Regulating DNA Transport through Solid-State Nanopores to Study Non-Canonical DNA Structures and Analyze Gene Synthesis Reactions

by Spencer Carson

B.S. in Physics and Mathematics, Gordon College
M.S. in Physics, Northeastern University

A dissertation submitted to

The Faculty of
the College of Science of
Northeastern University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

February 1, 2016

Dissertation directed by

Meni Wanunu
Professor of Physics and Chemistry and Chemical Biology
Dedication

to Sara, for being my constant
Acknowledgments

First and foremost, I would like to acknowledge Dr. Meni Wanunu, my dissertation advisor, for his continuous guidance and encouragement over the past four years. I arrived at Northeastern as a graduate student not knowing exactly what I would research for my dissertation, and Meni made a serious commitment to helping me learn many new concepts and techniques in biophysics throughout my time in lab. His strong passion for research most certainly rubbed off on me and will stay with me in my future pursuits. I will always be grateful for the time he invested passing along his knowledge to me. I additionally thank my thesis committee members Mark Williams, J. Timothy Sage, and Armen Stepanyants for their willingness to serve on my committee.

I would like to acknowledge Aleksei Aksimentiev and James Wilson for the MD simulations they conducted for this dissertation and for their dedication and perseverance in making the many adjustments and changes I often needed. I thank Murugappan Muthukumar for sharing his knowledge of polymer physics, which influenced me greatly in my graduate career. Doc Edge is a colleague who was vital in writing and troubleshooting the bootstrapping codes I used in many of my studies. Carlos Aguilar, a collaborator from MIT Lincoln Laboratories, put us on the path of studying gene synthesis reactions using nanopores and without his guidance this study would have been impossible to conduct. Peter Weigele, a collaborator at New England Biolabs was intimately involved in the epigenetic portion of my dissertation and prepared all the samples we needed for those experiments. Aaron Fleming, a postdoc at University of Utah, was generous enough to prepare samples needed for our G-quadruplex project and answered the
many questions that arose during experimentation. Andrey Ivankin and Mirela Mustata, postdocs when I first joined the lab, deserve thanks for providing me with advice and guidance early in my doctoral research. I thank the members of the Wanunu Lab (past and present) for their constant friendship and collaboration, and for making our lab such a fun environment to work in every day: Joseph Larkin, Robert Henley, Pradeep Waduge, Bedanga Sapkota, Benjamin Cressiot, Vivek Jadhav, Daniel Stoloff, Martin Langecker, Will Benoit, Brian Kamp, Ana Vazquez Pagan, Michael Toscano, Moroke Igue, and last but not least Baram Sosis. Specifically, I thank Joe Larkin for being a great resource to me whenever I had questions about virtually anything and for his enthusiasm for discussing Boston sports with me. I also want to specifically acknowledge Robert Henley for all the work he put into our analysis software and proofreading this dissertation, and for drafting Peyton Manning in fantasy football (yeah, sorry about that). I will certainly miss the conversations (or debates) the three of us had ranging from the definition of irony to whether champ was there.

Lastly, I would like to acknowledge my family and friends who were such a strong support to me during my entire doctoral study. The friends that helped me through this doctorate are too numerous to list here, but I thank all of them for being moral support and an oftentimes necessary diversion from my studies. I want to acknowledge my “newest” siblings Katie, Maureen, Patrick, and Tim for always taking an interest in my research and for being a source of relaxation when needed. I thank Mary Anne and Tim for being enthusiastic about my research and for supporting me like a son through my entire doctorate. I acknowledge my sister Mikala for her constant support in all my academic pursuits and for sharing my passion for Lost and Roger Federer. My parents, who nurtured me through my entire academic life, constantly inspired me to pursue my passions in life. They constantly demonstrate to me the rewards of
hard, honest work and I could not ask for better role models as parents. Finally, my wife and best friend Sara deserves many thanks for enduring the long hours and late nights that a doctorate requires, and for sticking by my side and supporting me through all of it. I most certainly could not have completed this dissertation without her constant support and encouragement.
Abstract of Dissertation

Voltage-driven transport of double-stranded DNA through nanopores holds much potential for applications in quantitative molecular biology and biotechnology, yet the microscopic details of translocation have proven challenging to decipher. Earlier experiments showed strong dependence of transport kinetics on pore size: fast regular transport in large pores (> 5 nm diameter), and slower yet heterogeneous transport time distributions in sub-5 nm pores, which imply a large positional uncertainty of the DNA in the pore as a function of the translocation time. In this dissertation, we show that this anomalous transport is the result of DNA self-interaction, a phenomenon which is strictly pore-diameter dependent. We identify a regime in which DNA transport is regular, producing narrow and well-behaved dwell time distributions that fit a simple drift-diffusion theory. This observation of smooth DNA translocation is then used to study the effect of epigenetic modifications on DNA transport dynamics and to analyze gene synthesis reactions. Additionally, a preliminary study is presented reporting on the detection of G-quadruplex structures using solid-state nanopores.
# Table of Contents

Dedication ...................................................................................................................................... ii

Acknowledgments ........................................................................................................................ iii

Abstract of Dissertation ............................................................................................................... vi

Table of Contents ........................................................................................................................ vii

List of Figures ................................................................................................................................ x

List of Tables ................................................................................................................................... xii

Chapter 1 – Introduction ............................................................................................................. 1

Chapter 2 – Smooth DNA Transport Through a Narrowed Pore Geometry ............................... 6

2.1 DNA translocation through a small nanopore ............................................................................. 7

2.2 Voltage fine-tunes DNA transport kinetics through small pores ................................................ 8

2.3 Impact of pore diameter on DNA translocation ......................................................................... 11

2.4 Mechanism for smooth DNA transport in small pores ............................................................ 16

2.5 Smooth transport of long DNA ................................................................................................ 18

2.6 Single-pulse discrimination of short DNA fragments ................................................................. 21

2.7 Discussion .................................................................................................................................. 22

2.8 Materials and Methods ............................................................................................................. 24

Chapter 3 – Hydroxymethyluracil Modifications Enhance the Flexibility and
Hydrophilicity of Double-Stranded DNA .................................................................................... 27

3.1 Introduction ............................................................................................................................... 28

3.2 Transport of hmU-containing DNA variants through small nanopores .................................. 29

3.3 Melting curve analysis of DNA variants .................................................................................... 32
3.4 Molecular dynamics simulations reveal hmU increases DNA flexibility .................. 33
3.5 Discussion .................................................................................................................. 37
3.6 Materials and Methods ............................................................................................ 38

Chapter 4 – Direct Analysis of Gene Synthesis Reactions Using Solid-State Nanopores .... 43
4.1 Introduction .............................................................................................................. 44
4.2 Rapid detection of small DNA defects ................................................................. 46
4.3 Differentiation of structural defects ................................................................. 50
4.4 Discrimination between correctly assembled gene products and misassembled gene products ........................................................................................................................................................................... 53
4.5 Discussion .............................................................................................................. 57
4.6 Materials and Methods ............................................................................................ 58

