; Holden, D. A.; White, H. S. *J. Am. Chem. Soc.* **2011**, *133*, 13300.

47. Pfeifer, G. P.; Besaratinia, Hum. Genet. 2009, 125, 493–506.

48. Schibel, A. E. P.; Fleming, A. M.; Jin, Q.; An, N.; Liu, J.; Blakemore, C. P.; White, H. S.; Burrows, C. J. *J. Am. Chem. Soc.* **2011**, *133*, 14778-14784.

49. Butler, T. Z.; Gundlach, J. H.; Troll, M. *Biophys. J.* 2007, 93, 3229.

50. Mathé, J.; Aksimentiev, A.; Nelson, D. R.; Schulten, K.; Meller, A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 12377–12382.

51. Purnell, R. F.; Mehta, K. K.; Schmidt, J. J. *Nano Lett.* **2008**, *8*, 3029-3034.

52. Lipscomb, L. A.; Peek, M. E.; Morningstar, M. L.; Verghis, S. M.; Miller, E. M.; Rich, A.; Essigmann, J. M.; Williams, L. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92* , 719–723.

53. Bockelmann, U.; Viasnoff, V. *Biophys. J.* **2008**, *94*, 2716-2724.

54. Chinyengetere, F.; Jamieson, E. R. *Biochemistry* ***2008****, 47*, 2584-2591.

55. Dudko, O. K.; Hummer, G.; Szabo, A. *Phys. Rev. Lett.* **2006**, 96, 108101.

56. Jia, L.; Shafirovich, V.; Shapiro, R.; Geacintov, N. E. Broyde, S. *Biochemistry* **2005**, *44*, 13342-13353.

CHAPTER 3

FINE-TUNING THE KINETIC PROPERTIES OF dsDNA STRAND DISSOCIATION IN A NANOPORE USING MODIFIED BASE PAIRS

3.1 Introduction

In single-molecule force experiments, the time-dependent trajectory describing the response of a biomolecule to an external force can be used to extract kinetic information concerning structures and interactions of individual molecules.[1,2] Kinetic studies using single-molecule force techniques have covered a variety of biological processes, including intramolecular and intermolecular interactions of DNA, RNA, and proteins.[3-6] Among them, the unzipping process of duplex DNA has been extensively investigated for the purpose of understanding the enzymatically meditated process of strand separation that universally occurs during DNA replication and transcription.[7-9] As an emerging technique of single-molecule force experiment, the nanopore method provides a fast and linker-free approach for molecular manipulation that overcomes the drawbacks of optical tweezers and atomic force microscopy (AFM).[10-12] In a nanopore experiment using the α-hemolysin (α-HL) nanopore,

individual DNA duplexes are driven electrophoretically into the pore, with only the overhang part, if any, being able to fully thread through the nanopore.[13] In this circumstance, the electrical field imposes a localized rupture force on the duplex, causing strand dissociation to occur.[14,15] From the moment of entry to that of rupture, the duplex generates a temporal blockade to the ionic current flowing through the channel. The duration of this current blockage is used to extract kinetic information from the process of strand dissociation.[16-18] Typically, the unzipping duration of a double-stranded DNA is 2 - 3 orders of magnitude longer than the translocation time of a single-stranded DNA (ssDNA) under 120 mV applied voltage across the α-HL.[19,20]

The duplex unzipping, as well as other forms of obstacle removal such as hairpin unfolding and protein-DNA detachment, can be described by a process of overcoming an energy barrier.[21,22] The existence of a pronounced energy well leads to a multistep mechanism for the force-induced separation, which was reported in the studies of unraveling RNA pseudoknots, stripping two binding proteins off one RNA strand, and unzipping duplexes that include mismatch base pairs.[23-25] Previous work from our laboratories in Chapter 2 discovered that the force-induced strand dissociation of lesion-containing duplexes follows either a first-order reaction (Type I) or two sequential first-order reactions (Type II), depending on the extent of destabilizing effect imparted by the damage.[26]

In this work, the fine-tuning of duplex kinetics is achieved by tweaking the degree of destabilization caused by the modified base pair X:Y to the otherwise fully matched dsDNA. Twelve X:Y base pairs were chosen to study in this work,

generating a broad range of gradual variation in the extent of destabilization. In the X:Y base pair, 8-oxo-7,8-dihydroguanine (OG), an oxidized lesion of guanine (G), or one of its second-step oxidized products, guanidinohydantoin (Gh) or spiroiminodihydantoin (Sp) was placed at X to introduce destabilization of different levels, while cytosine (C), adenine (A), or 2,6-diaminopurine (D) was paired with the lesion to further adjust the extent of destabilization.[27-30] D, a base analog of A, increases the H-bonding in X:Y compared with A pairing, thus being used as a base pair modulator. Figure 3.1 demonstrates the base-pairing schemes for the modified base pairs X:Y. These X:Y pairs cover a variety of base-pairing structures, allowing us to understand which structural factors determine the kinetic stability of duplex.

3.2 Experimental Section

3.2.1 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. Each oligodeoxynucleotide was cleaved from the synthetic column after synthesis and then deprotected according to the manufacturer's protocols. The oligonucleotides were purified using an ion-exchange HPLC column with a linear gradient of 25% to 100% B over 30 min while monitoring UV/Vis absorbance at 260 nm (A = 10% CH_3CN/90% ddH_2O, B = 20 mM NaP_i, 1 M NaCl pH 7 in 10% CH_3CN/90% ddH_2O, flow rate = 1 mL/min).

65-mer: 5'-$(T)_{23}$-TTGGAGCTGXTGGCGTAGG-$(T)_{23}$

17-mer: 3'-ACCTCGACYACCGCATC

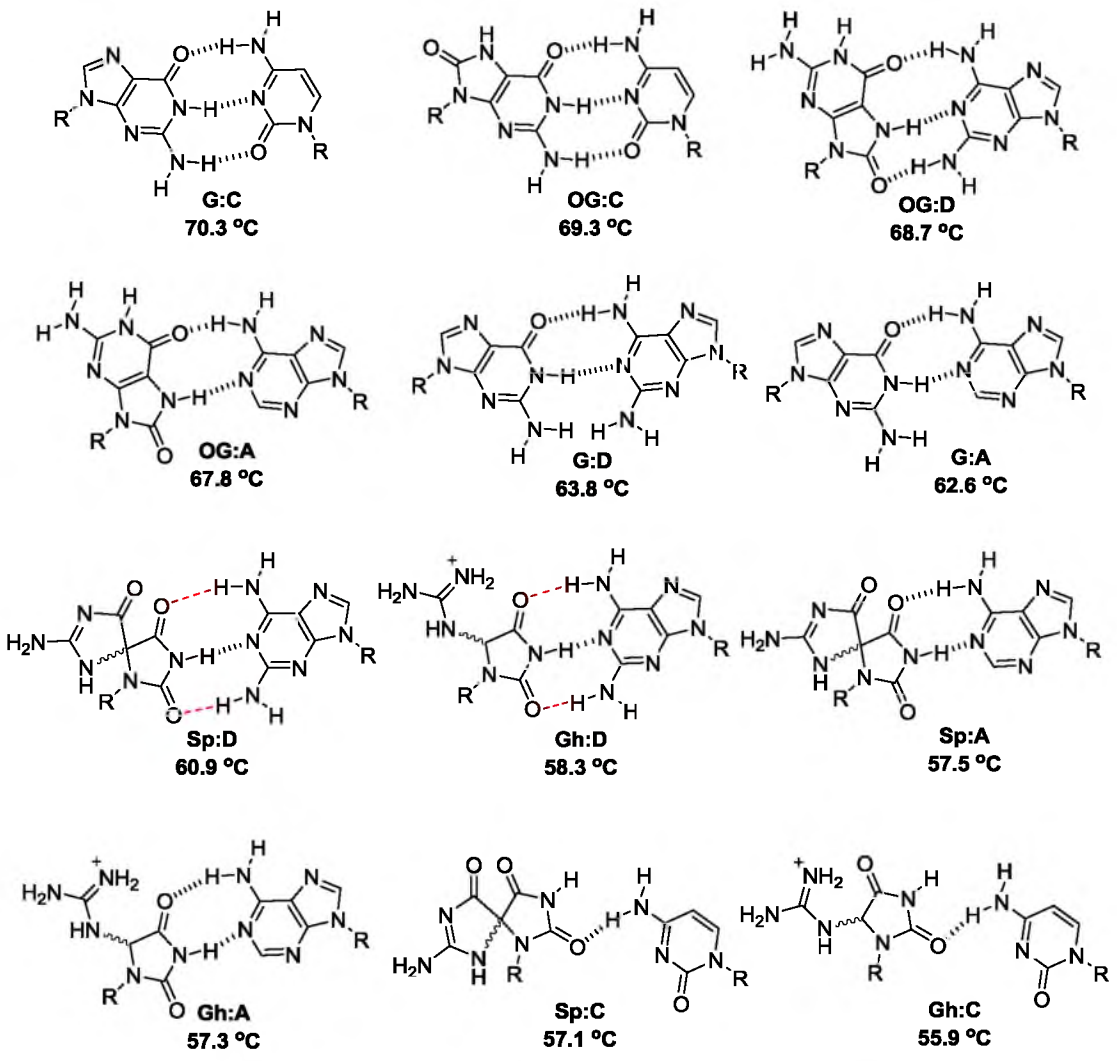


Figure 3.1. Base pairing schemes for X:Y (X = G, OG, Sp or Gh; Y = C, A, or D). The G and OG base pairs with A and C are drawn based on the reported structural data, while the hydantoin base pairs are drawn based on reported MD simulation data.[31-33] The D-containing base pairs shown are predicted based on their H-bonding capabilities discussed in refs. 34 and 35. The X:Y pairs are ordered in decreasing melting temperatures (from top left to bottom right). The melting temperatures listed under individual schemes change by a small degree for each X:Y-containing duplex and have an average standard deviation of 0.6 °C.

3.2.2 Thermal Denaturation Studies

In order to determine the T_m, thermal denaturation experiments were conducted with the truncated 23-mer strand, 5'- TT TTG GAG CTG XTG GCG TAG GTT-3', in which X = G, OG, Sp, or Gh. By removing the poly-dT tails of the 65-mer stand, the hyperchromic shift for the transition from dsDNA to ssDNA was more clearly observed. The dsDNA oligomers were formed by mixing the 23-mer and 17-mer (3'-ACC TCG ACY ACC GCA TC-5', Y = C, A or D) in a 1:1 ratio in 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH 7.4) at a concentration of 10 μM. The dsDNAs were then cooled to room temperature over 3 h. Thermal denaturation experiments were then conducted on diluted dsDNA samples at 1 μM in the same solution. Samples were thermally equilibrated at 20 °C for 20 min followed by heating to 75 °C at a rate of 0.5 °C/min. As the samples were heated, UV/vis absorbance readings were recorded at 260 nm every 30 sec. The T_m was determined using a two-point average analysis.

3.2.3 Chemicals and Materials for Nanopore Measurement

Wild-type α-hemolysin was purchased from List Biological Laboratories in a powder form of protein monomers. It was dissolved in water at 1 mg/mL and stored at -80 °C. The lipid (1,2-diphytanoyl-*sn*-glycero-3-phosphocholine, powder, Avanti Polar Lipids) was dissolved in decane at 10 mg/mL before use. Glass nanopore membranes (GNMs) fabricated using a bench-top method were chemically modified with 2% (v/v) (3-cyanopropyl) dimethylchlorosilane in acetonitrile to produce a hydrophobic surface as the support structure for the lipid

bilayer.[36-38] The 65-mer strands (5'-T_{23}-TTG GAG CTG XTG GCG TAG G-T_{23}-3', X = G, OG, Sp, or Gh) were mixed with the 17-mer strands (3'-ACC TCG ACY ACC GCA TC-5', Y = C, A or D) at 1:5 mole ratio (65-mer versus 17-mer) to form duplexes in 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH 7.4). The 65-mer and 17-mer mixture was heated in a 90 °C water bath for 5 min and then slowly cooled to room temperature over 3 h.

3.2.4 Current-Time Recordings

All current-time recordings were carried out on a custom built amplifier and data acquisition system (Electronic Bio Sciences, San Diego, CA). The same solution to dissolve DNA (1 M KCl, 10 mM PBS, and 1 mM EDTA at pH 7.4) was used as the electrolyte for ion-channel. -120 mV (*cis* vs. *trans*) was applied across the orifice of the GNM between two Ag/AgCl electrodes placed inside and outside of the GNM capillary. The electrode inside the capillary was grounded. A lipid bilayer was formed by painting the lipid/decane solution across the GNM orifice. A positive pressure was applied to inside of the GNM capillary, allowing the lipid bilayer to be functional for protein channel reconstitution.[39,40] Protein monomers were added to the *cis* side of the nanopore and self-assembled into the lipid bilayer as a heptamer. After protein insertion, the duplex DNA sample was added to the electrolyte reservoir at a final concentration of 5 μM. A minimum of 300 single-molecule unzipping events were collected for each DNA sample. Data were filtered at 100 kHz and sampled at 500 kHz. All samples were analyzed at least twice using ion-channel recordings.

3.2.5 Data Analysis

The blockades that lasted longer than 0.5 ms and reduced the channel conductance to -40 to 0 pA at -120 mV were analyzed as duplex unzipping events. The events shorter than 0.5 ms represent the translocation of the excess 17-mer strands due to the 1:5 mixing ratio (65-mer versus 17-mer) and were not analyzed. The unzipping process of a duplex that contains a single lesion site follows a path of either (I) a first-order reaction or (II) two sequential first-order reactions. The shape of histograms of unzipping duration (t) was used to determine which model the duplex follows during unzipping. The unzipping rate constant (k) was obtained based on fit of the t histogram using the corresponding kinetic equation for each model.

The kinetic equation for the Type I model is

$$C/T = \tau e^{-t}\Delta \tag{3.1}$$

where C/T is event counts in a time increment Δt centered at time t divided by the total counts. In the Type I model, the unzipping time constant $\tau = k^{-1}$.

The kinetic equation for the Type II model is

$$C/T = \frac{k_1 k_2}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t})\Delta \tag{3.2}$$

where k_1 and k_2 correspond to the rate constants for the first and second unzipping steps. In the Type II model, the time constants for the first and second steps τ_1 and τ_2 can be calculated from $\tau_1 = k_1^{-1}$ and $\tau_2 = k_2^{-1}$. Because k_1 and k_2 are interchangeable in eq 3.2, the fit to the equation cannot determine the rate constants for each step. In order to coordinate k_1 and k_2 with individual steps, the unzipping experiments on truncated duplexes are necessary (experimental details in Chapter 2 Table 2.2).

The histograms of unzipping durations for the G:C, OG:C, OG:D, OG:A, G:D, G:A, Sp:D, Gh:D, and 3'-Sp:A base-paired duplexes were fit using the Type I model as their shapes indicated, while whose for the 5'-Sp:A, Gh:A, Sp:C, and Gh:C base-paired duplexes were fit using the Type II model. Events were extracted using QuB (version 1.5.0.31). Histograms were plotted and fit using OriginPro (version 8.5.1). Density plots were generated using data analysis programs provided by Electronic Bio Sciences, San Diego, CA.