Chapter 5 – Detection of G-Quadruplex Structures within Gene Promoter Regions ....... 61
5.1 Introduction .............................................................................................................. 62
5.2 Detection of ILPR G-quadruplex structures in nanopores ........................................ 64
5.3 Effect of oxidation on VEGF promoter GQ formation ........................................... 66
5.4 Materials and Methods ............................................................................................ 69

Chapter 6 – Conclusions and Future Outlook ............................................................... 70

Appendices .................................................................................................................... 73

Appendix A - Supporting Material for Chapter 2 ........................................................... 73
A.1 Materials and Methods ............................................................................................ 73
A.2 Effect of low-pass filter frequencies on event detection ........................................ 75
A.3 Time stability and pore-to-pore reproducibility of DNA translocation ................. 77
A.4 Determination of pore diameter and thickness based on open and blocked pore ionic current ($I_o$ and $I_b$) ......................................................................................................................... 78
List of Figures

Figure 2.1 - Single-molecule DNA electrophoresis ........................................................................................................... 8
Figure 2.2 - Voltage-tuned DNA transport ..................................................................................................................... 10
Figure 2.3 - The influence of pore diameter on DNA transport time ............................................................................... 12
Figure 2.4 - Dependence of \( v \) and \( D \) with respect to pore diameter \( d \) ................................................................. 14
Figure 2.5 - Contour plots of \( \log V \) from finite element simulations for nanopores of two diameters show the contrast in residual capture field when DNA threads through the pore ............................................................................. 17
Figure 2.6 - Scaling of transport time \( t_d \) as a function of DNA length \( N \) ........................................................................ 19
Figure 2.7 - Single-pulse discrimination between two DNA fragments ............................................................................ 21
Figure 3.1 - Nanopore detection of thymine variants in DNA ........................................................................................ 30
Figure 3.2 - Translocation of thymine-modified DNA through a 2.8 nm diameter pore .................................................. 31
Figure 3.3 - Melting curves of modified DNA variants for different buffer conditions ................................................ 33
Figure 3.4 - Results of molecular dynamics simulations characterizing structural fluctuations of the DNA variants .................................................................................................................................................................. 34
Figure 3.5 - MD simulations of ion atmosphere in a solid-state nanopore occupied by DNA variants ......................... 37
Figure 4.1 - DNA translocation through a solid-state nanopore .................................................................................... 46
Figure 4.2 - Single mismatch detection using nanopores ............................................................................................ 48
Figure 4.3 - Gene synthesis misassemblies have unique translocation profiles .......................................................... 51
Figure 4.4 - Nanopore-based monitoring of gene synthesis products for the HIV-Protease gene .................................... 54
Figure 5.1 – G-quadruplex geometry and cation dependencies ....................................................................................... 63
Figure 5.2 – Nanopores allow the detection of G-quadruplexes in duplex DNA ............................................................ 65
Figure 5.3 - Oxidation of guanine disrupts G-quadruplex formation ........................................................................... 68
Figure A1 - Effect of signal bandwidth on masking DNA collision spikes .................................................................... 76
Figure A2 - A representative, continuous 40-second current trace for 500 bp transport at \( V = 200 \) mV through a pore of \( d = 3.0 \) nm .................................................................................................................... 77
Figure A3 - Contour plots of 500 bp DNA translocation at \( V = 200 \) mV for three different pores .............................. 78
Figure A4 - Finite element simulation geometry of a nanopore .................................................................................. 81
Figure A5 - An agarose gel of DNA lengths 1 kbp – 20 kbp

Figure A6 - Continuous current traces for 100 bp, 1 kbp, 3.5 kbp, 10 kbp and 20 kbp DNA for nanopores with $d = 2.8$-$3.0$ nm

Figure A7 - Concatenated current traces of consecutive events for the data shown in Figure A6 overlaid with analysis fits

Figure A8 - Scatter plots of $\Delta I/I_o$ vs. $t_d$ for 11 DNA fragments through different nanopores

Figure B1 - Sequence and restriction map of the amplicon used

Figure B2 - Nanopore translocation dataset for thymine-modified DNA samples

Figure B3 - The effect of point hmU modifications on simulated DNA

Figure B4 - The effect of point phmU modifications on simulated DNA flexibility

Figure C1 - PAGE reveals high affinity binding of MutS to DNA mismatches

Figure C2 - Power spectral density (PSD) plots based on 500 ms of open pore current data

Figure C3 - Gel and nanopore analysis of common gene synthesis misassemblies

Figure C4 - Inter-event time histograms of gene synthesis misassemblies

Figure C5 - Oligo pool information for HIVPr gene assembly
List of Tables

Table 3.1 - Experimental and Simulated Translocation Data for Thymine-Modified dsDNA ................. 32
Table 4.1 - Capture Rates of Gene Synthesis Misassemblies ................................................................. 53
Table A1 - Dwell Time vs. DNA Length Data ......................................................................................... 87
Chapter 1 – Introduction

Cellular DNA is generally found in the nucleus in the form of a long double-stranded DNA (dsDNA) helix. During transcription and other DNA metabolic processes, specialized proteins can slide along the DNA contour in search of a specific sequence, chemical modifications, or chemical damage. Recently, nanopores in synthetic materials and engineered protein channels have emerged as single-molecule tools for similarly identifying sequence and other structural features along DNA contours.\(^1-5\) In this method, DNA molecules are electrophoretically pulled into an electrolyte-filled nanopore by applying a small voltage bias (\(i.e., < 500 \text{ mV}\)) across it. DNA entry and passage through the pore are detected by monitoring the ion current signal across the pore. Temporal fluctuations in the signal can report on the presence of epigenetic modifications,\(^6-11\) alteration in sequence,\(^12-16\) length variation,\(^17, 18\) secondary structure,\(^18, 19\) and conformation\(^20-22\) of the DNA molecule. Prospects for nanopores as DNA sequence scanners, high-resolution DNA sequence mappers, and epigenetic modification detectors have fueled research that aims to understand the details of DNA translocation and its limitations. In comparison to existing technologies, scanning DNA fragments at high-throughput is potentially more attractive than optical DNA stretching and mapping\(^23-25\) because: 1) electronic measurement devices can be more easily miniaturized, 2) electronic measurements can be made at MHz bandwidths, allowing faster processes to be probed, and 3) the resolution of nanopore measurements is ideally dictated by its geometry (2-5 nm), whereas optical methods are constrained by undesirable photophysical features of dye molecules and optical diffraction limits (~100s of nm).
Despite ample promise, two major shortcomings of nanopores with respect to dsDNA analysis have been: 1) Mean transport speeds are 10-100 times faster than required for statistical averaging of ion current data from short DNA regions inside the pore,\textsuperscript{5} and 2) DNA position vs. time in the pore is not well-known or otherwise regulated.\textsuperscript{3, 26, 27} While regulation has been achieved for single-stranded DNA using enzymes as molecular stepper motors,\textsuperscript{9, 10, 28-32} regulated motion of dsDNA has yet to be demonstrated. Voltage-driven dsDNA translocation through 5-15 nm diameter pores in thin (20-50 nm) solid-state materials proceeds with \textit{mean} velocities of 10-100 ns/bp, and often in large pores multiple DNA strands enter simultaneously,\textsuperscript{18, 20, 33, 34} which complicates single-file readout of information that is encoded in the linear sequence. Proteins such as RecA from \textit{E. coli} form filaments around the DNA that slows DNA transport and prevents its folding,\textsuperscript{27, 35, 36} although this approach inherently masks chemical information contained within the DNA, such as the presence of DNA chemical modifications\textsuperscript{7, 37} or small bound drug/reporter molecules.\textsuperscript{38-41} Explorations of the effects of parameters such as the electrolyte viscosity,\textsuperscript{42, 43} salt type,\textsuperscript{44, 45} membrane material,\textsuperscript{46, 47} applied pressure,\textsuperscript{48, 49} and chemical composition inside\textsuperscript{50-53} and outside\textsuperscript{54} the pore have yielded only moderate DNA retardation factors.

In contrast, the use of smaller diameter pores (< 5 nm) results in \textit{mean} DNA velocities that are reduced easily by ~2 orders of magnitude\textsuperscript{5, 55-57} while simultaneously preventing folded DNA transport. However, the statistics of DNA transport times through sub-5 nm pores point to a highly complex process: a pure single-length DNA fragment can produce multiple event populations that are characterized by broad distributions of dwell times and blockade amplitudes.\textsuperscript{20, 56-59} Based on their dependence on various parameters, these distinct sub-populations have been associated with various processes such as non-transport DNA collisions.
with the pore and translocation governed by strong interactions. The broad and overlapping nature of these sub-populations is a grand impediment of DNA mapping applications as it implies a greatly smeared DNA positional trajectory during translocation.

In Chapter 2 of this dissertation, a systematic study is presented that pinpoints and overcomes the source of irregular DNA transport through nanopores. Our findings have led us to conclude that these anomalous transport dynamics are due to a residual electric field, which causes self-interaction between the DNA coil outside of the pore and the pore mouth, leading to stop-and-go motion during DNA translocation. These interactions are effectively extinguished by restricting the pore diameter, leading to reproducible transport time distributions with unprecedentedly low scatter. Molecular dynamics (MD) simulations of DNA diffusion within the pore constraint and quantitative fits of our experimental data support a greatly reduced DNA diffusion in sub-3 nm diameter pores. Mean transport times are then related to DNA length in the range of 35 – 20,000 bp through a single superlinear power law, which when compared to prior studies appears to be independent of pore diameter. Finally, we demonstrate discrimination between two short DNA fragments in a mixture from a single electrical pulse with >98% accuracy.

Chapter 3 of this dissertation builds upon these initial findings to explore the effects of DNA epigenetic modifications in the form of a thymine to 5-hydroxymethyluracil (hmU) oxidation, a type of DNA damage correlated with cancer and other chronic diseases. We discovered that hmU modifications cause dsDNA to be more flexible and hydrophilic, leading to faster transport through small nanopores in short DNA fragments (i.e., ~100 bp).
experimental results and hypotheses were confirmed by free solution and nanopore-constrained MD simulations of each DNA variant.

Chapter 4 discusses the differentiation of DNA structures commonly found in gene synthesis reactions using solid-state nanopores. Initially, we discovered that our nanopore system has the sensitivity to detect a single base mismatch in short dsDNA molecules (i.e., 70 bp) by binding a DNA repair protein. Next, we compared the nanopore transport dynamics of three common DNA misassemblies in gene synthesis reactions to a properly assembled dsDNA of the same length and found that each molecule had a distinct translocation “fingerprint”. Finally, we conducted nanopore experiments with gene synthesis reaction samples before and after PCR purification and successfully detected the emergence of an event population corresponding to the correct DNA gene assembly in the purified sample.

Chapter 5 describes preliminary research using solid-state nanopores for the detection of G-quadruplexes in duplex (genomic form) DNA. G-quadruplexes are four-stranded DNA structures that form in guanine-rich regions of the human genome. The G-quadruplex (GQ) has become a hot topic in cancer genomics because many putative GQ sequences are located in telomeres (i.e., ends of chromosomes) and promoter regions of notable genes, many of which have implications in cancer. Due to their prevalence in promoter regions, GQs have been directly linked to gene expression, which has made them a potential drug target to prevent the over- or underexpression of particular genes. We aim to study the promoter regions of two such genes using nanopores to detect the presence of GQ-containing duplex DNA, and explore the effect of DNA damage on GQ formation in one of these genes.
Brief conclusions and an outlook on how this research could be expanded upon in future experiments are presented in Chapter 6.
Chapter 2 – Smooth DNA Transport Through a Narrowed Pore Geometry

2.1 DNA translocation through a small nanopore

Figure 2.1a illustrates a molecular model of a ~150 bp dsDNA fragment being voltage-driven through an ultrathin pore with a diameter of \( d \sim 3 \) nm. The length of dsDNA shown is one persistence length, intentionally chosen to illustrate the stiff nature of the polymer with respect to the pore dimensions. DNA introduction into the \( cis \) (top) chamber and application of positive bias to the \( trans \) (bottom) chamber results in capture and electrophoresis of individual DNA molecules across the pore. Transport of a DNA molecule produces a transient downward spike in the electrical signal, a result of the temporary restriction in ion flux through the pore. In Fig. 2.1b we show two continuous traces obtained when 60 nM of 500 bp DNA was added to the \( cis \) chamber and driven using a 200 mV voltage through a 2.9 nm pore and a 6 nm pore. The all-point histograms on the right reveal for each pore two predominant current levels, the higher corresponding to open pore (\( I_o \)) levels and the lower corresponding to DNA-occupied levels. Post-acquisition analysis of the traces (OpenNanopore, EPFL\textsuperscript{66}) identifies the events as rectangular pulses, a sample of which are shown for both pores in Fig. 2.1c. Scatter plots of fractional current blockade \( \Delta I/I_o \) vs. dwell time \( t_d \) are displayed in Fig. 2.1d, showing a broad population centered at \( t_d \sim 25 \) µs and \( \Delta I/I_o \sim 0.12 \) for the 6 nm pore (\( n_{\text{total}} = 1,798 \)), and \( t_d \sim 890 \) µs and \( \Delta I/I_o \sim 0.56 \) for the 2.9 nm pore (\( n_{\text{total}} = 2,593 \)). We note that we can only observe dwell times and current blockades larger than our detection limits (dashed red lines in the plot). Further, only when reducing the bandwidth of our signal do we observe DNA collisions with the pore (see Fig. A1), the vast majority of which produce short-lived, low amplitude pulses. The low scatter in DNA transport times in the smaller of the two pores, which in the course of our experiments has been reproduced for different DNA fragments using over 50 pores, has motivated our study to pinpoint the reasons for smooth DNA transport through pores in the \( d = \)
2.8 – 3.0 nm regime (see Figs. A2 and A3 for information on pore stability and pore-to-pore reproducibility).