3.3 Results and Discussion

3.3.1 Blockage Currents for Duplexes Containing Modified Base Pairs

In this study, the 65-mer target strand that contains X (X = G, OG, Sp, or Gh) was hybridized with its complementary 17-mer containing Y (Y = C, A, or D), forming the duplex with the base pair X opposite to Y. The resultant duplex was denoted as X:Y. Ion-channel recordings were performed during unzipping of single molecules of DNA duplexes through the α-HL channel under an electrical field (Figure 3.2). The blockage due to occupation of the channel by the

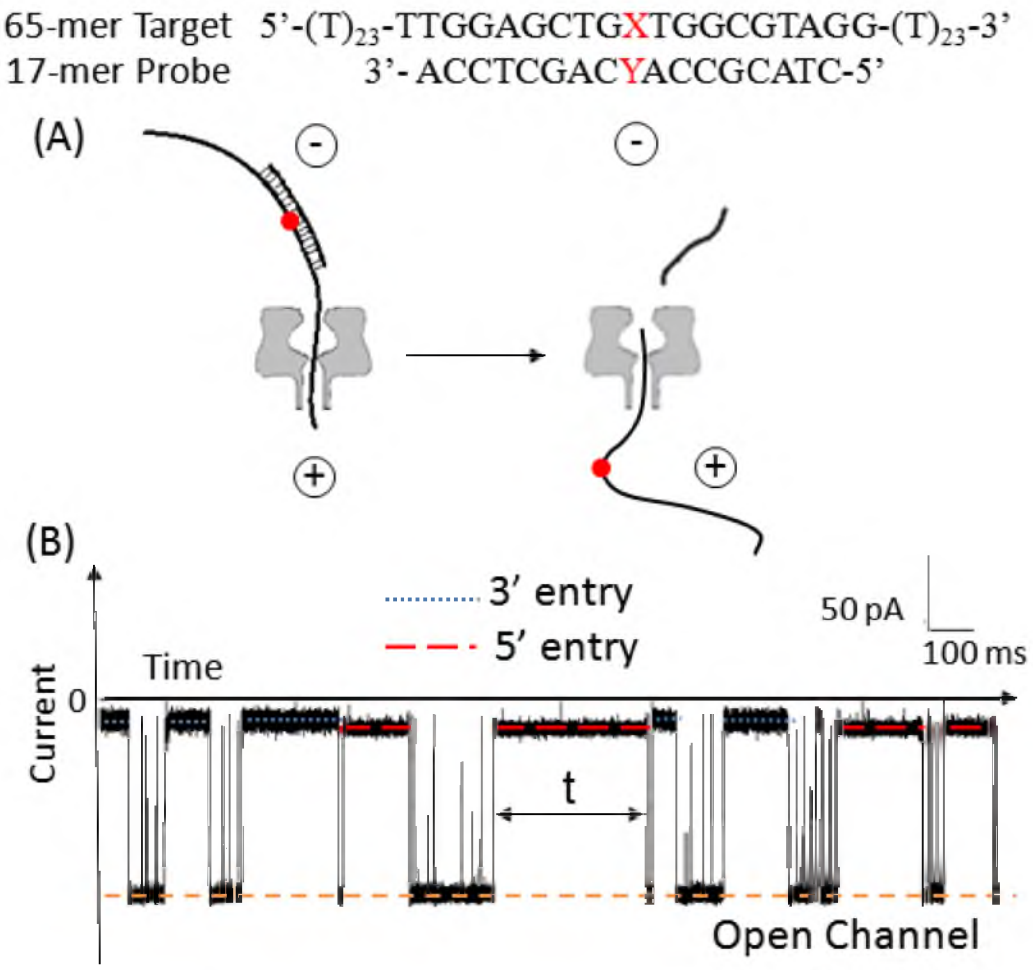


Figure 3.2. Unzipping of a DNA duplex using an α-HL nanopore under an electrical field. (A) Schematic of strand separation in a nanopore. The red dot in the target strand defines the relative position of X (X = G, OG, Sp, or Gh) opposite to Y (Y = C, A, or D) in the duplex. The + and – signs show the polarity of electrodes. (B) Example current-time trace under an applied voltage of -120 mV (*cis* versus *trans*). The blockades that last up to hundreds of milliseconds correspond to unzipping events of duplexes (unzipping duration shown as *t*). Events with duration shorter than 0.5 ms correspond to translocation of excess ssDNAs in the mixture of the 65-mer and 17-mer strands (1:5 mole ratio).

unzipping duplex yielded two distinct current levels (Figure 3.3), depending on which side of the duplex overhang threaded into the channel and consequently initiated the strand dissociation: the more blocking current level is associated with 3' entry of the duplex and the less blocking level with 5' entry (the 3' and 5' directions refer to the ends of the 65-mer target strand).[26] As shown in Figure 3.4, current histograms extracted from unzipping events for each double-overhang duplex did not display a general preference of entry direction, though the biased entry over one direction was reported for ssDNA and hairpins.[15,41] Figure 3.4 also indicates that the current blockage levels are independent of the identity of X:Y for G-, OG-, and Gh-containing duplexes. The Sp-containing duplexes, on the other hand, always generated shallower current blockages than the other duplexes by 3-5 pA, regardless of the paring base or the entry direction. Broader current distributions were observed for all of the Gh- and Sp-containing duplexes. This is best explained by the conformation variation originated from coexistence of diastereomers for both Gh and Sp.[29,42]

3.3.2 Unzipping Kinetics for Duplexes Containing Modified Base Pairs

Unzipping durations (*t)* sorted by the entry direction were used to investigate kinetics of strand separation for each X:Y duplex, which follows either a first-order reaction path (Type I) or a path of two sequential first-order reactions (Type II). In Figure 3.5, histograms of unzipping durations for the X:Y duplexes (X = G, OG, Sp, or Gh, and Y = C, or A) show a transition from the Type I shape to the Type II shape as the T_m decreases, with less stable duplexes tending to

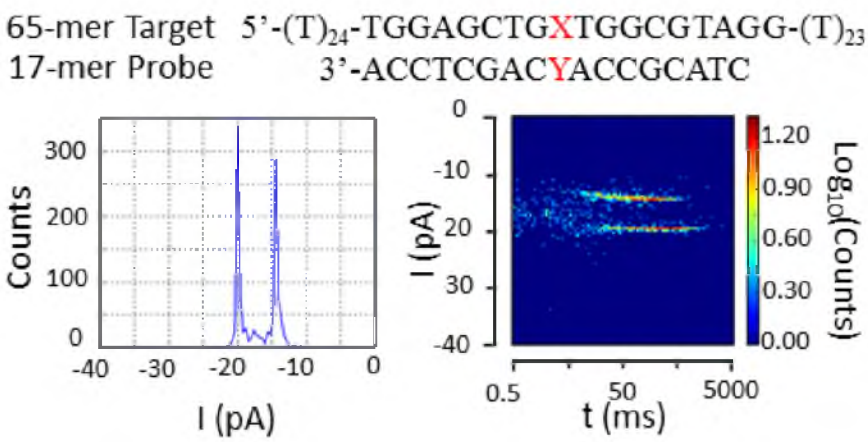


Figure 3.3. The 65mer-17mer duplex can thread into the α-HL channel either from the 3'-overhang or the 5'-overhang, each generating a distinct current blockage level. (Left) Example histogram of current blockage levels (I) generated by the duplex containing X:Y = OG:C. (Right) Event density for the current blockage level (I) of the OG:C duplex as a function of unzipping duration (t). It has been demonstrated by the previous work of our lab in main text ref 26 that 3'-threading generates a deeper blockade than 5'-threading and the unzipping time constant (τ) via 3' and 5' entry are different.

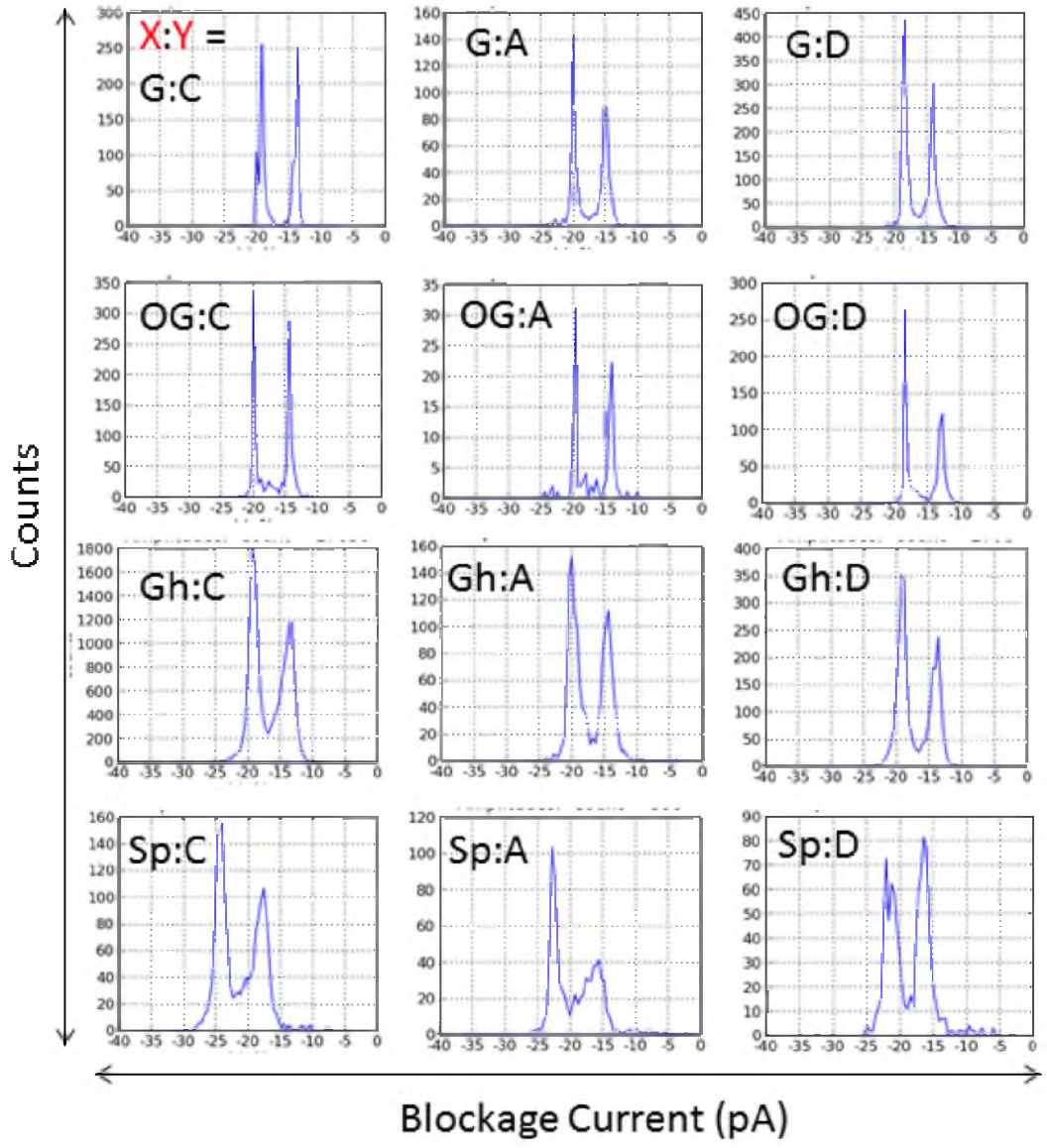
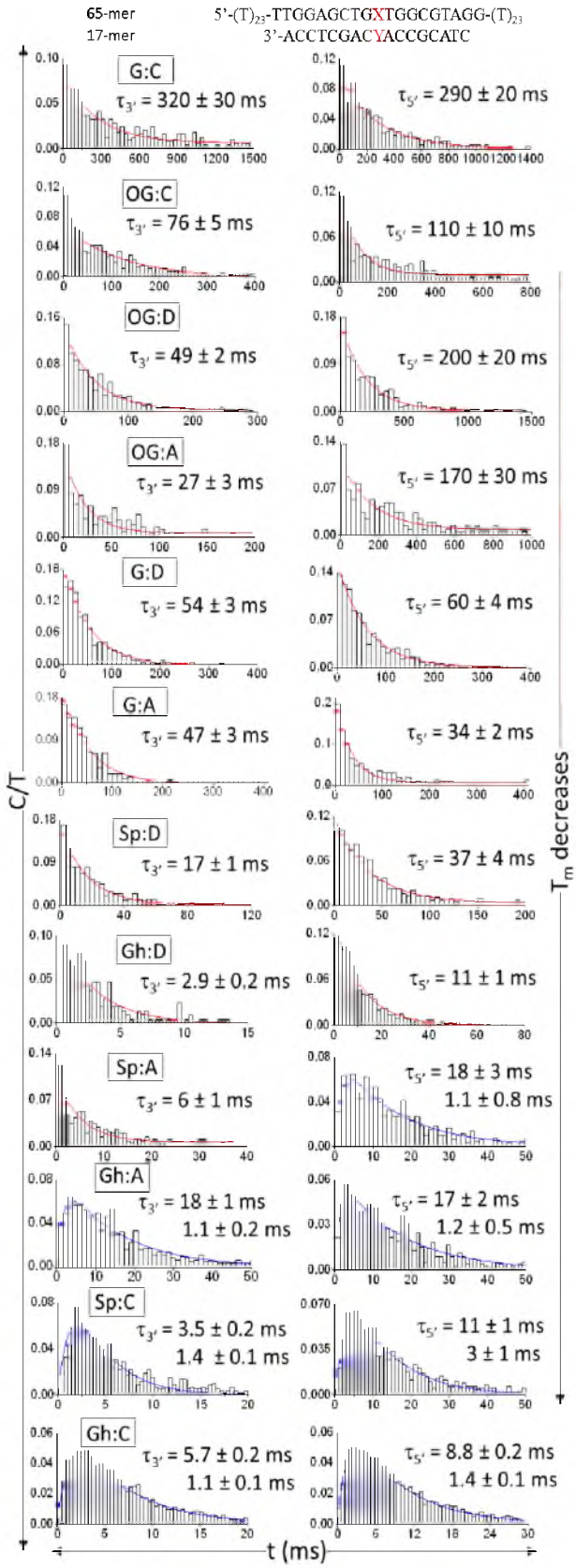


Figure 3.4. Histograms of blockage current for the duplex formed by hybridizing the X-containing 65-mer with the Y-containing 17-mer at -120 mV. X:Y represents the modified base pair.

Figure 3.5. Histograms of unzipping time (t) via 3' entry (left column) and 5' entry (right column) at -120 mV for the duplex that contains X:Y (where X = G, OG, Gh, or Sp, and Y = C, A, D). The bordered text above the histograms represents the identity of X:Y. The unzipping time constant (τ) was obtained based on the fit (red or blue curve) of the t histogram using the corresponding kinetic model, either a first-order reaction (Type I, red) or two sequential first-order reactions (Type II, blue). If the strand dissociation follows the Type II model, the unzipping time constants for each step are presented without assignment of the first and second steps. The histograms are plotted on different scales in order to emphasize the shape as a determinant for the model type.


65-mer 5'-(T)23-TTGGAGCTGXTGGCGTAGG-(T)23
17-mer 3'-ACCTCGACYACCGCATC
G:C
τ3' = 320 ± 30 ms
τ5' = 290 ± 20 ms
OG:C
τ3' = 76 ± 5 ms
τ5' = 110 ± 10 ms
OG:D
τ3' = 49 ± 2 ms
τ5' = 200 ± 20 ms
OG:A
τ3' = 27 ± 3 ms
τ5' = 170 ± 30 ms
G:D
τ3' = 54 ± 3 ms
τ5' = 60 ± 4 ms
G:A
τ3' = 47 ± 3 ms
τ5' = 34 ± 2 ms
Sp:D
τ3' = 17 ± 1 ms
τ5' = 37 ± 4 ms
Gh:D
τ3' = 2.9 ± 0.2 ms
τ5' = 11 ± 1 ms
Sp:A
τ3' = 6 ± 1 ms
τ5' = 18 ± 3 ms
1.1 ± 0.8 ms
Gh:A
τ3' = 18 ± 1 ms
1.1 ± 0.2 ms
τ5' = 17 ± 2 ms
1.2 ± 0.5 ms
Sp:C
τ3' = 3.5 ± 0.2 ms
1.4 ± 0.1 ms
τ5' = 11 ± 1 ms
3 ± 1 ms
Gh:C
τ3' = 5.7 ± 0.2 ms
1.1 ± 0.1 ms
τ5' = 8.8 ± 0.2 ms
1.4 ± 0.1 ms
C/T
t (ms)
Tm decreases

unzip in two steps. The clearly defined Type II shapes of duration histograms were observed for Gh:C and Sp:C. The lowest stability for Gh:C and Sp:C among the studied base pairs justify them as most capable of generating a pronounced energy well for the existence of an intermediate state. In contrast, the duration histograms for the G- and OG-containing duplexes have the shape of the Type I model, indicating that they have relatively stable duplex structures despite the presence of slightly destabilizing base pairs. For Sp:A/D and Gh:A/D whose stability falls between the above two groups, their histograms display the shapes of either the Type I or the Type II model, consistent with the tendency of model evolution in terms of stability.

Our previous work in Chapter 2 demonstrated that when X:Y = Gh:C, the two-step strand dissociation includes a faster separation process for the 8-bp duplex subsection on the 5' side of Gh:C than that on 3' side regardless of entry direction.[26] It is uncertain, however, whether this side-dependent stability of the subsections around other X:Y base pairs remains the same as Gh:C. Assuming the strand dissociation follows a two-step model, the unzipping time constants are presented in Figure 3.5 without assignment of the first and second steps. Additional experiments using truncated duplexes that place X:Y asymmetrically in the sequence, as demonstrated in Chapter 2, are necessary to determine the unzipping time constants for each step. The Unzipping time constant τ as a function of melting temperature T_m for the 65mer-17mer duplexes is shown in Figure 3.6.

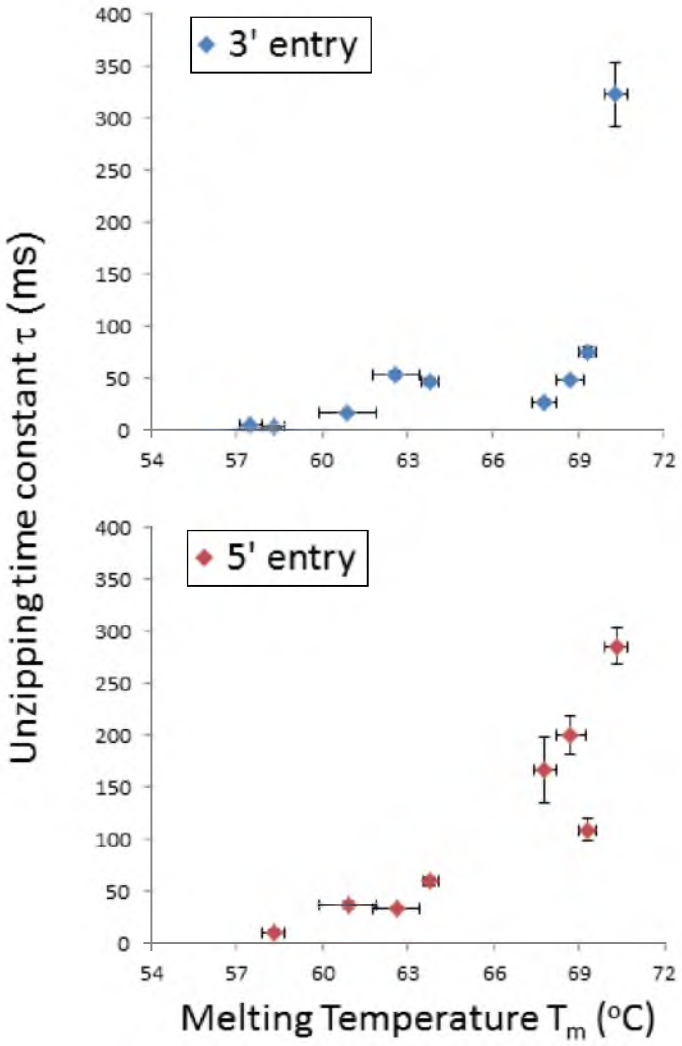


Figure 3.6. Unzipping time constant τ as a function of melting temperature T_m for the 65mer-17mer duplexes each contains a different X:Y base pair (X = G, OG, Sp, or Gh. Y = C, A, or D) at -120 mV for both 3' and 5' entry. Only the unzipping time constants for the Type I model are plotted. The x-axis and y-axis error bars are based on errors of thermal denaturation experiments and the fit errors of the unzipping time histograms into the Type I model.