### 2.2 Voltage fine-tunes DNA transport kinetics through small pores

DNA is a uniformly charged polymer, and therefore its transport kinetics is expected to depend on applied bias. We have chosen to conduct our systematic study using 500 bp dsDNA because it is in the rod-to-coil regime, *i.e.*, its 3+ persistence lengths are sufficiently long to frequently adopt 360° loop configurations at equilibrium (j-factor of ~10 nM).\(^6^7\) In Fig. 2.2a we present dwell time distributions for 500 bp DNA transport through a 3.0 nm diameter pore in the 200 – 350 mV voltage range. The dwell time distributions are asymmetric in shape, and are characterized by sharply increasing peaks at early times followed by broader decays at longer times. The distribution shapes were fit to first-passage time distributions obtained...
from the one-dimensional (1D) Fokker-Planck equation,\textsuperscript{68,69}

\[ f(t_d) = A \cdot b / (4\pi Dt_d)^{3/2} \exp[-(b - vt_d)^2 / (4Dt_d)] \] \hspace{1cm} (1)

where \( b \) is the trajectory length of the molecule, \( D \) is the diffusion coefficient, \( v \) is the drift velocity, and \( A \) is a normalization constant. The model of Eq. 1 describes the 1D motion of a particle that starts at the origin (\( x = 0 \)) of a semi-infinite capillary with an absorbing boundary located at \( x = b \). Given that the complete contour length \( L_C \) of DNA is longer than the pore length \( b_{\text{eff}} \), we set the absorbing boundary to be at \( b = b_{\text{eff}} + L_C \).\textsuperscript{43,69,70} Phenomenological estimates of \( b_{\text{eff}} \) for each pore were based on \( I_o \) and \( \Delta I \) values measured during DNA translocation\textsuperscript{71,72} (see Section A.4), whereas \( L_C \) values were calculated from \( L_C = 0.34N \) nm, where \( N \) is the number of base pairs. Best fits to Eq. 1 for each distribution are shown in Fig. 2.2a (shaded curves). The percentages of events that fit the model in Eq. 1 lie in the range 84 – 94% (see Section A.5), where the remaining events outside of the fit are mostly scattered at longer dwell times.

In Fig. 2.2b we plot the fractional current blockade (\( \Delta I/I_o \)) as a function of voltage. We observe a moderate increase in \( \Delta I/I_o \) with increasing voltage (6% change from 200 – 350 mV), an effect that was observed in a recent study by the Hall group.\textsuperscript{59} Drift velocities (\( v \)) obtained from the fits are shown in Fig. 2.2c. Displayed standard errors for \( v \) (and later for \( D \)) were calculated by resampling our data for 10,000 iterations using a standard bootstrapping procedure (see Section A.1.3). As previously observed using larger pores\textsuperscript{42,73} and for free-solution DNA electrophoresis,\textsuperscript{74} we find that \( v \) increases with voltage. However, \( v \) does not depend linearly on voltage and cannot be extrapolated to the origin, suggesting a voltage-dependent electrophoretic
mobility ($\mu$), i.e., $\mu = f(V)$. Mean values of $\mu$ obtained in our voltage experiments are on the order of $\sim 10^{-7}$ cm$^2$ V$^{-1}$ s$^{-1}$, three orders of magnitude smaller than bulk $\mu$ values for DNA in free-solution electrophoresis ($\sim 10^{-4}$ cm$^2$ V$^{-1}$ s$^{-1}$). The obtained values of $v$ yield mean velocities that correspond to one DNA base pair moving through the pore per $\mu$s (dashed red line). We highlight that these are mean values because a constant DNA velocity through a nanopore is unlikely, and further was not proven by this experiment.

The diffusion coefficients ($D$) of the DNA extracted using Eq. 1 appear to slightly depend
on voltage, increasing from 0.5 to 1.5 nm²/µs (see Fig. 2.2d). A similar trend was recently observed for translocation of stiff fd virus filaments.\textsuperscript{76} Moreover, as for a rod-like stiff fd filament ($L_C < p$, where $p$ is a persistence length), values for $D$ were comparable to the axial component of their bulk translational diffusion coefficients. In the case of a semi-flexible DNA chain, we argue that a relevant length scale for correlated axial DNA motion is one-half of a persistence length, as opposed to the full DNA contour length. Using reference bulk $D_o$ values for 75 bp (45 nm²µs\textsuperscript{-1}),\textsuperscript{77-82} we find that the axial component of $D$ for our experiment is $1.5D_o$, or 68 nm²/µs (see Section A.6 for details). Fig. 2.2d appears to contradict our hypothesis above, since for all voltages tested we find that $D$ is at least 40 times smaller. In the next section, we will show that the restricted pore geometry is responsible for the greatly reduced $D$ values of DNA, as we recently found for protein electrophoresis in the regime $d_{protein} \sim d_{pore}$.\textsuperscript{83} Finally, in Fig. 2.2e we show the dependence of the ratio $v/D$ on voltage for this study. Apart from scatter that is mainly a result of the dispersity in measured $D$ values, we find that the ratio $v/D$ is independent of voltage (0.43 ± 0.13 nm\textsuperscript{-1}). This suggests that an intricate relationship between $v$ and $D$ governs our observed superlinear dependence of $v$ on voltage, one that is a result of tight molecular confinement and/or interactions within the pore volume.

### 2.3 Impact of pore diameter on DNA translocation

In Fig. 2.3a-e we present five representative dwell time distributions for 500 bp DNA transport through 2.6 – 6 nm diameter pores. Additionally, inset to each distribution we show contour plots of fractional current blockades vs. log dwell times, all to the same scale for comparison.
FIGURE 2.3 - The influence of pore diameter on DNA transport time. (a-e) Dwell-time distributions with drift-diffusion fits included (shaded purple curves) for 500 bp DNA translocation through five different pores of diameter $d$ ranging from 2.6 nm to 6 nm. Our model fits the data by the designated percentages with the shown 2.9 nm pore matching the best, an 85% fit. Pores with $d > 3.4$ nm have distributions with noticeably long tails that lie outside our model curves, as denoted by the red brackets, indicating translocation events with erratic transport through the pore. Insets: Heat maps of fractional current blockade vs. the log of the dwell time for each pore diameter. The color scale bars displayed are in log scale. (f) By plotting the fit percentages for 14 different pores as a function of $d$, we see that the optimal pore size for regular DNA transport lies between $d = 2.8 – 3.0$ nm (green shaded region). Red dotted line is a guideline for viewing the trend in $d$.

First, the inset contour plots reveal a striking 500-fold shift to longer dwell times upon reducing the pore diameter from 6.0 to 2.6 nm (~10 µs to ~5 ms, respectively). In addition to the longer observed dwell times with decreased pore diameters, their relative scatter (i.e., $<t_d>/\sigma_{td}$) decrease. Further, upon fitting our data to Eq. 1, as shown by the purple shaded curves, we find for the 3.4 – 6.0 nm pores a secondary population of long-lived dwell times, indicated by red asterisks in the figure. These doubly populated dwell time distributions have been previously
observed in larger pores of various materials, and suggest a different translocation mechanism than the main population. Lastly, in Fig. 2.3f we plot the fraction of events that fit the model in Eq. 1 for different pore diameters, as obtained by integrating the experimental distributions and the fits. To arrive at these values we converted the dwell time histograms and fits into cumulative distribution functions (CDFs), and then estimated the percent of excess (i.e., non-fitting) events that lie above the aligned experimental and theoretical CDFs. By dividing the area corresponding to this dwell time by the total area of the histogram curve, the percentage of events that agree with our model is determined (see Section A.5 for further discussion). The guideline drawn in Fig. 2.3f (dashed red line) shows a plateau of events that fit Eq. 1 for $d = 2.8 – 3.0 \text{ nm (shaded in green)}$, whereas a regular emergence of a second process is observed for larger pores. In the optimum regime we can therefore say that the predominant transport mechanism is a smooth one-dimensional first-passage process. Interestingly, for $d = 2.6 – 2.7 \text{ nm (n = 7)}$ we find that a significantly lower fraction of events fit the model than for $d = 2.8 – 2.9 \text{ nm (n = 4)}$, presumably due to overwhelming DNA/pore interactions.

Next, in Fig. 2.4a we present a compiled set of \( v \) and \( D \) values for 500 bp DNA transport as a function of pore diameter (\( n = 14 \), effective thickness \( b_{\text{eff}} = 7 – 10 \text{ nm} \)). The trends for both parameters have a striking resemblance: as pore diameter increases from 2.6 – 3.0 nm we observe a drastic increase in both \( v \) and \( D \), whereas for \( d > 3 \text{ nm} \) both parameters approach asymptotic limits (\( v_{\text{max}} \approx 20 \text{ nm/\mu s}; D_{\text{max}} \approx 80 \text{ nm}^2/\mu \text{s} \)). Under our experimental conditions, we could not accurately detect 500 bp DNA through \( d > 6 \text{ nm} \) pores. To explain the convergence of \( v \) and \( D \) values for larger pores we turn to the simplistic model of the DNA drift-diffusion problem. By fitting our dwell-time distributions to Eq. 1 we seek a solution of first-passage times for rod-like DNA segments to traverse a full DNA contour length through the pore (illustrated in
As argued earlier in the manuscript, for a rod-like DNA length of \sim 0.5 \mu m (or 75 bp) we expect an axial diffusion coefficient of \sim 68 \text{nm}^2/\mu s. As shown by an asterisk in Fig. 2.4a, this value is in fair agreement with our observed values of $D$ for $d > 3$ nm (40-100 \text{nm}^2/\mu s).

In the $d < 3$ nm regime, we find a trend that implies strongly dampened DNA diffusion due to confinement. Effects of pore confinement on $D$ are well known,\textsuperscript{26, 87, 88} and have been recently observed for protein transport in confining pores.\textsuperscript{83} We independently assessed the
impact of confinement on the diffusion coefficient of dsDNA by employing all-atom molecular dynamics (MD) simulations. Specifically, we constructed five atomic-scale models of the experimental system, each containing a nanopore in a 7 nm-thick silicon nitride membrane, a 36 bp fragment of dsDNA, and an electrolyte solution (see Fig. 2.4b). Each nanopore had an hourglass shape and the nanopore diameters were 2.5, 2.7, 2.9, 3.5, or 4.5 nm. We simulated mechanical pulling of dsDNA through the nanopore while measuring the average force $F$ required to maintain the prescribed translocation velocity $v$ (see Section 2.8 for details). The measured force-velocity dependence was used to estimate the diffusion coefficient via the Einstein relation: $D = k_BT v / F$. The results of these simulations clearly have a striking, nearly quantitative resemblance to our experimental observations: Simulated $D$ values for pores with $d > 3.5$ nm were quantitatively close to our asymptotic values, whereas in the confined DNA regime systematic decreases in $D$ were observed. Notably, for $d = 2.5$ nm, the diffusion coefficient obtained through the Einstein relation was found to depend on the pulling velocity due to a noticeable friction between the pore wall and the DNA. This is coincidental with our experimental observations, which showed more scatter in $D$ and $v$ values for pores with $d = 2.6 – 2.7$ nm.

While the experimental and simulation-based dependence of $D$ on pore size agrees qualitatively (and perhaps semi-quantitatively), a comparison of experimental $v$ values in the unconfined regime to values based on free-solution electrophoresis measurements yields a very poor agreement: Stellwagen measured for ~75 bp DNA a mobility of $\mu = 3.2 \times 10^{-4}$ cm$^2$ V$^{-1}$ s$^{-1}$, which, given the electric field in our pores ($E \sim 1.8 \times 10^5$ V/cm, corresponding to 200 mV applied across a 8 nm long nanopore), should yield a drift velocity of ~580 nm/μs. This is a factor of ~26 higher than our largest value of $v$ ($d = 6.0$ nm, $v = 22.5 \pm 0.5$ nm/μs). We argue that
this discrepancy is an outcome of three major differences from bulk measurements: First, only a fraction (about one-third) of the ~75 bp rod-like DNA segment (i.e., a half-persistence length of DNA, or 25 nm) is actually driven by the electric field since our pores are ~8 nm thick. Second, in free electrophoresis the electric fields are typically uniform on the scale of a molecule, whereas the large field gradient within a pore cannot act to drive all monomers in the pore equally, further reducing the effective driving force. Finally, dsDNA effective charge is reduced in nanopores to ~0.5 $e$/bp due to electroosmotic effects (one-fourth of formal charge), as determined via single-molecule force measurements.\textsuperscript{90, 91} These combined factors argue for greatly reduced $\mu$ values in nanopores, although a quantitative assessment of the reduction requires models that are beyond the scope of this work.

In the more confined regime of $d < 3$ nm we find a sharp and systematic decrease in $v$ values by up to two orders of magnitude, a trend that remarkably resembles that of $D$. In Fig. 2.4c we compute for the series of 14 pores the ratio $v/D$. Strikingly, as we found to be voltage-independent in Fig. 2.2, the parameter $v/D$ remains fairly diameter independent ($0.15 \pm 0.10 \text{ nm}^{-1}$), despite pore-to-pore variance of up to 70%. Given the increased variance in $v/D$ for sub-3 nm diameter pores, we attribute it to small differences in geometry among the different pores. Therefore, the $v/D$ ratio qualifies as a useful metric when contemplating pore-to-pore variations in experiments.