3.3.3 Interpretation of the Kinetic Stability of Duplexes Containing Modified Base Pairs

The T_m order was used in Figure 3.5 to indicate the general trend of thermodynamic stability; nevertheless, we could not unequivocally predict the unzipping model based merely on T_m, especially for duplexes of moderate stability. We understand that the correlation between the unzipping duration and T_m has exceptions when the local kinetic stability of a duplex molecule is very different from the thermal stability for an ensemble of duplex molecules.[43-46] Hereby, instead of explaining the time duration results using T_m data, we rationalize the results in terms of two factors related to the local structures, the backbone distortion as the major factor, and the number of hydrogen bonds between the modified base pair as the minor factor. These two factors are discussed below.

In considering the influence of backbone distortion on duplex stability, we found that the unzipping duration is very sensitive to even a slight backbone displacement relative to the perfectly matched G:C duplex. The unzipping duration decreased by a factor of 3-4 as a result of substituting G:C with OG:C, in which the small backbone distortion originates from the repulsion between the C8 oxo group and the phosphate group in the backbone.[32] In terms of the resolution to discriminate G:C and OG:C, the nanopore method outrivals thermal denaturation methods, which show a modest 1 ^{o}C T_m difference between these two. From OG:C to OG:D, the kinetic stability did not change unambiguously based on 3' and 5' entry. Although no structural details are available, we propose that OG adopts the *syn* conformation when paired with D forming three H-bonds

(Figure 3.1).[34] Consequently, for OG:D the tension between the C8 oxo group and the phosphate backbone is eased by rotating the C8 oxo group into the minor groove, but not to the degree of a full recovery. Further, by substituting OG:D/A with G:D/A, a decrease in unzipping time was expected, due to destabilization from widening of the phosphate backbone at the G:D/A mismatch site.[47] For the Gh- and Sp-containing duplexes, the phosphate backbones are severely distorted, making them unzip faster than G- and OG-containing duplexes.[31]

As the second consideration, in cases where backbone distortion remains similar, the number of H-bonds is another possible defining feature for the interpretation of unzipping stability. The decrease in unzipping duration for OG:A versus OG:D, G:A versus G:D, and Sp:A versus Sp:D can be explained by reduction in the number of hydrogen bonds. Nevertheless, an exception occurs for Gh:A versus Gh:D. With fewer hydrogen bonds, Gh:A unzips slower than Gh:D, though the two-step model for Gh:A indicates that it is more destabilized than Gh:D. For Sp:C and Gh:C, in addition to forming poor H-bonds between Gh/Sp and C, the incorporation of these two base pairs dramatically alters the duplex backbone, resulting in the most severe destabilizing effect observed. In view of the influence of backbone distortions and hydrogen bondings, backbone distortions that generate a large-scale structural perturbation should facilitate unzipping more proficiently than removal of H-bonds whose effect is constrained most within the base pair.[48] Thus, a higher weight in determining the kinetic

stability was assigned to the influence of backbone distortion over the number of hydrogen bonds in X:Y.

The results in Figure 3.5 confirmed that the sequence-dependent unzipping duration is also conditioned to the unzipping direction. [49] From Figure 3.5, when the difference in unzipping time constants for 3' and 5' entry is resolvable, 3'-unzipping always proceeds faster than the 5'-unzipping. It has also been found that 3' entry is related to a faster passage event compared to 5' entry in ssDNA translocation and hairpin unraveling.[41, 50] Further, we found that the oxidized lesions destabilize the duplex in the 3'-orientation to a larger degree than the 5'-orientation, as evidenced by the more notable change of duration at 3' entry for modified duplexes versus G:C; the origin of the orientation-dependent destabilizing effect is not entirely understood.

In addition to studying the kinetics of strand separation in the nanopore, the results of unzipping durations also provide insight into detection of DNA oxidative damage products.[51-55] The mutagenic potentials of OG, Sp, and Gh make it important to detect them for the purpose of disease diagnosis and management.[56-60] Each of the 65-mer target strands studied here contains either G or one of its oxidative lesions placed at a mutational hot spot among a portion of the K-ras gene in codon 12.[61,62] We are able to differentiate target strands containing X = G, OG, Sp, or Gh using three different 17-mer probes, each contains Y = C, A, or D. The C probe destabilizes the target strand that contains oxidative lesions relative to the undamaged strand and causes the unzipping duration to decrease by 4 times for OG versus G and up to 60 times for Sp/Gh

versus G. The stabilizing effect of D on OG relative to G enhances the detection selectivity toward OG and increases the unzipping duration for G:D relative to OG:D by a factor of 3 at 5' entry. Specifically, Sp can be distinguished from G, OG, and Gh by inspecting the blockage current that yields from its corresponding duplex.

3.4 Conclusions

Our nanopore-based work has provided an example of using the damaged base pairs X:Y to manipulate the duplex kinetics during stand separation in α-HL. A kinetic evolution has been observed from a single-step path of first-order reaction to a path of two sequential first-order reactions, with the duplex containing the more destablizing X:Y prone to unravel in a two-step fashion. The kinetic properties of unzipping have been interpreted in terms of the duplex local structures based on the degree of backbone distortion caused by X:Y as well as the number of hydrogen bonds within X:Y. Among the two mentioned factors here, backbone distortion plays a major role in determining the stability of the duplex in α-HL. This study on various duplexes over a broad range of stability lead us toward better understanding of the physical models for the strand-separation process and information about DNA local structures and interactions, which are often concealed in the ensemble-averaged measurements. In addition, the wide selection of pairing bases, including artificial ones, allows us to have abundant options for tuning the properties of DNA and RNA complexes in nucleic acids research.

Our work also sheds light on damage detection using the nanopore unzipping concept; the damage could be detected by interrogating a damaged target strand with a customized probe. Towards the assay of biological samples, DNA strands extracted from real cells need to be amplified to increase the number of analyte copies before nanopore experiments. Comparing with normal nucleotides, the population of damaged nucleotides is about 4 orders of magnitude smaller in a real cell. PCR methods intended for damaged nucleotides are now under development to ensure that the unpopulated damaged nucleotides will not be overlooked in the mixture with normal nucleotides.

3.5 References

1. Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. Reversible Unfolding of Individual Titin Immunoglobulin Domins by AFM. *Science* **1997**, *276*, 1109.

2. Bennink, M. L.; Leuba, S. H.; Leno, G. H.; Zlatanova, J.; de Grooth, B. G.; Greve, J. Unfoling individual nucleosomes by stretching single chromatin fibers with optical tweezers. *Nature Structrual Biology* **2001**, *8*, 606.

3. Wang, M. D.; Yin, H.; Landick, R.; Gelles, H.; Block, S. M. Stretching DNA with optical tweezers. *Biophys. J.* **1997**, *71*, 1335.

4. Li, P.T.X.; Bustamante, C.; Tinoco, I. Real-time control of the energy landscape by force directs the folding of RNA molecules. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7039.

5. Cecconi, C.; Shank, E. A.; Bustamante, C.; Marqusee, S. Direct observation of the three-state folding of a single protein molecule. *Science* **2005**, *309*, 2057.

6. Koch, S. J.; Shundrovsky, A.; Jantzen, B. C.; Wang, M. D. Probing protein-DNA interactions by unzipping a single DNA double helix. *Biophys. J.* **2002**, *83*, 1098.

7. Bockelmann, U.; Thomen, P.; Essevaz-Roulet, B.; Viasnoff, V. Heslet, F. Unzipping DNA with optical Tweezers: High sequence sensitivity and force flips. *Biophy. J.* **2002**, *82*, 1537.

8. Larson, M. H.; Greenleaf, W. J.; Landick, R.; Block, S. M. Applied force reveals mechanistic and energetic details of transcription termination. *Cell* **2008**, 132, 971.

9. Cocco, S.; Monasson, R.; Marko, J. F. Force and kinetic barriers to unzipping of the DNA double helix. *Proc. Natl. Acad. Sci.* **2001**, *98*, 8608.

10. Neuman, K. C.; Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature Methods* **2008**, *5*, 491.

11. Dudko, O. K.; Mathé, J.; Szabo, A.; Meller, A. Nanopore force spectroscopy tools for analyzing single biomolecular complexes. *Mehtods Enzymol.* **2010**, *475*, 565.

12. Dudko, O. K.; Mathé, J.; Szabo, A.; Meller, A.; Hummer, G. Extracting kinetics from single-molecule force spectroscopy: nanopore unzipping of DNA hairpins. *Biophys. J.* **2007**, *92*, 4188-4195.

13. Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouax, J. E. Structure of staphylococcal α-hemolysin a heptameric transmembrane pore. *Science* **1996**, *274*,1859–1865.

14. Renner, S.; Geltinger, S.; Simmel, F. C. Nanopore translocation and force spectroscopy experiments in microemulsion droplets. *Small* **2010**, 6, 190–4.

15. Muzard, J.; Martinho, M.; Mathè, J.; Bockelmann, U.; Viasnoff, V. DNA Translocation and Unzipping through a Nanopore: Some Geometrical Effects. *Biophys. J.* **2010**, *98*, 2170–2178.

16. Vercoutere, W.; Winters-Hilt, S.; Olsen, H.; Deamer, D.; Haussler, D.; Akeson, M.][PDF] Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat. Biotechnol.* **2001**, *19*, 248–252.

17. Vercoutere, W. A.; Winters-Hilt, S.; DeGuzman, V. S.; Deamer, D.; Ridino, S. E.; Rodgers, J. T.; Olsen, H. E.; Marziali, A.; Akeson, M. Discrimination among individual Watson–Crick base pairs at the termini of single DNA hairpin molecules. *Nucleic Acids Res.* **2003**, *31*, 1311–1318.

18. Mathè, J.; Visram, H.; Viasnoff, V.; Rabin, Y.; Meller, A. Nanopore unzipping of individual DNA hairpin molecules. *Biophys. J.* **2004**, *87*, 3205–3212.

19. Kasianowicz, J.; Brandin, E.; Branton, D.; Deamer, D. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13770–13773.

20. Akeson, M.; Branton, D.; Kasianowicz, J.; Brandin, E.; Deamer, D. Microsecond timescale discrimination among polycytidylic acid, poly-adenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* **1999**, *77*, 3227–3233.

21. Dudko, O. K.; Hummer, G.; Szabo, A. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* **2006**, *96*, 108101.

22. Dudko, O. K.; Hummer, G.; Szabo, A; Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 15755.

23. Lin, J.; Fabian, M.; Sonenberg, N.; Meller, A. Nanopore detachment kinetics of poly(A) binding proteins from RNA molecules reveals the critical role of C-terminus interactions. *Biophys. J.* **2012**, *102*, 1427.

24. Gerland, U.; Bundschuh, R.; Hwa, T. Translocation of structured polynucleotides through nanopores. *Phys. Biol.* **2004**, *1*, 19.

25. Sauer-Budge, A.F.; Nyamwanda, J. A.; Lubensky, D. K.; Branton, D. Unzipping kinetics of double-stranded DNA in a nanopore. *Phys. Rev. Lett.* **2003**, *90*, 238101.

26. Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. Unzipping kinetics of duplex DNA containg oxidized lesions in an α-hemolysin nanopore. *J. Am. Chem. Soc.* **2012**, *134*, 11006.

27. Steenken, S.; Jovanovic, S. V.; How Easily Oxidizable Is DNA? One-Electron Reduction Potentials of Adenosine and Guanosine Radicals in Aqueous Solution. *J. Am. Chem. Soc.* **1997**, *119*, 617–618.

28. Luo, W.; Muller, J. G.; Burrows, C. J. The pH-dependent role of superoxide in riboflavin-catalyzed photooxidation of 8-oxo-7,8-dihydroguanosine. *Org. Lett.* **2001**, *3*, 2801–2804.

29. Luo, W.; Muller, J. G.; Rachlin, E. M.; Burrows, C. J. Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-Oxo-7,8-dihydroguanosine. *Org. Lett.* **2000**, *2*, 613–616.

30. Luo, W.; Muller, J. G.; Rachlin, E. M.; Burrows, C. J. Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model. *Chem. Res. Toxicol.* **2001**, *14*, 927–938.

31. Jia, L.; Shafirovich, V.; Shapiro, R.; Geacintov, N. E.; Broyde, S. Structrual and thermodynamic features of spiroiminodyhydantioin damaged DNA duplexes. *Biochemitry* **2005**, *44*, 13342-13353.

32. Lipscomb, L. A.; Peek, M. E.; Morningstar, M. L.; Verghis, S. M.; Miller, E. M.; Rich, A.; Essigmann, J. M.; Williams, L. D. X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 719-723.

33. McAuley-Hecht, K. E.; Leonard, G. A.; Gibson, N. J.; Thomson, J. B.; Watson, W. P.; Hunter, W. N.; Brown, T. Crystal Structure of a DNA Duplex Containing 8-Hydroxydeoxyguanine-Adenine Base Pairs. *Biochemistry* **1994**, *33*, 10266-10270.

34. Cheong, C.; Tinoco Jr., I.; Chollet, A. Thermodynamic studies of base pairing involving 2,6-diaminopurine. *Nucleic Acids Res.* **1988**, *16*, 5115-5122.

35. Fleming, A.M.; Muller, J.G.; Dlouhy, A.C.; Burrows, C.J. *J. Am. Chem. Soc.*

36. Zhang, B.; Galusha, J.; Shiozawa, P.G.; Wang, G.; Bergren, A.J.; Jones, R.M.; White, R.J.; Ervin, E.N.; Cauley, C.; White, H.S. Bench-Top Method for Fabricating Glass-Sealed Nanodisk Electrodes, Glass Nanopore Electrodes, and Glass Nanopore Membranes of Controlled Size. *Anal. Chem.* **2007**, *79*, 4778–4787.

37. White, R. J.; Ervin, E. N.; Yang, T.; Chem, X.; Daniel, S.; Cremer, P. S.; White, H. S. Single ion-channle recordings using glass nanopore membranes. *J. Am. Chem. Soc.* **2007**, *129*, 11766-11775.

38. Lan, W.-J.; White, H. S. Diffusional motion of a particle translocating through a nanopore. *ACS Nano* **2012**, *6*, 1757.

39. Schibel, A. E. P.; Heider, E. C.; Harris, J. M.; White, H. S. Fluorescence Microscopy of the Pressure-Dependent Structure of Lipid Bilayers

Suspended across Conical Nanopores. *J. Am. Chem. Soc.* **2011**, *133*, 7810–7815.

40. Lan, W.-J.; Holden, D. A.; White, H. S. Pressure-dependent ion current rectification in conical shaped glass nanopores. *J. Am. Chem. Soc.* **2011**, *133*, 13300.

41. Mathé, J.; Aksimentiev, A.; Nelson, D. R.; Schulten, K.; Meller, A. Orientation discrimination of single-stranded DNA inside the α-hemolysin membrane channel. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12377–12382.

42. Ye, Y.; Muller, J. G.; Luo, W.; Mayne, C. L.; Shallop, A. J.; Jones, R. A.; Burrows, C. J. Formation of 13C-, 15N-, and 18O-Labeled Guanidinohydantoin from Guanosine Oxidation with Singlet Oxygen. Implications for Structure and Mechanism. *J. Am. Chem. Soc.* **2003**, *125*, 3926.

43. McNally, B.; Wanunu, M.; Meller, A. Electromechanical Unzipping of Individual DNA Molecules Using Synthetic Sub-2 nm Pores. *Nano Lett.* **2008**, *8*, 3418–3422.

44. Viasnoff, V.; Chiaruttini, N.; Bockelmann, U. Probing DNA base pairing energy profiles using a nanopore. *Eur. Biophys. J.* **2009**, *38*, 263–269.

45. Renner, S.; Bessonov, A.; Gerland, U.; Simmel, F. C. Sequence-dependent unfolding kinetics of DNA hairoins studied by nanopore force spectroscopy. *J. Phys.: Condens. Matter.* **2010**, *22*, 454119.

46. Viasnoff, V.; Chiaruttini, N.; Muzard, J.; Bockelmann, U. Force fluctuations assist nanopore unzipping of DNA. *J. Phys.: Condens. Matter.* **2010**, *22*, 454122.

47. Berashevich, J.; Chakraborty, T. Thermodynamics of G·A mispairs in DNA: Continuum electrostatic model . *J. Chem. Phys.* **2009**, *130*, 015101.

48. Renner, S.; Geltinger, S.; Simmel, F. C. Nanopore translocation and force spectroscopy experiments in microemulsion droplets. *Small* **2010**, 6, 190–4.

49. Schibel, A. E. P.; Fleming, A. M.; Jin, Q.; An, N.; Liu, J.; Blakemore, C. P.; White, H. S.; Burrows, C. J. 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.

50. Butler, T. Z.; Gundlach, J. H.; Troll, M. Ionic current blockages from DNA and RNA molecules in the α-hemolysin nanopore. *Biophys. J.* **2007**, *93*, 3229.

51. Mangal, D.; Vudathala, D.; Park, J.-H.; Lee, S. H.; Penning, T. M.; Blair,I. A. Analysis of 7,8-Dihydro-8-oxo-2′-deoxyguanosine in Cellular DNA during Oxidative Stress. *Chem. Res. Toxicol.* **2009**, *22*, 788–797.

52. Steenken, S.; Jovanovic, S. V.; How Easily Oxidizable Is DNA? One-Electron Reduction Potentials of Adenosine and Guanosine Radicals in Aqueous Solution. *J. Am. Chem. Soc* **1997**, *119*, 617–618.

53. Shibutani, S.; Takeshita, M.; Grollman, A. P. Insertion of specific bases during DNA synthesis past oxidation-damaged base-8-oxodG. *Nature* **1991**, *349*, 431-434.

54. Loeb, L. A.; Harris, C. C. Advances in chemical carcinogenesis: a historical review and prospective. *Cancer Research* **2008**, *68*, 6863-6872.

55. Schibel, A.E.P.; An. N.; Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. Nanopore detection of 8-oxo-7,8-dihydro-2′-deoxyguanosine in immobilized single-stranded DNA via adduct formation to the DNA damage site. *J. Am. Chem. Soc.* **2010**, *132*, 17992.

56. Duarte, V.; Muller, J. G.; Burrows, C. J. Insertion of dGMP and dAMP during in vitro DNA synthesis opposite an oxidized form of 7,8-dihydro-8-oxoguanine. *Nucleic. Acids. Res.* **1999**, *27*, 496–502.

57. Hailer, M.K.; Slade, P.G.; Martin, B. D,; Sugden, K. D. Nei deficient Escherichia coli are sensitive to chromate and accumulate the oxidized guanine lesion spiroiminodihydantoin. *Chem. Res. Toxicol.* **2005**, *18*, 1378–1383.

58. Delaney, S.; Neeley, W. L.; Delaney, J. C; Essigmann, J.M. The substrate specificity of MutY for hyperoxidized guanine lesions in vivo. *Biochemistry* **2007**, *46*, 1448–1455.

59. Henderson, P. T.; Delaney, J. C.; Muller, J. G.; Neeley, W. L.; Tannenbaum, S.R.; Burrows, C. J.; Essigmann, J. M. The hydantoin lesions formed from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors in vivo. *Biochemistry* **2003**, *42*, 9257–9262.

60. Mangerich, A.; Knutson, C. G.; Parry, N. M.; Muthupalani, S.; Ye, W.; Prestwich, E.; Gui, L.; McFaline, J. L.; Mobley, M.; Ge, Z.; Taghizadeh, K.;

Wishnok, J, S.; Wogan, G. N.; Fox, J. G.; Tannenbaum, S. R.; Dedon, P. C. Infection-induced colitis in mice causes dynamic and tissue-specific changed in stress response and DNA damage leading to colon cancer. *Proc. Natl. Acad. Sci. USA* **2012**, early edition.

61. Xie, Y.; Yang, H,; Cunanan, C.; Okamoto, K.; Shibata, D,; Pan, J.; Barnes, D. E.; Lindahl, T.; McIlhantton, M.; Fishel, R.; Miller, J. H. Deficiencies in Mouse Myh and Ogg1 Result in Tumor Predisposition and G to T Mutations in Codon 12 of the K-ras Oncogene in Lung Tumors. *Cancer Research* **2004**, *64*, 3096-3102.

62. Husgafvel-Pursiainen, K.; Hackman, P.; Ridanpaa, M.; Anttila, S.; Karjalainen, A.; Partanen, T.; Taikina-Aho, O.; Heikkila, L.; Vainio, H. K-ras mutations in human adenocarcinoma of the lung: association with smoking and occupational exposure to asbestos. *Int. J. Cancer.* **1993**, *53*, 250–256.

CHAPTER 4

MONITORING THE ENZYME ACTIVITY OF URACIL DNA GLYCOSYLASE IN THE LATCH SENSING ZONE OF α-HEMOLYSIN

4.1 Introduction

Cytosine deamination, one of the most common forms of DNA hydrolytic damage, occurs at a considerable rate of 100-500 times per cell per day.[1,2] The consequence is that the abnormal component uracil is incorporated in DNA, causing C:G to T:A transition mutations upon replication and potentially disturbing genome integrity if left unrepaired.[3] In addition, the coding of the resulting uracil into messenger RNA may give rise to transcriptional errors in protein expression systems.[4] To repair deamination lesions, uracil DNA glycosylase (UDG) initiates the base excision repair (BER) pathway by cleaving the N-glycosylic bond between the base Uracil and the sugar of the nucleotide, leaving an abasic site (AP, Figure 4.1).[5,6] The BER pathway is completed by the coordinated work of additional enzymes: the abasic site is subsequently removed by AP-endonuclease and deoxyphosphodiesterase and replaced with the correct nucleotide by polymerase and ligase.

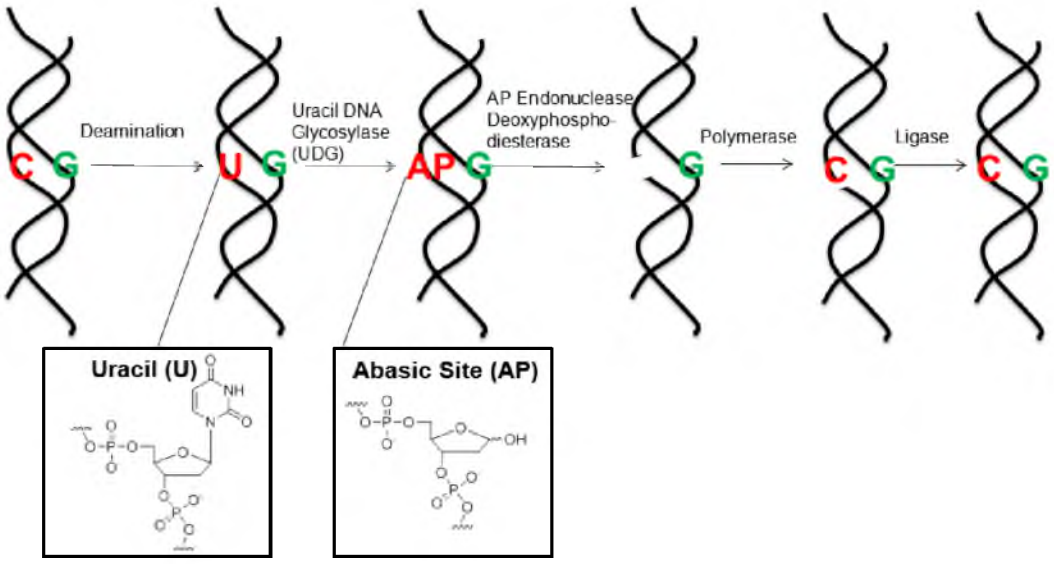


Figure 4.1. Base excision repair (BER) pathway that repairs the C-to-U deamination conversion.

Studies of enzyme kinetics have been performed on UDG to understand its catalytic mechanism and repair efficiency. The commonly used approach to measure the UDG activity involves quenching aliquots of the reaction solution at a series of time intervals followed by gel electrophoresis profiling in parallel lanes.[7-9] This method is complicated by lengthy preparation procedures of radioactive labeling as well as long gel development time. We demonstrate here a label-free and time-efficient method to monitor the UDG activity using nanopore ion-channel recordings. The protein nanopore α-hemolysin (α-HL) has been widely studied as a stochastic detector for nucleotide discrimination at the single-molecule level.[10,11] Driven by the electrophoretic force, a DNA strand is captured by the α-HL, causing a temporal blockage to the ion flow through the channel.[12] Duplex structures are required to unzip in order to translocate through the nanopore, since the duplex diameter (20 Å) is slightly larger than the narrowest constriction of the protein channel (14 Å).[13,14]

In this work, the UDG conversion of the U-containing duplex to the AP-containing duplex (sequences shown in Figure 4.2a and 4.2b) was continuously monitored by capture of the duplexes in an α-HL channel. The difference in current blockage of the AP- and U-containing duplexes during unzipping was used to determine the identity of the duplex. The quantitative conversion of the U- to AP-duplex, studied by single-molecule events of duplex unzipping, was used to extract kinetic parameters for the UDG digestion reaction. Additionally, the ability to distinguish an abasic site in a DNA duplex far from the narrowest constriction zone of α-HL indicates the presence of a well-defined, previously

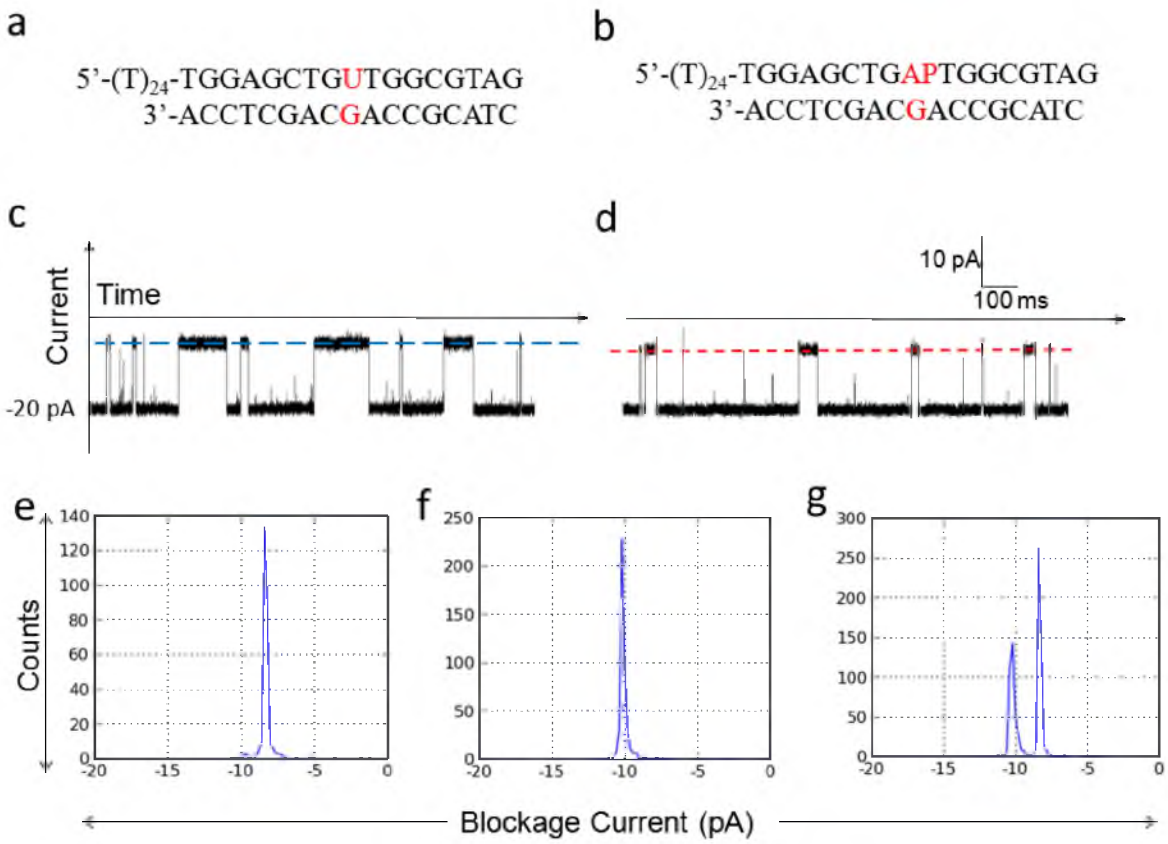


Figure 4.2. Single-nucleotide resolution is achieved between the U-containing duplex (a, c, and e) and the AP-containing duplex (b, d, and f) by ~ 2 pA based on the corresponding blockage current levels of the unzipping events. (a) Sequence of the starting material formed by a 41-mer U-containing strand hybridized to a 17-mer strand. (b) Sequence of the product containing AP. (c,d) Sample current-time traces for blockages generated by the U duplex (c) and the AP duplex (d). The blue and red lines indicate the current blockage levels used to determine the duplex identity. (e,f,g) Histograms of current blockage levels for the U duplex (e), the AP duplex (f), and a mixture of the U and AP duplexes (g, mole ratio = 2:1). Each histogram (smoothed) contains more than 250 events.

unrecognized sensing zone in this ion channel. The location of this new sensing zone coincides with the protruding region of the vestibule, or the "latch" as defined in Ref. 15.

4.2 Experimental Section

4.2.1 DNA Preparation and Purification Procedures

DNA was prepared from commercially available phosphoramidites (Glen Research, Sterling, VA) by the DNA Core Facility at the University of Utah. Afterwards, DNA was deprotected and purified using an ion-exchange column by HPLC with a linear gradient of B from 25% to 100% over 30 min while monitoring UV absorbance at 260 nm (A = 20 mM NaP_i, 1 M NaCl, pH 7 in 10% CH_3CN/90% ddH_2O, B = 10% CH_3CN/90% ddH_2O, flow rate = 1 mL/min). AP-containing duplex was prepared by converting U to AP by UDG digestion followed by HPLC purification. For mapping studies, tetrahydrofuran (THF) was used to represent an abasic site, and was incorporated into oligonucleotides during synthesis.

4.2.2 Chemicals and Materials for Nanopore Measurement

A 150 mM KCl, 20 mM Tris·HCl and 1 mM EDTA solution at pH 7.7 was used as the electrolyte for ion-channel recordings to measure the UDG activity. Because high ionic strength (> 200 mM, according to the manufacturer's instructions) reduces the catalytic activity of UDG by disrupting the electrostatic

interaction between UDG and the DNA substrate, the low salt solution was selected in comparison to the 1 M KCl solution used in Chapters 2 and 3.[16,17] A 1 M KCl, 20 mM Tris·HCl and 1 mM EDTA solution (pH = 7.5) was used to test the sensing capability of the latch zone. Wild-type α-hemolysin was purchased from List Biological Laboratories in the monomer form of lyophilized power and dissolved in water at 1 mg/mL. 1,2-diphytanoyl-sn-glycero-3-phospho-choline (DPhPC) was dissolved in decane at 10 mg/mL and used to form the bilayer. Uracil DNA glycosylase was purchased from New England Laboratories at 5000 units/mL. The bilayer was supported by a glass nanopore membrane (GNM), which was modified with a 2% (v/v) (3-cyano-propyl) dimethylchlorosilane in acetonitrile to create a moderately hydrophobic surface.[18] The DNA duplexes were annealed by mixing the 41-mer and 17-mer at a 1:5 mole ratio, followed by heating in a 90 °C water bath for 5 min and then cooling to room temperature over 3 h.

4.2.3 Current-Time Recordings

Current-time (i-t) recordings were performed at 22 ± 1 °C using a custom-built high-impedance and low-noise system (Electronic Biosciences Inc., San Diego, CA). The KCl solution was used as the electrolyte to fill the solution reservoir and the GNM capillary. A voltage was applied across the GNM between two Ag/AgCl electrodes placed inside and outside of the capillary. A lipid bilayer was deposited across the GNM orifice as indicated by a resistance increase from ~10 MΩ (associated with the open GNM) to ~100 GΩ.[19] A pressure of 20 to 40

mmHg was applied to the inside of the GNM capillary through a syringe, allowing the lipid bilayer to be functional for the protein channel reconstitution.[20] Next, 0.2 μL of α-hemolysin monomer solution at 1 mg/mL was added to the *cis* side of GNM (a volume of 300 μL). After protein reconstitution into the lipid bilayer, the duplex DNA (5 nmol for determination of current blockage levels or 15 nmol for measuring the UDG activity) was added to the solution reservoir. A voltage of -120 mV was applied (Ag/AgCl electrode placed at external solution vs. Ag/AgCl electrode inside the capillary). The i-t traces were filtered at 10 kHz and sampled at 50 kHz.

4.2.4 UDG Digestion

The UDG digestion was carried out by adding UDG (either 15, 21, or 30 units) to 15 nmol of duplex starting material in the bovine serum albumin (BSA) coated reservoir. BSA was used to prevent UDG from adsorbing to the reservoir wall made with polycarbonate. We found that the enzyme digestion took much longer and the digestion time was irreproducible without the BSA coating. The start time, t = 0 min, in monitoring the UDG reaction was set upon the enzyme addition. The concentration of the UDG stock solution (8 μM) was determined by Bradford protein assay as described in Appendix. 15, 21, and 30 units of UDG at 5000 units/mL each contains, respectively, 20, 28, 40 pmol of the enzyme.