2.4 Mechanism for smooth DNA transport in small pores

We have shown that DNA dwell time distributions of a 500 bp DNA fragment through pores in the regime $d < 3$ nm appear as single populations, whereas in the regime $d = 3.4 – 6.0$ nm two populations are observed (Fig. 2.3). To explain this observed behavior we point to the
electric field landscape in the pore vicinity during translocation. In Fig. 2.5 we present finite element numerical simulations of the electric potential ($V$) profiles around a $d = 4.0$ nm (Fig. 2.5a-c) and $d = 2.6$ nm pore (Fig. 2.5d-f). Correspondingly, below the open pores we show $V$ profiles for DNA-occupied pores (see Section A.7 for simulation details).

**FIGURE 2.5** - Contour plots of log $V$ from finite element simulations for nanopores of two diameters show the contrast in residual capture field when DNA threads through the pore (simulation conditions: 200 mV, 0.4 M KCl, 21°C). When the nanopore is open (a, d) there is a wider capture radius for the 4 nm pore than for the 2.6 nm pore. Once DNA is threading through the nanopore (b, e) there is a reduction in the electric field, but still enough residual remaining to affect the DNA coil outside the nanopore. Since the DNA occupies a larger percentage of the pore’s volume when the diameter is 2.6 nm, it’s residual field is reduced more significantly (~50%) than in the case of a 4 nm diameter (~14%). The dotted black line in each panel indicates where the electric potential is 1 mV. Insets: TEM images of SiNx nanopores with dsDNA overlaid to demonstrate occupied area during translocation (scale bar is 2 nm). (c, f) These panels illustrate the effect of coil interference in the case of each diameter. Sample dwell time distributions taken from Fig. 2.3b,e are displayed below to show the added longer dwell times (**red shaded region**) that are the result of DNA self-interaction during translocation.

We refer the reader to two key observations: 1) the “external” field above the pore mouth is more pronounced for the larger pore, and 2) DNA threading results in a reduction of this “external” field in both cases, although this diminution is greater for the small pore. The above two observations are trivial from experiments, since larger ion currents are observed for larger pores, and larger fractional blockades are seen for smaller pores, respectively. However, this also demands that the impact of the “external” electric field on the extra-pore DNA segments during
its transport is not equal for different pore diameters. For the 2.6 nm pore much of the external field upon DNA threading is self-extinguished, whereas for the 4.0 nm pore DNA causes only a mild reduction. To orient the reader we show using dashed semicircles the 1 mV equipotential lines for all simulations in the figure. For the 4.0 nm pore, the 1 mV contour line is found at a reduced distance from the pore by 14% upon DNA threading, whereas for the 2.6 nm pore the reduction is 50%. As a DNA molecule travels through a larger pore, an extra-pore DNA region of the currently translocating molecule is more likely to be voltage-driven towards the mouth of the larger pore, causing self-interference that stochastically slows/stops the DNA, as observed by the smeared dwell-time distribution at longer timescales than predicted by the drift-diffusion model (see Fig. 2.5c). In contrast, this self-interference is inhibited in the regime $d = 2.8 – 3.0$ nm, yielding smooth transport (see Fig. 2.5f). We note that the illustrations in Fig. 2.5c,f are not to scale, and the pseudo-loop that the DNA forms requires several persistence lengths. As we have shown here, a 500 bp DNA length is sufficient to yield complicated translocation dynamics unless the pore size is restricted to a sub-3 nm diameter.

2.5 Smooth transport of long DNA

In Fig. 2.6a we show representative sets of concatenated events for different DNA lengths (see Fig. A7 for more traces). Clearly, the events uniformly increase in duration for longer DNA molecules, yet the current blockades remain constant (see Fig. A8) in contrast with a prior report in which DNA length dependence on the conductance blockade fraction was observed for DNA lengths above 2 kbp. In Fig. 2.6b we present log dwell time distributions for transport of 11 DNA lengths in the range 35 – 20,000 bp through nanopores with $d = 2.8 – 3.0$ nm. The distributions are represented as color maps, where increasing color intensity represents more populated bins ($n > 280$ for each length shown, see Table A1). Remarkably, we observe a
uniform increase in the peak dwell time position with increasing DNA lengths, with the exception of DNA lengths in the range 6 – 20 kbp, for which there is a significant scattered trail of events with shorter dwell times. After having confirmed using gel electrophoresis that our long DNA fragments are not contaminated with shorter DNA fragments (see Fig. A5), we believe that transport of DNA fragments longer than 6 kbp often involves other processes that are a result of the coil size (e.g., DNA shearing, DNA escape, etc.). However, since this only occurs for a minor fraction of the events, the analytical power of small pores is not entirely

![Graphs showing the scaling of transport time as a function of DNA length.](image)

**FIGURE 2.6** - Scaling of transport time $t_d$ as a function of DNA length $N$. (a) Sample concatenated traces of consecutive events for 100 bp, 1 kbp, and 10 kbp with included analysis fits (black solid lines). (b) The normalized dwell time histograms for $N = 35$ bp – 20 kbp in 2.9 ± 0.1 nm SiN nanopores (200 mV, pH 7.9, 21°C). Notably, each DNA length less than 6 kbp has only one event population in contrast to past results with small nanopores. (c) A logarithmic plot of $<t_d>$ vs. $N$ for DNA translocation through pores with $d = 2.9 ± 0.1$ nm. By fitting our data using a power law function ($t_d \sim N^{\alpha}$) for 35 bp $< N < 20,000$ bp we extract a power exponent $\alpha = 1.37$. 

19
limited to DNA lengths below 6 kbp. After having extracted peak dwell times from curve fitting the data for each DNA length, we plot the characteristic dwell-time $t_d$ vs. $N$ in the regime $35 < N < 20,000$ bp (see Fig. 2.6c). The uncertainty in dwell time displayed for each data point is the standard error of the mean (see Table A1). Fitting the results to a power law ($i.e., t_d \sim N^\alpha$) we obtain an exponent of $\alpha = 1.37$ with a Pearson’s chi-squared of $R^2 = 0.98$. Power laws such as this have been previously obtained using Monte Carlo simulations$^{97,98}$ as well as Langevin and molecular dynamics simulations$^{99-102}$ for dsDNA transport through synthetic pores, although quite a broad range of power laws has been suggested.$^{103}$ However, the similarity among different reports is striking: experimental studies using $d \sim 10$ nm pores obtained similar values of $\alpha = 1.27^{17}$ and $\alpha = 1.34^{18}$ to our results, and similarly, for $d = 4$ nm, $\alpha = 1.4$ was found for short DNA fragments.$^{56}$ In contrast with the latter cited power law, in our experiments we have found a single power law relationship for the entire DNA length range tested. This universal power law further suggests a single transport process that does not suffer from stalling mechanisms due to DNA coil self-interference. Reducing the pore diameter to the 2.8 – 3.0 nm regime helps to eliminate this adverse and anomalous motion of DNA transport. However, a standing question that remains is: why is a scaling of $\alpha = 1.3 – 1.4$ for dsDNA transport so persistent for such a wide range of nanopore diameters and experimental conditions? While there are two distinct power law regimes that are expected, namely, a scale-free mean velocity ($\alpha = 1$) in the limit of strongly interacting pores$^{104}$ and a parabolic dependence ($\alpha = 2 – 2.5$) in the diffusive regime (bias-free escape),$^{105,106}$ power laws that are pore diameter-independent imply that for voltage-driven dsDNA translocations hydrodynamic interactions play a pronounced role.
2.6 Single-pulse discrimination of short DNA fragments