4.2.5 Data Analysis

Based on previous reports, i-t blockades that lasted longer than 2 ms were identified as DNA unzipping events. The current amplitude of each blockade was used to determine the identity of duplex (AP-or U-containing duplex).[21] Shorter events were attributed to translocation of excess single-stranded DNAs (ssDNAs). A 5:1 mole ratio (17-mer versus 41-mer) was used to anneal the DNAs, driving the equilibrium between single strands and duplex to the side of duplex formation. Events were extracted using QuB (version 1.5.0.31). Histograms of unzipping durations were plotted using data analysis programs provided by Electronic Biosciences Inc., San Diego, CA. The percentage of AP-containing duplex (i.e., the product of the UDG reaction) was obtained by calculating the peak area of the AP-containing duplex in the current blockade histograms and ratioing this to the total peak area of both the AP- and U-containing duplexes in the same current blockade histograms.

4.3 Results and Discussion

In these experiments, a centrally placed U base pairing with G was embedded in the middle of a 17-mer:41-mer heterosequence duplex as the starting material of the UDG reaction (Figure 4.2a). A poly(dT) tail was included at the 5'-end to facilitate threading into the α-HL. The voltage-driven unzipping of DNA duplexes was initiated by pulling the 5'-tail of the molecule into an α-HL. The electrical signature of the unzipping duplex was used as the determinant of duplex identity, either containing U or AP. Our previous work discovered distinct

current levels from 3' and 5' entry if the duplex has two single-stranded DNA tails.[21,22] Thus, the 3'-end was left without a tail to avoid complication of different blockage currents that originate from the directionality effect of entry. Additionally, duplex unzipping does not occur in the absence of a 5' or 3' tail (see Appendix).

4.3.1 Single-Nucleotide Discrimination between U and AP in a Duplex

The capability of α-HL to yield single-nucleotide discrimination between U and AP was tested initially. Nanopore unzipping experiments were performed in solutions containing either the U duplex or the AP duplex (Figure 4.2a and 4.2b). Both duplexes generate a uniform level of current blockage during unzipping, with the U duplex blocking the channel ~2 pA more than the AP duplex. The unzipping durations of these two duplexes overlap; the duration histograms display a first-order exponential decay with time constants of 110 ± 15 ms and 14 ± 2 ms, respectively, for the U- and AP-containing duplexes (Figure 4.3). Accordingly, in monitoring the UDG conversion of U- to AP-containing duplex, the amplitude of the current blockage was used to determine the identity of the captured duplex.

4.3.2 Monitoring the UDG Reaction in a Nanopore

The U-containing duplex was treated with UDG directly in the nanopore solution reservoir and the UDG conversion of the U- to AP- containing duplex was monitiored by recording single-molecule unzipping events in the nanpore. A single blockage level was observed before the UDG addition, corresponding to

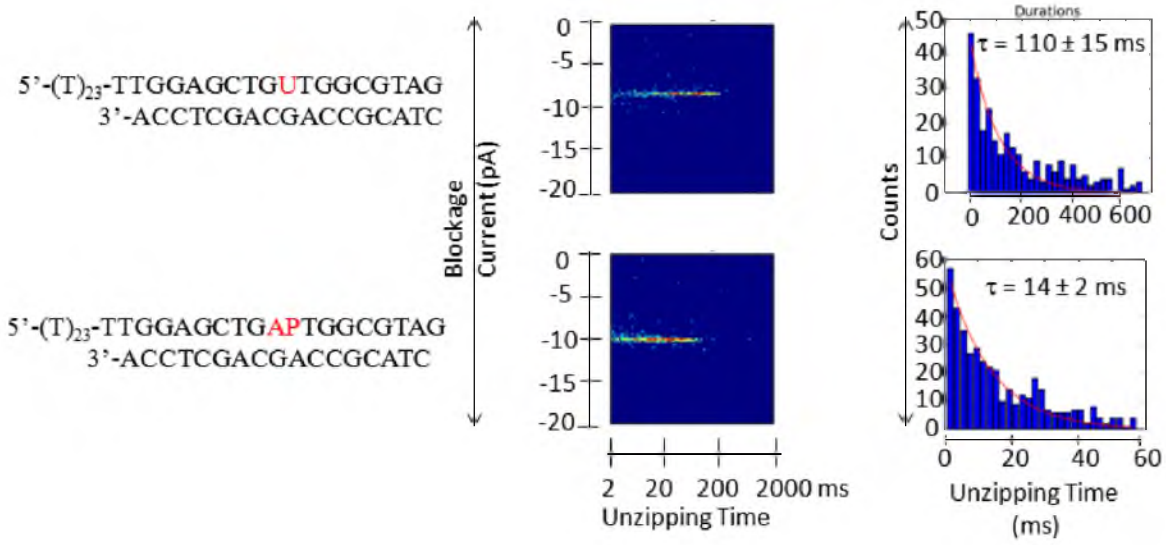


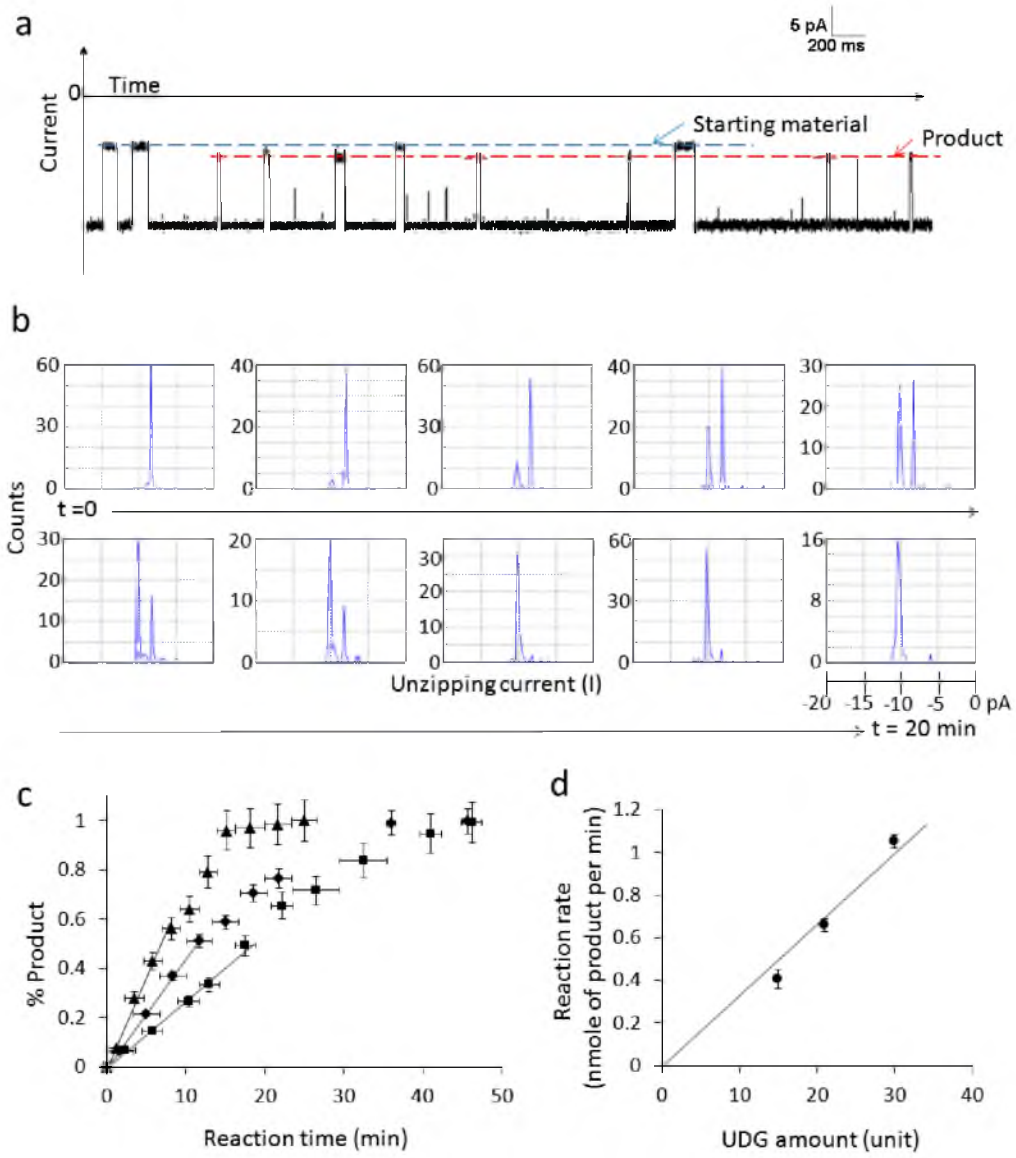
Figure 4.3. Density plots and histograms of unzipping events for the duplexes that contain either U (top) or AP (bottom) at point of interest. The strand dissociation follows a first-order exponential reaction path (rate constant τ), leading to a wide distribution of unzipping times.

the blockage caused by the U duplex. After addition of UDG to the solution, a second current blockage level corresponding to the AP-containing duplex was detected, indicating the generation of product. The identity of the duplex can be determined by examining the current amplitude of the unzipping events, with -8.3 ± 0.2 pA being attributed to the starting material and -10.2 ± 0.2 pA to the product. The progress of the reaction was demonstrated by the time-dependent histograms of blockage currents (Figure 4.4 b), in which the relative peak heights for the two species change as a function of time. The histograms were prepared from single-molecule events collected over 2-3 time intervals.

The histograms in Figure 4.4b show that completion of reaction occurs within 25 minutes. The results of these nanopore experiments are consistent with the result obtained from the gel electrophoresis method (gel result will be attached and compared quantitatively with the nanopore results).

The enzyme kinetic curves, specifically percentage of generated product as a function of reaction time, attained from the same amount of DNA substrate treated with three different UDG amounts were plotted in Figure 4.4c. From the enzyme kinetic curves, UDG produces the AP duplex at a nearly linear rate at the very beginning of the reaction (see Appendix). Afterwards, the reaction continuously slowed down until completion as indicated by the decreasing derivatives of the kinetic curves as the reaction goes on. The reaction rate of UDG was obtained from the slope of the enzyme kinetic curve at the very beginning of the reaction. In addition, there is a linear dependence of UDG

Figure 4.4. Monitoring the UDG reaction in a nanopore. (a) A 5-second long *i-t* trace collected 7 min after UDG addition. The two current blockage levels (blue and red dashed lines) are associated with starting material and product, respectively. (b) Time-dependent histograms of blockage currents correspond to the progression of the enzymatic reaction. Each histogram includes unzipping events that occur within a 2-3 min interval. (c) Enzyme kinetic curves obtained from 15 nmol of U-containing duplexes treated with 15 (purple dots), 21(blue dots), and 30 (red squares) units of UDG. The horizontal and vertical error bars are based on the time domain used to obtain each histogram and the estimated error in computing histogram peak areas. Black lines imply the linear fit to obtain the reaction rate at the beginning of the reaction. (d) Reaction rates based on the slope of black lines in (c) were plotted as a function of UDG amount. The error bars are from the linear fit of kinetic curves in (c).


a
5 pA
200 ms
Time
0
Current
Starting material
Product
b
Counts
t =0
Unzipping current (I)
-20 -15 -10 -5 0 pA
t = 20 min
c
% Product
Reaction time (min)
d
Reaction rate
(nmole of product per min)
UDG amount (unit)

amount on the reaction rate as shown in Figure 4.4d. This indicates that the enzyme is the rate-limiting element of the reaction, in agreement with the fact that the amount of substrate is in large excess compared to that of enzyme (15 nmol of DNA substrate was treated with 20, 28, and 40 pmol of UDG, respectively. Calculation details in Appendix). The time-dependent density plots of blockage current versus unzipping are shown in Figure 4.5.

4.3.3 Sensing Zone for a Duplex in the α-HL

The open channel current for wild type α-HL is ~-20 pA at 150 mM KCl; thus, the ~2 pA difference generated by substituting U by AP corresponds to a ~10% difference in residual current. Remarkably, this single-nucleotide discrimination has been performed on unlabeled DNA duplexes in wild type α-HL, in contrast with other elaborate methods being used to increase the resolution up to 2.5 – 8%, such as using engineered pores and modified bases.[23,24] It is interesting to consider the structural origin of the high resolution with which it is possible to discriminate between uracil and an abasic site.. The position of DNA structural modification (where U or AP is located) is incorporated into the the duplex section of the oligonucleotide, 9 bases from the 5' terminus of the 17-mer probe. Assuming each nucleotide is 3.4 Å long, and the duplex is driven by the electric field up to the central constriction of α-HL, the location of the U or AP within the duplex should be in vicinity of the protruding region of the vestibule, or the "latch" as defined in Ref. 15. Thus, the work reported here implies that there may be a sensing zone for duplexes in the latch region of α-HL. Until now, major efforts to

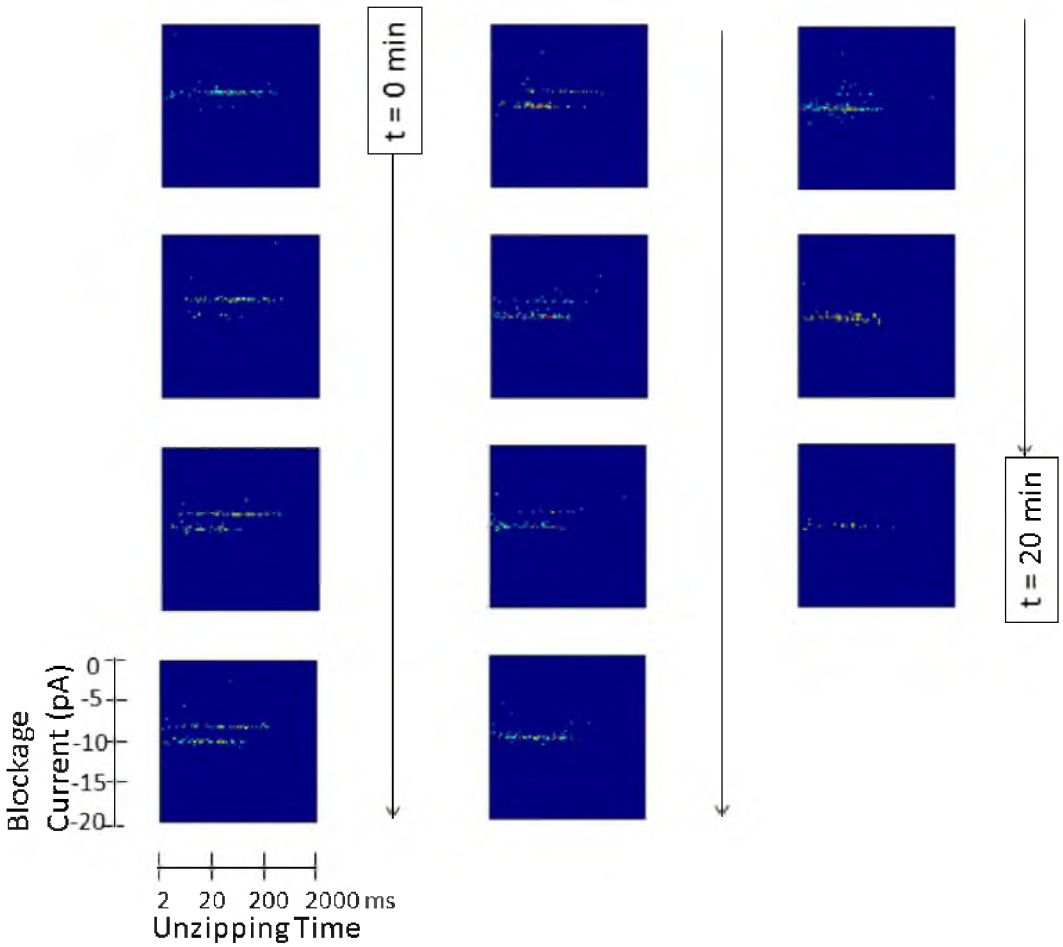


Figure 4.5. Density plots of unzipping events demonstrate the proceeding of the UDG reaction. Each plot includes events that occur within 2-3 min.

study DNA in α-HL are focused on sensing zones at central constriction or within the β-barrel section of the channel. [25-27]

To test our hypothesis, as shown in Figure 4.6, a series of duplexes containing single tetrahydrofuran (THF or F) bases, as analogues to AP sites, were used to map the blockage currents at the latch region. The F base was moved through a section of the duplex that is positioned within the latch region when the duplex is captured, in order to determine the sensitivity of residual current to the position of the duplex relative to the protein structure. F was used here instead of an abasic site for synthetic convenience, and differs from an abasic site only in the absence of an -OH group on the 1` carbon atom.

Nanopore unzipping experiments have showed that the substitution of an abasic site with F does not yield any change in current (Figure 4.7). A set of duplexes with F at positions 6-13 relative to 3'-end of the shorter strand were studied, covering a spatial range around the latch, with position 6 in the wider space of the vestibule, position 13 out of the vestibule, and position 10 located at the narrowest place. (Position 9 is the spot in which the UDG activity was measured.) In Figure 4.6, the differences in residual currents for these duplexes during unzipping were plotted against the position 13 duplex at 150 mM and 1 M KCl.

It appears that in 150 mM KCl, positions 7-12 define a recognition zone for F in the duplex context, supporting our assumption of a sensing zone for the duplex near the latch region. Our discovery of the sensing capability of the latch region coincides with a study one decade ago that the nucleotides in the hairpin

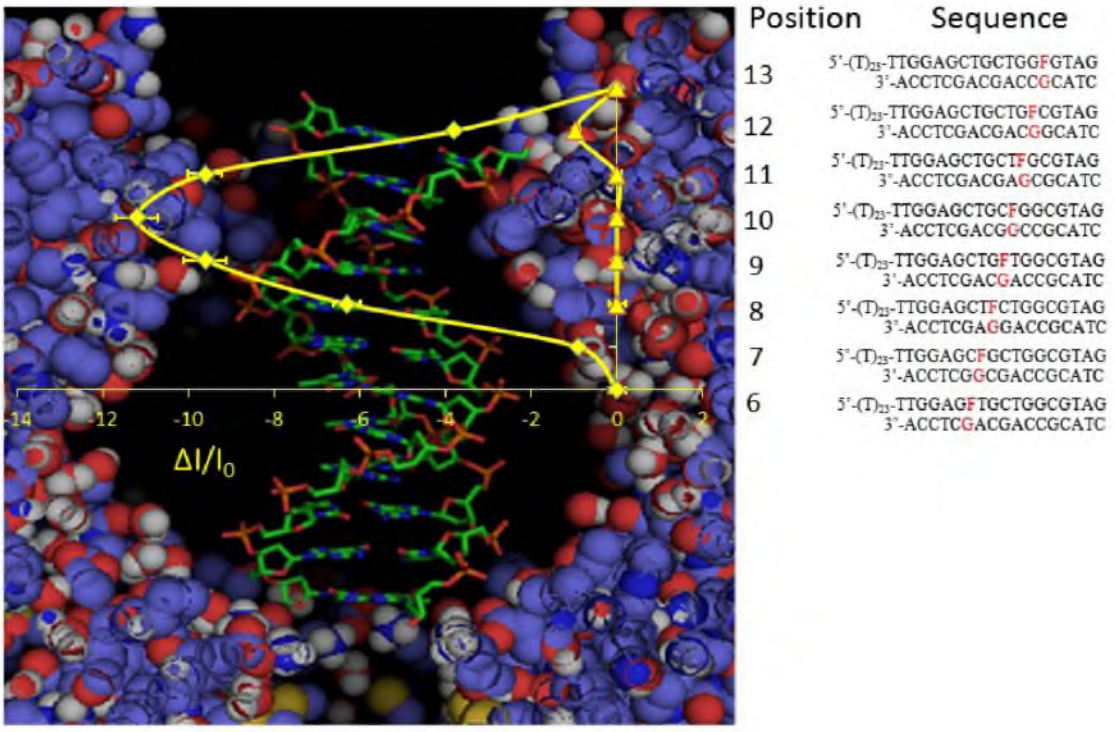


Figure 4.6. Defining a sensing zone for a duplex at the latch region of α-HL. (Left) the inner space of the α-HL vestibule and dsDNA residing within it. (Right) a set of duplexes with F:G base pair placed at positions 6-13 were examined in the nanopore experiments at 150 mM KCl (diamonds) and 1 M KCl (triangles). Positions are numbered relative to the 3'-end of the shorter strand. Residual currents of these duplexes during unzipping were plotted against the position 13 duplex ($\Delta I/I_0$). The y-axis positional data are placed at their corresponding locations in the α-HL. The error bars are based on standard deviations of the means of current blockage levels (Appendix). In the figure, dsDNA up to position 12 was shown. The actual dsDNA in the experiments extended to position 17.

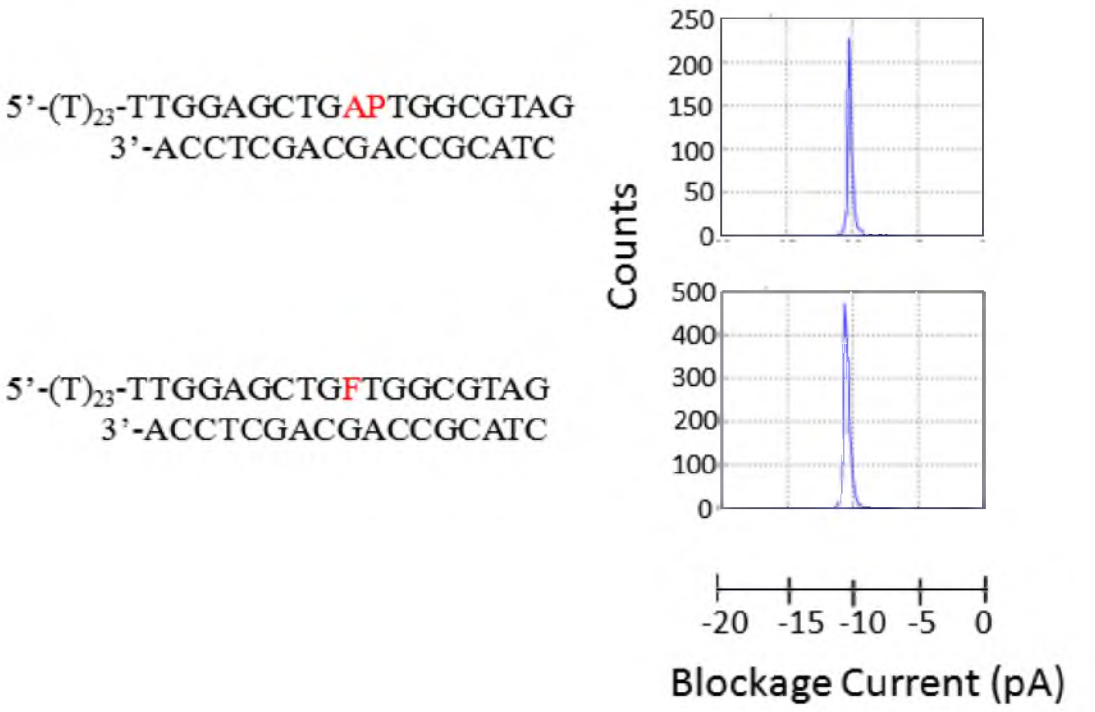


Figure 4.7. Substitution of abasic site (AP) with tetrahydrofuran (F) does not cause any change in terms of unzipping current.

loop were differentiated when the loop was placed near the α-HL latch in 1 M KCl.[25] Interestingly, in 1 M KCl, we did not observe a pronounced sensing zone for F at the latch. The origins of the latch sensing zone are discussed below.

The detection sensitivity yielded at positions 7-12 at 150 mM KCl may be attributed to the size exclusion effect of dsDNA within the channel. The proximity of the cross-section size for the duplex (2.0 nm) and the latch region (2.4 nm) suggests that the overall current blockage may be dominated by duplex filling most of the volume in the latch region, thereby constricting the ion flux in that region.[15] Generally, a smaller inner space is associated with a stronger effect of ion constriction, thus better sensitivity.[24,26,27]

Other factors besides the effect of size exclusion need to be considered to explain the resolution yielded at the latch. Admitting that the interactions between the amino acid wall and DNA can influence blockage current levels, it is not easy to rationalize the change of blockage current levels for a specific base by protein-DNA interactions due to the complicated nature of interactions.[27] We propose that the richness of neutral N-terminal groups around the latch may allow for abundant polar interactions, such as through dipole-dipole, dipole-ion interactions and hydrogen bonds, between DNA and the amino wall, potentially contributing to the unique signal yield at that position. In 1 M KCl, the polar interactions are weakened as the salt screens molecular dipoles. The loss of sensitivity in 1 M KCl at the latch might be related to this screening effect. We are currently investigating other salt concentrations to determine the salt dependence on the sensing capability of the latch.

4.4 Conclusions

In conclusion, we present a label-free and fast readout method of measuring the UDG enzyme activity using the nanopore approach in an α-HL nanopore. The UDG reaction was monitored based on the difference in current blockage levels generated by the unzipping duplexes captured into the nanopore. This work can also be adapted to monitor the activity of other enzymes if there is a change in DNA electrical signatures of the duplexes before and after the enzyme action.

The exceptional single-nucleotide resolution in monitoring the UDG reaction was is attributed to a recognition site at the latch in the vestibule of α-HL, specific to dsDNA in contrast to the previously acknowledged ssDNA recognition sites in the central constriction and β-barrel sections. The mapping of the latch region indicates that this region corresponds to a new sensing zone for duplexes. The discovery of a sensing zone at the latch will embark on new research concerning DNA characterization in α-HL. The present body of work addressed the detection of abasic sites in low salt conditions within wild-type α-HL. It would be interesting to see if the detection sensitivity yielded at the latch can be extended to other nucleotides or lesions in a wider range of experimental conditions.

The proposal of the new sensing zone also sheds light on hot spots of protein mutagenesis in α-HL. Mutant proteins provide possibilities of tailoring the discrimination properties of the latch for various other bases and lesions. Further, our work demonstrated a sensing zone in wild-type α-HL that spans over five nucleotides. By altering the amino acids at the latch in mutant α-HL, the sensing

zone can potentially be sharpened to cover less nucleotides or even a single nucleotide.

4.5 Appendix

4.5.1 Bradford Protein Assay to Determine the Concentration of the UDG Solution

This assay is based on a shift of UV absorbance to 595 nm caused by the dye Commassic Blue binding to the protein. The Commassic Blue dye reagent (0.1%) was purchased from USA Corp. and diluted with distilled water by 5 folds. A BSA solution at 1 mg/mL was chosen as the protein standard to obtain the calibration curve. 300 μL of diluted dye reagent was mixed with various amounts of BSA solution (1, 2, 3, 4, 5 μL, respectively). After 15 min, the UV absorbances at 595 nm for the mixed solutions were measured. The calibration curve is shown in Figure 4.8.

Next, 2.2 μL of the UDG solution was mixed with 300 μL of diluted dye reagent and the resulting absorbance was monitored at 595 nm. It was calculated that the UDG solution has a concentration of 8 μM. The three UDG amounts used in the experiments correspond to 20, 28, and 40 pmol of UDG, respectively.

4.5.2 Current Blockage Levels and Their Errors for the Duplexes with F Placed in Positions 6-13

For each duplex with F placed in a specific position (duplex A), it was tested on an α-HL by itself and then examined with a reference duplex on the same α-

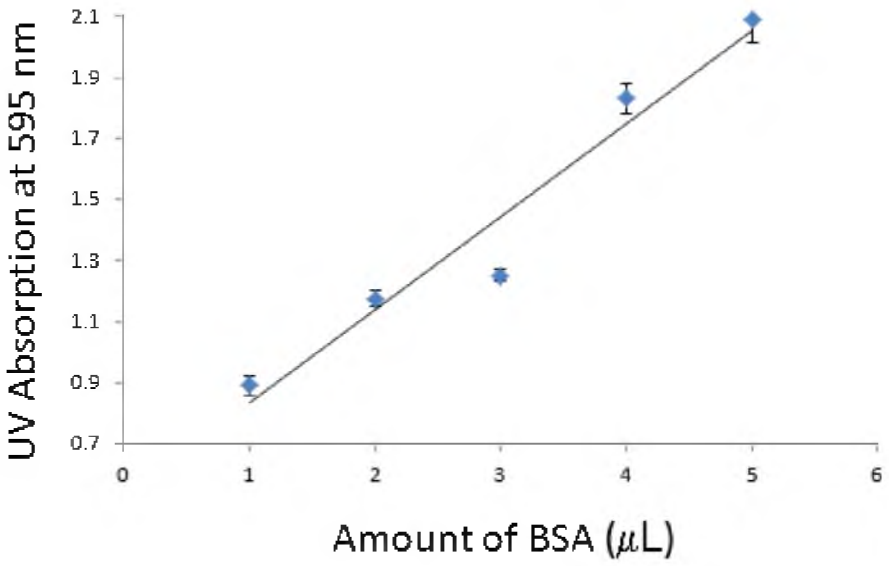


Figure 4.8. UV absorption as a function of BSA amount.

HL. The current-time traces were collected for duplex A first, followed by the addition of a reference duplex to the solution reservoir. Subsequently, current-time traces were obtained for the mixture of duplex A and the reference.

In different experiments, it is not unusual for protein channels to have a slight variation in terms of current. This is a systematic error inherited in the ion-channel measurement using α-HL. The reference functions as an internal standard and helps eliminate the interference from protein channels. The current difference between duplex A and the reference were examined. Individual nanopore experiments for each duplex A plus the reference at 150 mM KCl are shown in Figures 4.9-4.14.

How to obtain Errors of Residual Currents is discussed below.

Theories of the mean, standard deviation, and standard deviation of the mean are introduced first. (Reference: Sprinthall, R. C. Basic Statistic Analysis, 7th ed.; Pearson Education Group, Inc.: Massachusetts, 2003; pp 150-155.)

In a sample set $\{x\} = \{x_1, x_2, \ldots, x_N\}$, the mean is defined by

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

The standard deviation of {x} is calculated as

$$\sigma = [\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2]^{\frac{1}{2}}$$

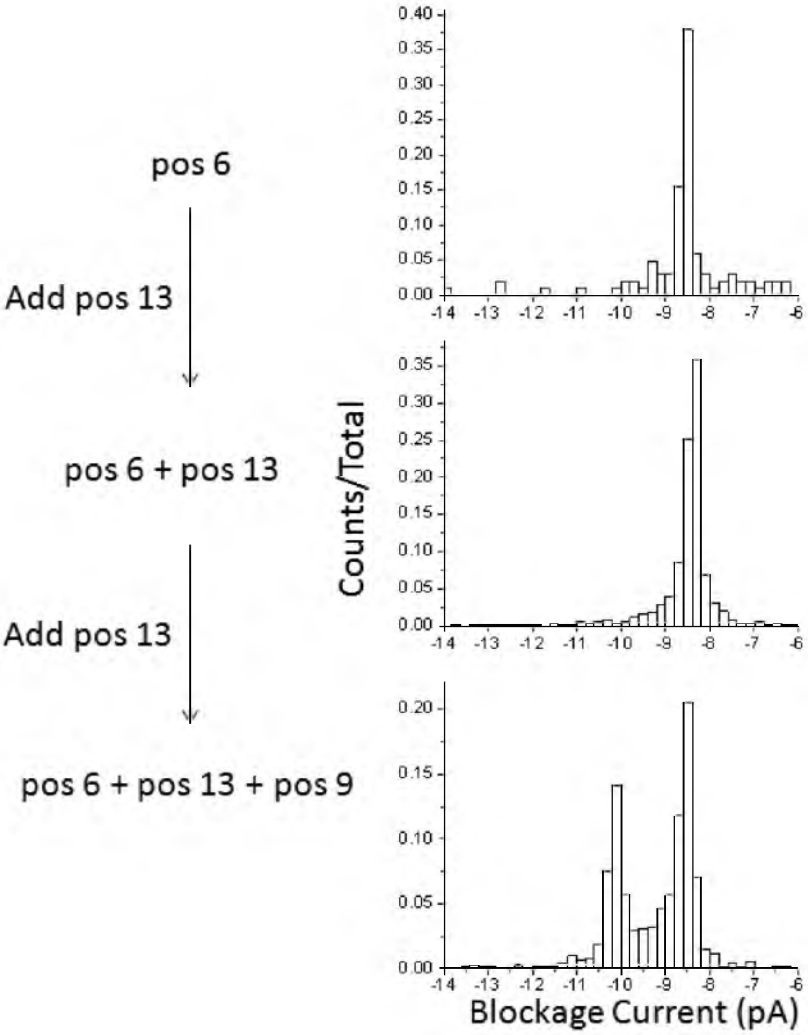


Figure 4.9. Current blockage histograms for pos 6, 13 and 9. (Note: pos X represents the duplex with F placed at position X.)

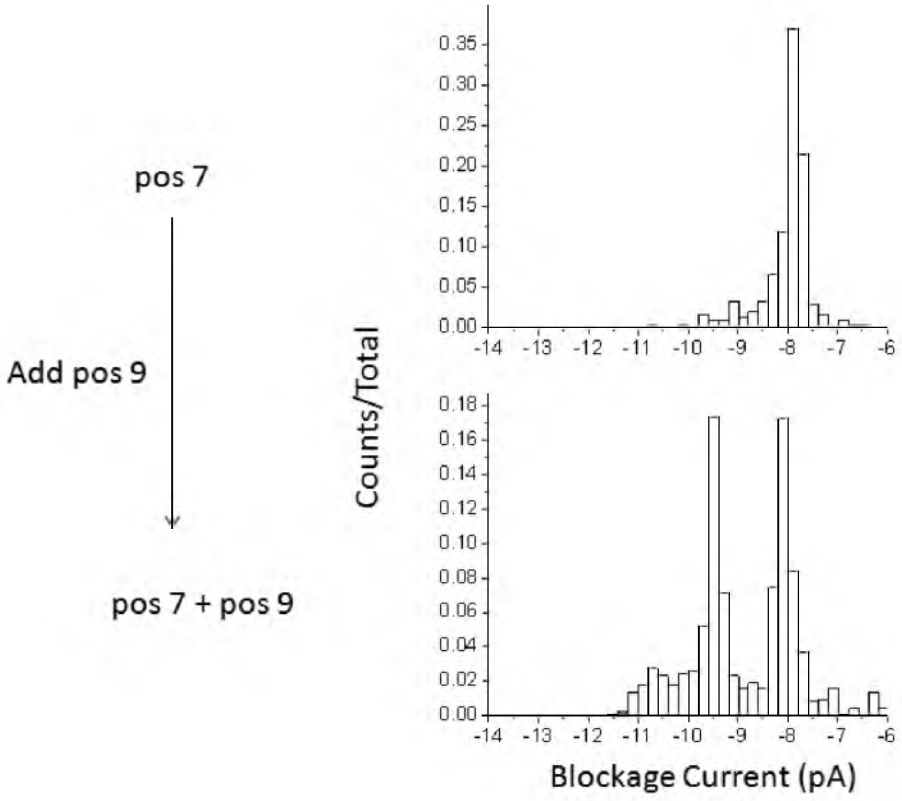


Figure 4.10. Current blockage histograms for pos 7 and 9.