As a demonstration of the high resolution of our nanopores regarding DNA contour length, we used a pore of \( d = 2.6 \text{ nm} \) and \( b_{\text{eff}} = 6 \text{ nm} \) to differentiate between two short DNA lengths: 100 bp and 500 bp. Upon the addition of a mixture of 100 bp and 500 bp (1:1 ratio, 30 nM of each length) to the cis chamber, it was apparent in real-time that two dwell time populations existed as highlighted by asterisks in Fig. 2.7a. When binning the transport time data, as shown in Fig. 2.7b, we clearly observe two distinct populations for 100 and 500 bp with an overlap of less than 3%. The calculated mean dwell times of 190 \( \mu \text{s} \) and 970 \( \mu \text{s} \) correspond well to the values displayed in Fig. 2.6c for pores with \( d = 2.8 - 3.0 \text{ nm} \), which is due to the use of a thinner pore (6 nm in this case, as compared to 8 nm above). As an additional confirmation of identifying the correct DNA lengths, we ran the same DNA samples individually in pores of identical dimensions to see if we would obtain similar values for dwell time \(<t_d>\). As illustrated in the inset of Fig. 2.7b, we found that 100 bp and 500 bp samples yielded highly distinguishable

![Figure 2.7](image_url)

**FIGURE 2.7** - Single-pulse discrimination between two DNA fragments. (a) A continuous current trace for a 60 nM equimolar mixture of 100 bp and 500 bp in a small, thin nanopore (i.e., \( d = 2.6 \text{ nm}, b_{\text{eff}} = 6 \text{ nm} \)) shows a clear distinction between the two different lengths as denoted by the red and blue asterisks (100 bp and 500 bp). (b) Dwell time distributions of translocation data from panel (a) show distinct peaks for 100 bp and 500 bp. Inset: Scatter plots of the fractional current blockade vs. dwell time for the translocation of 100 bp (red) and 500 bp (blue) separately and together in an equimolar mixture (black) in pores of comparable size (\( d = 2.6 \text{ nm} \)). Using pores of this size, we are able distinguish between 100 bp and 500 bp events with an accuracy of 98.4% and 97.6% respectively when dividing the populations at 450 \( \mu \text{s} \) (pink dotted line). In the 60 nM mixture, we determine there are 955 and 1347 translocation events for 100 bp and 500 bp, respectively.
populations with $<t_d>$ of 210 μs and 880 μs respectively. When we set a threshold dwell time of 450 μs (pink dotted line), we find that 98.4% of 100 bp events fall below this limit while 97.6% of 500 bp events lie above this cutoff, demonstrating a ~98% accuracy in differentiating these DNA lengths. When evaluating the relative capture rate in this mixture we find that the capture rate for 500 bp is 68% greater than for 100 bp. This too is in agreement with past results stating that in the presence of an energy barrier for capture due to a small pore, the larger DNA is more likely to be captured.93

2.7 Discussion

We have identified here an optimum pore geometry that leads to smooth transport of double-stranded DNA. The DNA trajectory statistics can be well modeled by first-passage time distributions derived from a 1D drift-diffusion model. The model we have used describes the biased escape of a particle from a trajectory under subjection to a uniform field. Clearly, this 1D model is not sufficiently elaborate for describing a complex process such as semi-flexible polymer translocation through thin pores (i.e., non-uniform fields), and other models that capture various phenomena such as polymer-pore interactions and coil effect should be developed and/or implemented to better explain our data. However, we have found that this simple two-parameter model adequately explains our results by providing for each experiment a characteristic axial diffusion coefficient and velocity. Quantitative fits of our data to this model using a bootstrapping algorithm shed light on the DNA translocation process through small pores, and we found that transport through nanopores with $d \sim 3$ nm is smooth, i.e., governed by a drift-diffusion process. Experiments using larger pores in the range 3.4 – 6.0 nm reveal an additional population with longer dwell times than predicted by the model, which we attribute to DNA coil self-interference that is driven by the external field in larger pores.
We caution the reader that our use of “smooth” to describe the translocation process does not imply that DNA velocity is constant. For both a single molecule and an ensemble of molecules, translocation velocities are not generally uniform but subject to various types of random forces that overall prescribe normally distributed velocity profiles (similar to how dwell times are distributed according to Eq. 1). We show that the electrophoretic mobility of DNA is smaller than DNA mobility in bulk solution and is a function of applied voltage and pore diameter. Turning to the diameter dependence, experiments and simulations reveal that axial DNA diffusion in small pores is greatly reduced due to confinement, and that this reduced diffusion proportionally reduces the mean velocities to values that are below 0.34 nm/µs (i.e., \( v/D \) is constant). The dependence of translocation time on DNA length reveals a power law scaling with an exponent of 1.37, which is in good agreement with other experimental results using a range of pore diameters. However, for DNA lengths longer than 6 kbp we have found noticeable (though not major) fractions of observed events with shorter dwell times than the major population, which point to coil-induced complications in the transport process. Fine-tuning our pores for smooth DNA transport enabled us to discriminate between 100 bp and 500 bp in a mixture from individual pulses with >98% accuracy. This regulation of double-helical DNA transport through nanopores provides a steady control that could prove useful in detecting other structures such as non-canonical DNA motifs, DNA-protein interactions, and epigenetic modifications.
2.8 Materials and Methods

2.8.1 Nanopore experiments.

For all details about nanopore fabrication and information about our experimental setup, please refer to Section A.1. Unless otherwise indicated, all experiments were performed at room temperature (21°C) and using 0.40 M KCl electrolyte tris-buffered to pH 7.9.

2.8.2 Data acquisition and analysis.

Experimental data was collected using the Chimera Instruments VC100 (New York, NY) at a sampling rate of 4.167 MHz, and further digitally low-pass filtered at 200 kHz prior to analysis in order to reduce the capacitive noise. The high-bandwidth increases our time resolution and ensures that events as fast as 2.5 μs are detected. All DNA samples used in these studies were obtained from Thermo Scientific (Waltham, MA). Analysis of all nanopore data was performed using MATLAB-based OpenNanopore software\textsuperscript{66} developed by the Radenovic group at EPFL, which uses a cumulative sums algorithm to detect individual events in the raw current signals. Standard errors shown for $D$ and $v$ were estimated using a bootstrapping procedure that is described thoroughly in Section A.1.