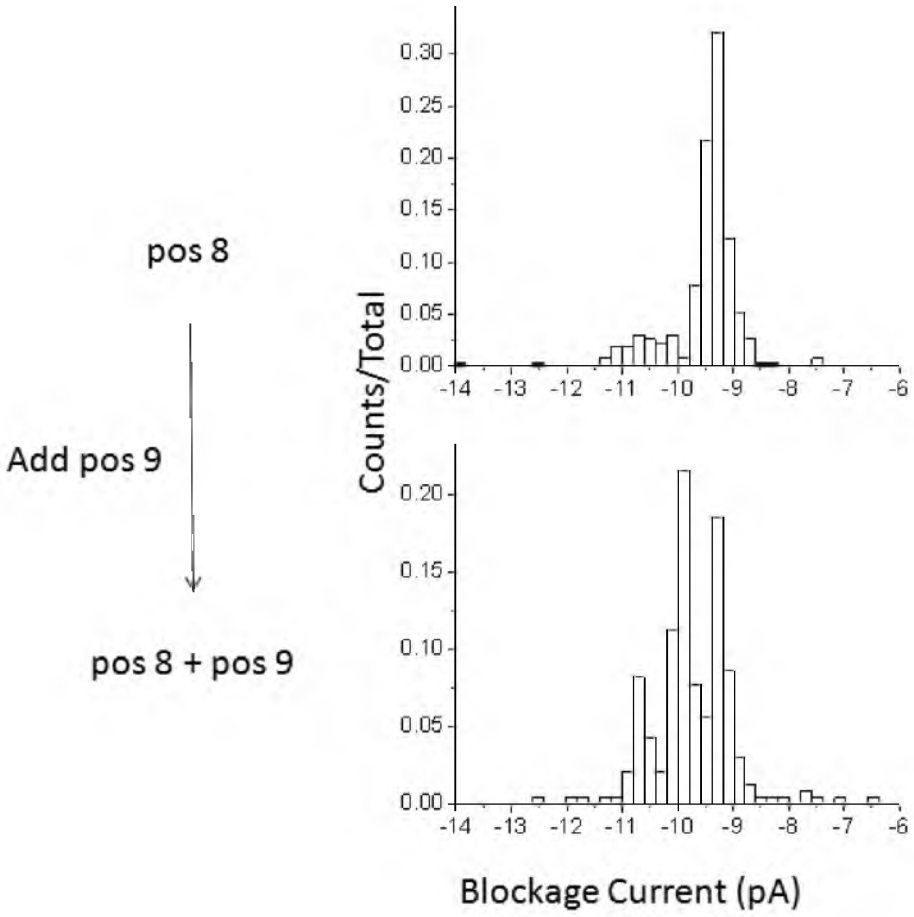


Figure 4.11. Current blockage histograms for pos 8 and 9. Interestingly, the current histogram for pos 8 displays a major peak and a less-blocking minor peak that accounts for less than 10% of events. When calculating the residual current shown in Figure 4.4, only the major peak was considered.

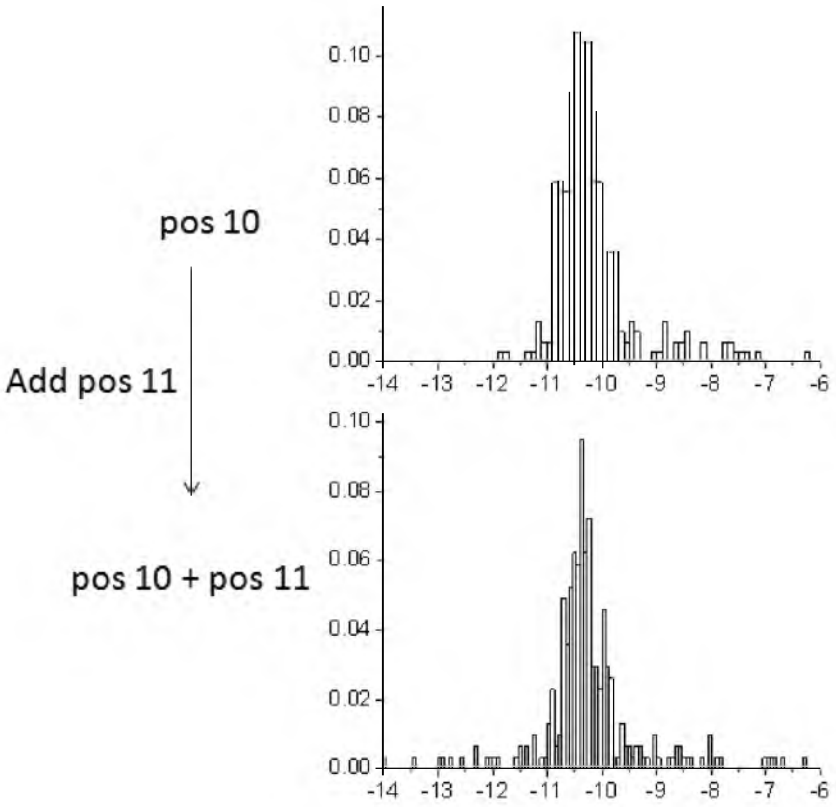


Figure 4.12. Current blockage histograms for pos 10 and 11.

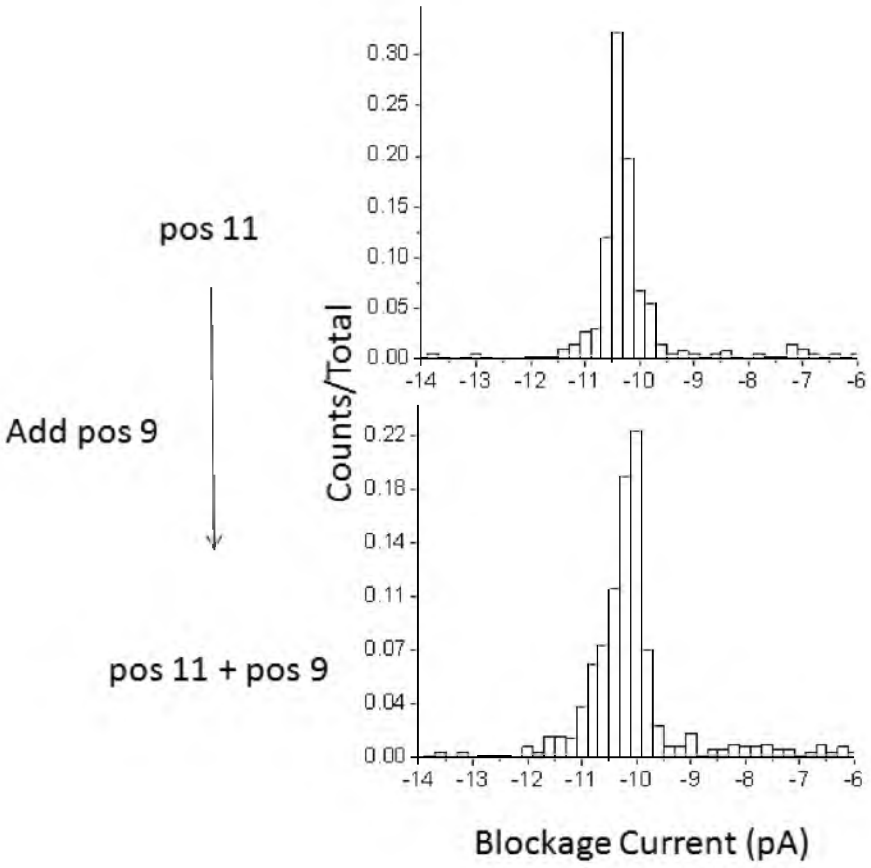


Figure 4.13. Current blockage histograms for pos 11 and 9.

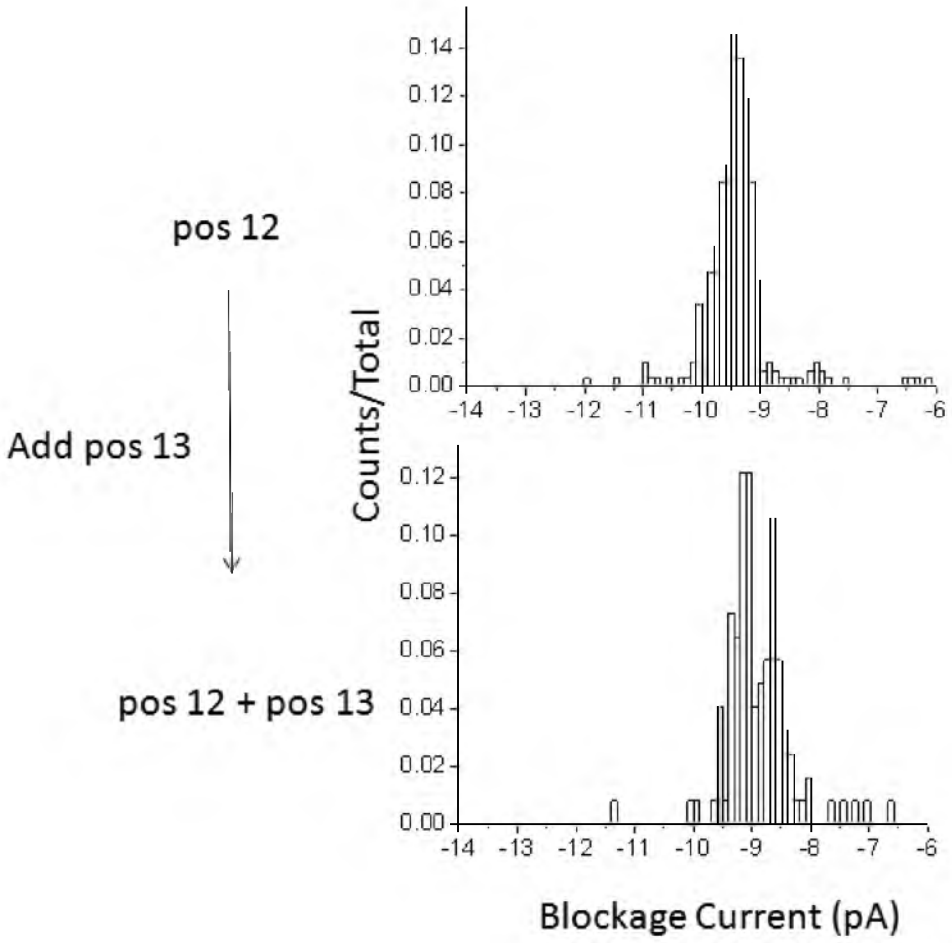


Figure 4.14. Current blockage histograms for pos 12 and 13.

where σ shows how much variation there is between x_i and mean.

Noticing that the sample set is a portion of the whole population, we expect the mean of the sample set ($\bar{x}$) to deviate from the average of the whole population (μ). This deviation is indicated by the standard deviations of the means (SDOM), a measure of errors from sampling variability

$$\text{SDOM} = \frac{1}{\sqrt{N}}\sigma$$

SDOM decreases as the size of the sample set increases. When an infinite number of samples are measured ($N \rightarrow$), SDOM is approaching 0 and $\bar{x} = \mu$.

In a nanopore experiment, we collected a finite number of samples (n) regarding captured duplexes out of all of the duplexes in the solution. SDOM describes the variability between the sample mean of blockage current (I_m) obtained from n events and the averaged blockage current for the whole population. x and N in the above theories correspond to blockage current (I) and number of events (n) in this case. As more events are included in the sample set of blockage currents {I}, the accuracy of I_m to the real average of the whole population is improved.

In Figure 4.4, the error bars of $\Delta I/I_0$ were obtained from SDOM of blockage currents. Below, an example is given to demonstrate the calculation of the error bar of $\Delta I/I_0$ at pos 9.

Based on the separation of peaks in the lower histogram of Figure 4.9, we assume -9.5 pA as the dividing line between pos 6/13 and pos 9. (Note: the peaks of blockage currents for pos 6 and pos 13 are overlapped. These two duplexes were discussed together as pos 6/13). In the total collected unzipping events, 531 events (n_1 = 531) belong to pos 6/13 and 383 events (n_2 = 383) belong to pos 9. Although the mole ratio of pos 6/13 vs. pos 9 is 2:1 in the mixture, the events collected in nanopore did not follow the same ratio. This discrepancy might be attributed to uneven mixing when adding samples to the solution or the inherent difference of entry rate for different duplexes.

The means of blockage currents were obtained from the averages of n_1 and n_2 events.

$I_{m,n1}$ = - 8.27 pA, $I_{m,n2}$ = -10.22 pA

The standard deviations of these n_1 and n_2 events were calculated.

σ_{n1} = 1.36 pA, σ_{n2} = 0.76 pA

SDOM were obtained by dividing σ by the square root of n_1 or n_2.

$SDOM_{n1} = \sigma_{n1}/(n_1)^{1/2}$ = 1.36/(536)$^{1/2}$ pA = 0.06 pA,

$SDOM_{n2} = \sigma_{n2}/(n_2)^{1/2}$ = 0.76/(383)$^{1/2}$ pA = 0.04 pA

Thus, the blockage currents of I_{n1} and I_{n2} are

$I_{n1} = I_{m,n1} \pm SDOM_{n1} = -8.27 \pm 0.06$ pA,

$I_{n2} = I_{m,n2} \pm SDOM_{n2} = -10.22 \pm 0.04$ pA

The difference in blockage current is

$\Delta I = I_{n2} - I_{n1} = -1.95 \pm 0.07$ pA

The propagation of errors was considered in estimating the error of ΔI. ΔI was then normalized by open channel current to yield $\Delta I/I_0$ as the result shown in Figure 4.4. The SDOM of open channel current (< 0.0005 pA) was ignored.

When two peaks in the current histograms were partially overlapped, such as in Figure 4.11, 4.12, and 4.14, the lowest bar between the two peaks was selected as the position of the dividing line.

To emphasize the resolution yielded at the latch positions, the blockage current at position 13 (an insensitive position away from the latch) was arbitrarily set as 0. In Figure 4.4, the residual currents relative to position 13 for each positional duplex were plotted. Tables 4.1 and 4.2 display the means and standard deviations for each positional sample at 150 mM and 1 M KCl.

Table 4.1 Means and SDOM for each positional sample at 150 mM KCl.

Exp #	tested pos #	Note	pos # for I_m and s.d.	number of events n	I_m	s.d. of events	SDOM	Pos # to yield ΔI	ΔI	s.d. of ΔI
1	6		6	103	-8.59	1.15	0.11			
2	6,13	overlap	6,13	1557	-8.52	0.71	0.02	6 vs 13	0	0.03
3	6,13,9	6 and 13 overlap	6,13	531	-8.27	1.36	0.06	6,13 vs 9	-1.95	0.07
		9 is separated	9	383	-10.22	0.76	0.04			
4	7		7	295	-8.01	0.52	0.03			
5	7,9	separated	7	333	-7.94	0.47	0.02	7 vs. 9	-1.82	0.04
			9	358	-9.76	0.56	0.03			
6	8		8	268	-9.52	0.65	0.04			
7	8,9	separated	8	76	-9.19	0.18	0.02	8 vs. 9	-0.69	0.03
			9	107	-9.88	0.17	0.02			
8	11		11	396	-10.19	1.13	0.06			
9	11,9	overlap	11,9	558	-9.85	1.56	0.07	11 vs. 9	0	0.10
10	10		10	297	-10.18	0.73	0.04			
11	10,11	separated	10	201	-10.23	0.58	0.04	10 vs. 11	0.36	0.11
			11	84	-9.87	0.94	0.10			
12	12		12	335	-8.82	1.77	0.10			
13	12,13	separated	12	46	-8.41	0.48	0.07	12 vs. 13	-0.83	0.08
			13	65	-9.24	0.36	0.04			

Table 4.2. Means and SDOM for positional sample at 1 M KCl.

Exp #	tested pos #	Note	pos # for I_m and s.d.	number of events n	I_m	s.d. of events	SDOM	Pos # to yield ΔI	ΔI	s.d. of ΔI
1	8, 13	not separable	8, 13	489	-21	1	0.09	8 vs 13	0	0.13
2	9, 13	not separable	9, 13	689	-20	1	0.05	9 vs 13	0	0.08
3	10, 13	not separable	10, 13	7735	-19	2	0.03	10 vs 13	0	0.04
4	11, 13	not separable	11, 13	4358	-20	2	0.04	11 vs 13	0	0.08
5	12, 13	slightly separable	12	2044	-18.6	0.3	0.006	12 vs. 13	-1.18	0.01
			13	3052	-17.4	0.6	0.01			

4.5.3 Constant Reaction Rate at the Beginning of an Enzyme Reaction

A typical enzyme reaction involves the formation of the enzyme-substrate complex and the breakdown of it to yield the product. The mechanism can be presented as

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES$$

$$ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$$

where E, S, and ES each represents enzyme, substrate, and enzyme-substrate complex. k_1 and k_{-1} are the forward and reverse rate constants of the ES formation reaction. k_2 and k_{-2} are the forward and reverse rate constants of the ES breakdown reaction.

The formation rate of ES is

$$k_1([E_t] - ES])[S]$$

where $[E_t]$ is the total concentration of enzyme. The formation of ES from E + P is neglected since k_{-2} is small.

The breakdown rate of ES is

$$k_{-}[ES] + {}_2[ES]$$

At the steady state, the formation rate is equal to the breakdown rate. Therefore,

$$k_1([E_t]- ES])[S] = [ES] + [ES]$$

The above equation can be rearranged and yields

$$[ES] = \frac{[E_t][S]}{[S] + k_2 +)/k_1}$$

The rate of product generation is

$$\frac{d[P]}{dt} = [ES] = \frac{k_2[E_t][S]}{[S] + k_2 +)/k_1}$$

At the beginning of the reaction, DNA substrate amount is in large excess compared with enzyme, thus [S] can be assumed as a constant. The entire right side of the above reaction can therefore be considered as a constant. (Reference: Lehninger, A. L. Principles of Biochemistry; Worth Publishers, Inc.: New York, 1982; pp 214-215.)

4.5.4 Nanopore Experiments with No-Tail Duplexes

Truncated DNA target strands without poly(dT) tails (5'-TGGAGCTGUTGG CGTAG-3') were annealed with 17mer probe strands (3'-CCTCGACGACCGCA TC-5') to form the no-tail duplexes. The mixing ratio for target vs. probe is 1:5. Nanopore experiments were carried out on no-tail duplexes samples that contain duplexes and single stranded probe strands in excess at 120 mV, 150 mM KCl. In Figure 4.15, the current-time trace was compared with that obtained from 5'-tail duplexes.

The no-tail duplex sample did not show any events longer than 2 ms associated with duplex unzipping. The short events as displayed above were due to translocation of excess single strands. The unzipping events analyzed in main text using 5'-tail duplexes were longer than 2 ms and should be attributed to 5' entry.

4.5.5 Tethering Experiments with DNAs of Different Length

Current-time traces for strept-biotin-20mer, 30mer, and 40mer under -120 mV (*cis* vs *trans*) are shown in Figure 4.16. The structures and noise levels for these DNAs are demonstrated in Figure 4.17. At -120 mV, strept-biotin-20mer can enter into and escape from the channel by itself, generating the fluctuation of current in the upper current-time trace. The time that the 20mer complex could stay in the channel is always shorter than 300 ms. For the longer ssDNA, once being captured, the DNA complex can be immobilized indefinitely in the channel, therefore voltage switch is necessary to release the DNA complex out of channel.

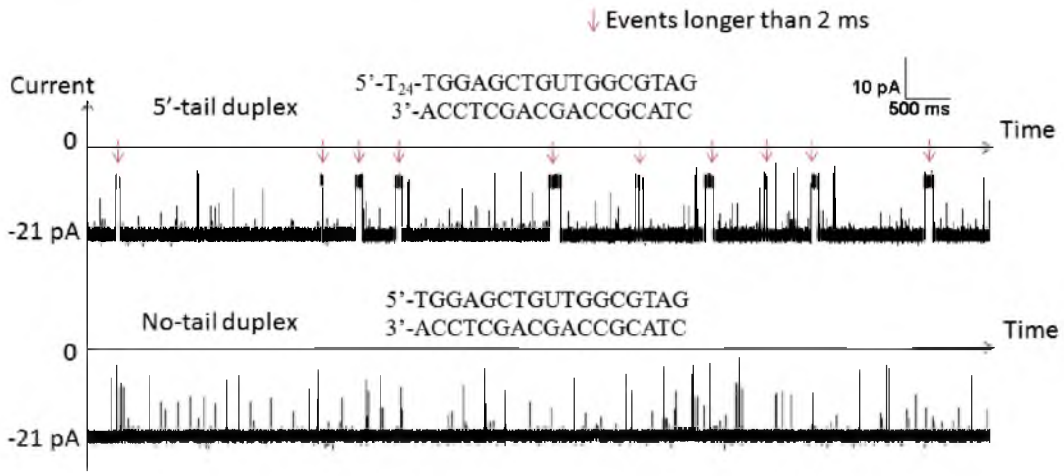


Figure 4.15. Current-time traces for the 5'-tail duplex and no-tail duplex in 150 mM KCl at 120 mV.

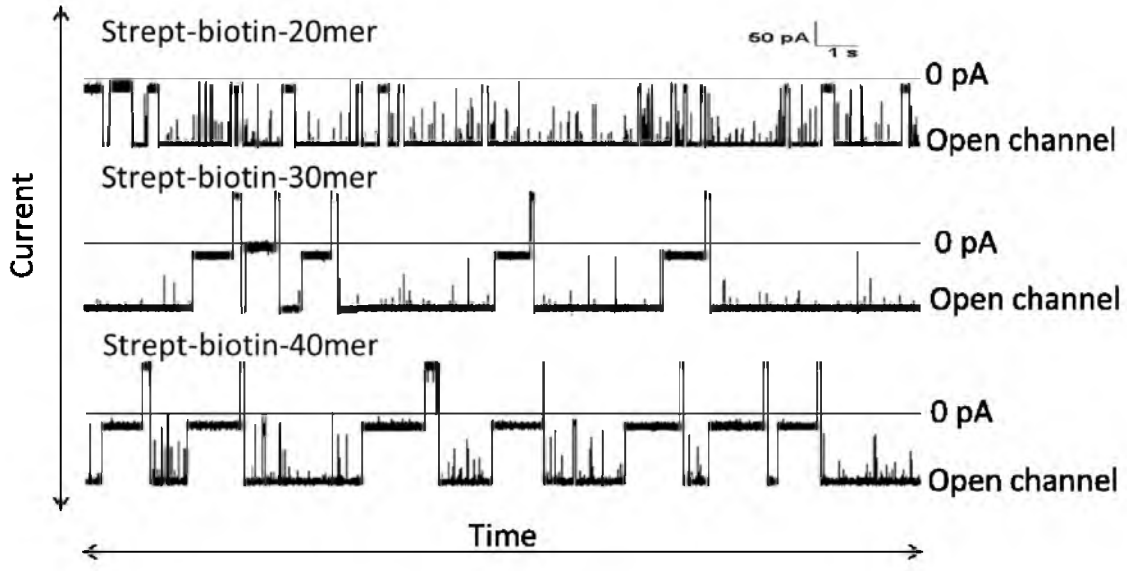


Figure 4.16. Current-time traces for strept-biotin-20mer, 30mer, and 40mer under -120 mV (*cis* vs *trans*).

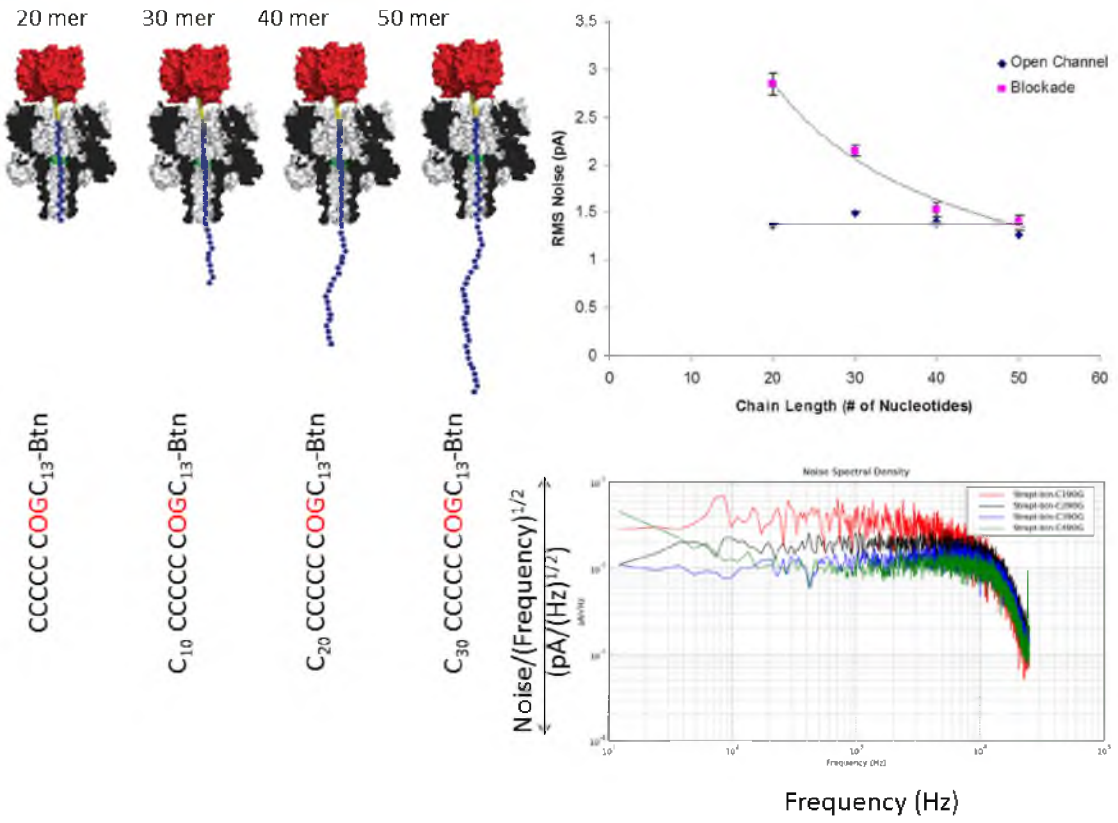


Figure 4.17. The structures and noise levels for strept-biotin-20mer, 30mer, and 40mer at -120 mV. (left) Structures of these ssDNA bound to strept-biotin anchors. (top right) RMS noise as a function of chain length of the tethered DNA. (lower right) Noise spectral density of these strept-biotin-ssDNAs.

In the middle and bottom traces, voltage was switched to + 120 mV if the blockage lasted longer than 800 ms in order to eject the strept-biotin-ssDNA complex out of the channel. After staying at +120 mV for 100-300 ms, the voltage was switched back to -120 mV, allowing the next capture event to occur. 20mer, 30 mer and 40 mer all have homogeneous poly(dC) sequences with a single substitution of OG at ω14 position from 3'-end.

4.6 References

1. Lindahl, T. *Nature* **1993**, *362*, 709.
2. Frederico, L. A.; Kunkel, T. A.; Shaw, B. R. *Biochemistry* **1990**, *29*, 2532.
3. Duncan, B. K.; Miller, J. H. *Nautre* **1980**, *287*, 560.
4. Imai, K.; Slupphaug, G.; Lee, W.I.; Revy, P.; Nonoyama, S.; Catalan, N.; Yel, L.; Forveille, M.; Kavli, B.; Krokan, H. E.; Ochs, H. D.; Fischer, A.; Durandy, A. *Nature Immunology* **2003**, *4*, 1023.
5. David, S. S.; Williams, S. D. *Chem. Rev.* **1998**, *98*, 1221.
6. Lindahl, T. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3649.
7. Souza-Pinto, N. C.; Harris, C. C.; Bohr, V. A. *Oncogene* **2004**, *23*, 6559.
8. Neddermann, P.; Jiricny, J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1642.
9. Blaisdell, P.; Warner, H. *J. Biol. Chem.* **1983**, *258*, 1603.
10. Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H. *Nat. Nanotechnol.* **2009**, *4*, 265.
11. White, R. J.; Ervin, E. N.; Yang, T.; Chen, X.; Daniel, S.; Cremer, P. S.; White, H. S. *J. Am. Chem. Soc.* **2007**, *129*, 11766-11775.
12. Kasianowicz, J.; Brandin, E.; Branton, D.; Deamer, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13770.

13. Sauer-Budge, A. F.; Nyamwanda, J. A.; Lubensky, D. K.; Branton, D. *Phys. Rev. Lett.* **2003**, *90*, 238101.

14. Schibel, A. E. P.; Fleming, A. M.; Jin, Q.; An, N.; Liu, J.; Blakemore, C. P.; White, H. S.; Burrows, C. J. *J. Am. Chem. Soc.* **2011**, *133*, 14778-14784.

15. Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouax, J. E. *Science* **1996**, *274*, 1859–1865.

16. Jiang, Y. L.; Ichikawa, Y.; Song, F.; Stivers, J. T. *Biochemistry* **2003**, 42, 1922.

17. Fromme, J. C.; Banerjee, A.; Verdine, G. L. *Curr. Opin. Struct. Biol.* **2004**, 14, 43.

18. Zhang, B.; Galusha, J.; Shiozawa, P. G.; Wang, G.; Bergren, A. J.; Jones, R. M.; White, R. J.; Ervin, E. N.; Cauley, C.; White, H. S. *Anal. Chem.* **2007**, *79*, 4778–4787.

19. White, R. J.; Ervin, E. N.; Yang, T.; Chen, X.; Daniel, S.; Cremer, P. S.; White, H. S. *J. Am. Chem. Soc.* **2007**, *129*, 11766.

20. Schibel, A. E. P.; Heider, E. C.; Harris, J. M.; White, H. S. *J. Am. Chem. Soc.* **2011**, *133*, 7810–7815.

21. Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. *J. Am. Chem. Soc.* **2012**, *134*, 11006.

22. Jin, Q.; Fleming, A. M.; Ding, Y.; Burrows, C. J.; White, H. S. manuscript in preparation.

23. Schibel, A. E. P.; An. N.; Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. *J. Am. Chem. Soc.* **2010**, *132*, 17992.

24. Stoddart, D.; Heron, A. J.; Mikhailova, E.; Maglia, G.; Bayley, H. *Proc. Natl. Acad. Sci.* **2009**, *106*, 7702.

25. Vercoutere, W.; Winters-Hilt, S.; Olsen, H.; Deamer, D.; Haussler, D.; Akeson, M. *Nautre Biotechnol.* **2001**, *19*, 248.

26. Stoddart, D.; Maglia, G.; Mikhailova, E.; Heron, A. J.; Bayley, H. *Angew. Chem. Int. Ed.* **2009**, *48*, 1.

27. Stoddart, D.; Heron, A. J.; Klingelhoefer, J.; Mikhailova, E.; Maglia, G.; Bayley, H. *Nano Lett.* **2010**, *10*, 3633.

CHAPTER 5

CONCLUSIONS

In this dissertation, nanopores have been used in biophysical studies to investigate the kinetics of dsDNA strand dissociation (Chapters 2 and 3) as well as in enzymatic assays to monitor the UDG activity (Chapter 4). As a single-molecule manipulation method, the nanopore approach has been emerging as a label-free and modification-free technique to study the dissociation of bio-relevant complexes. In real cells, force-induced strand dissociation mediated by enzymes occurs universally. The nanopore provides a promising tool to mimic enzymatic force using electrical force to trigger dissociation of the biomolecular complexes. Just like an enzyme functioning on a DNA strand on a single-molecule basis, a nanopore offers a single-molecule approach to study the biomolecular complexes. Nanopores will undoubtedly be employed in a broader range of biomolecular complexes, such as enzyme or protein-bound DNA, dsDNA and hairpins, DNA-RNA duplexes, DNA-drug complexes and DNA secondary structures.

Though α-HL has been used as a stochastic sensor to characterize DNA for two decades, its sensing capability has only been studied in the central constriction and β-barrel sections. The new sensing zone for dsDNA in the latch

region was discussed in Chapter 4. This discovery embarks on a new direction for DNA detection in α-HL by highlighting a previously unrecognized region of protein sensor. In Chapter 4, the detection was performed on abasic sites in duplex context at 150 mM KCl in wild type α-HL. It would be interesting to explore the potential of the latch region in detecting other bases or lesions, or at a wider range of experimental conditions. Protein mutagenesis that alters the amino acids at the latch region can change the size of the channel or the interaction between the channel wall and DNA, thus potentially useful to enhance the detection sensitivity. Noticing that the new sensing zone in wild type α-HL spans over five nucleotides, it would be exciting to examine the options of mutant proteins in order to sharpen the sensing zone to one nucleotide. The new sensing zone also provides insight into sequencing if the dsDNA can be slowed down and pass through the latch region.

As demonstrated in Chapter 4, nanopores can be used as an analytical tool in assays to characterize damaged DNA. The fast readout speed makes it possible to scan all of the DNA nucleotides from a single cell in a day using nanopore arrays. To move in the direction of biological sampling, increasing the detection sensitivity to a bio-relevant level is challenging as the damaged DNA is very unpopulated compared to normal DNA in *vivo.* (One damaged DNA cccurs per 10,000 normal ones. Reference: Loeb, L. A.; Harris, C. C. Cancer Res. 2008, 68, 6863-6872.) Thus, a method similar to PCR is necessary to amplify a few damaged DNA strands to millions of copies. In addition, to handle large quantities of DNA in biological samples, high throughout nanopore platforms

need to be improved to simultaneously keep each nanopore under working conditions considering the problem of pore clogging as well as short lifetime of lipid bilayers and protein channels. Automatic systems that control individual nanopores can be programed to clean the pore, build the bilayer, obtain a protein channel and repeat these procedures when the pore misbehaves. It will be interesting to see what nanopore devices, biological pores or solid-state pores (Section 1.1.1), will reach commercialization in future for nanopore bioassays. Regardless of the devices, auto-correction that ensures the function of individual nanopores is necessary for successful nanopore arrays.