2.8.3 Molecular dynamics simulations of diffusion in pore-confined DNA.

MD simulations were performed using the NAMD\textsuperscript{107} software package. Periodic boundary conditions were applied, and particle mesh Ewald electrostatics\textsuperscript{108} governed the long-range interactions. Multiple time-stepping\textsuperscript{109} was used so that local interactions were calculated every time step, and the full electrostatic calculation was performed every 3 time steps. A 2 fs time step was used with RATTLE\textsuperscript{110} and SETTLE\textsuperscript{111} algorithms applied to covalent bonds involving hydrogens in DNA and water, respectively. The van der Waals forces were smoothly
cut off starting at 7Å and were completely cut off at 8Å. The CHARMM36\textsuperscript{112} force field was used for nucleic acids, water, and ions, with NBFIX corrections for ions,\textsuperscript{113} and a custom force field for Si\textsubscript{3}N\textsubscript{4}.\textsuperscript{114} The NPT simulations used a Nose-Hoover Langevin piston pressure control,\textsuperscript{115} and the temperature was controlled by a Langevin thermostat acting on the membrane atoms with a damping constant of 1.0 ps\textsuperscript{-1}. During NPT simulations, the DNA and Si\textsubscript{3}N\textsubscript{4} were harmonically restrained with a spring constant of 695 pN/Å. In all NVT production simulations, the surface atoms of the membrane were harmonically restrained with a spring constant of 695 pN/Å, and the interior atoms of the membrane were harmonically restrained with a spring constant of 69.5 pN/Å.

Atomic-scale models of silicon nitride nanopores were built following a previously described protocol.\textsuperscript{21} The pores were made by removing atoms from a crystalline silicon nitride membrane. A 36 bp fragment of dsDNA was introduced into the pore, with its axis collinear with the pore axis. Each DNA/nanopore system was solvated; potassium and chloride ions were added to produce electrically neutral systems of target KCl concentration (1.0 or 0.40 M). Each system (about 120,000 atoms) underwent 1,000 steps of energy minimization, followed by 1 ns equilibration in the constant area, pressure (1 bar) and temperature (295 K) ensemble maintained using a Langevin piston. All production simulations were carried out in a constant volume/temperature ensemble. For information on the script detailing the pulling of DNA molecules see Section A.1.

\textbf{2.8.4 Finite element simulations.}

All simulations presented were computed using COMSOL Multiphysics 4.3b (Burlington, MA) with custom geometries (see Section A.7 for more information).
Chapter 3 – Hydroxymethyluracil Modifications Enhance the Flexibility and Hydrophilicity of Double-Stranded DNA

This chapter is adapted with permission from Carson, S., Wilson, J., Aksimentiev, A., Weigele, P. R., and Wanunu, M. "Hydroxymethyluracil modifications enhance the flexibility and hydrophilicity of double-stranded DNA." Nucleic Acids Research (2015): gkv1199. Copyright 2015 Oxford University Press.
3.1 Introduction

Epigenetic modifications in eukaryotic DNA, which are most commonly manifested in cytosine (C) taking the form of 5-methylcytosine (mC) and its oxidized product 5-hydroxymethylcytosine (hmC), have been implicated in gene expression, genetic imprinting, and other various genetic processes. Less abundant than cytosine modifications, several oxidized products of thymine (T) including 5-hydroxymethyluracil (hmU), formed by ionizing radiation, reactive oxygen species, or Tet enzymes, have been correlated with chronic inflammatory diseases and cancers. Although these thymine lesions can be detected in bulk solution using mass spectroscopy, little is known about their impact on DNA structure. Nanopores have emerged as single-molecule probes for high-resolution determination of base composition in double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA), which have made nanopores a prime candidate for next-generation DNA mapping and sequencing applications.

A nanometer-size aperture in a thin membrane—a nanopore—provides a solvent-filled passage between two solution compartments separated by the membrane. When subjected to a trans-membrane voltage, individual molecules of DNA can transit from a negatively biased compartment into a positively biased one through the nanopore, as depicted in Fig. 1a. The presence of DNA in a nanopore is experimentally detected as a transient drop in ionic current flowing through the nanopore. Each translocation event is characterized by the duration of the ionic current blockade, or dwell time ($t_d$), and the amount of ion current reduction, or current blockade ($\Delta I$). The mean values of $t_d$ and $\Delta I$ determined from a population of single-molecule transport events (typically at least 1,000) reflect properties of the molecules, e.g., their contour length, cross-sectional diameter, and microscopic conformation.
Recently, various approaches to differentiating dsDNA fragments that contain cytosine and modified cytosines have been pursued using small diameter (~4 nm) silicon nitride (SiN) nanopores. In these studies, DNA flexibility accounted for differences in transport dynamics, and has been previously investigated using experimental and computational tools. In addition, nanopore identification of single cytosine modifications was demonstrated using methylation-specific binding proteins, streptavidin-based ssDNA immobilization, and DNA polymerases acting as molecular ratchets. In this article, we characterize the transport kinetics of the following DNA variants: (I) canonical DNA containing no thymine modifications (t-DNA), (II) DNA where all thymine nucleotides were replaced by hmU (h-DNA), and (III) h-DNA molecules that were further modified to have all TG sites phosphorylated (p-DNA) (see Section 3.6.2 for sample preparation information and Fig. B1 for DNA sequences). We show that thymine conversion to hmU and phosphorylated hmU (phmU) induces noticeable structural differences that affect transport of dsDNA through ultrathin SiN pores. Molecular dynamics (MD) simulations characterize the effect of the modifications on the equilibrium properties of DNA and relate the effect of modifications to the outcome of nanopore measurements.

3.2 Transport of hmU-containing DNA variants through small nanopores

Figure 3.1 schematically illustrates the process of DNA translocation through a nanopore (Fig. 3.1a), the chemical structures of the modified bases (Fig. 3.1b-d), and the gel electrophoresis image of the products (Fig. 3.1e). A low voltage bias (i.e., 200 mV) applied across a nanopore generates a steady-state ion current. Addition of DNA to the cis (-) chamber leads to a stochastic set of downward current spikes caused by individual DNA molecules transporting across the pore. Detection and analysis of each event using Python
(www.github.com/rhenley/Pyth-Ion) yields the dwell time $t_d$ and the fractional current blockade $\Delta I/I_0$, which reflect DNA properties such as diameter, length, base content, and flexibility.

Figure 3.2 summarizes the results of DNA translocation experiments performed at $V = 200$ mV, 0.40 M KCl, pH 7.9, and $T = 21^\circ$C (an additional dataset for a smaller diameter pore is shown in Fig. B2). Figure 3.2a shows a continuous trace for t-DNA along with a sample of concatenated, consecutive events and Fig. 3.2b displays the scatter plots that summarize our data. Dwell time histograms shown in Fig. 3.2c reveal that h-DNA passes through the pore noticeably faster than both t-DNA and p-DNA, with the latter two molecules having similar mean dwell times. Fitting the dwell time distributions to a 1D drift-diffusion model (black curve)\textsuperscript{43, 68, 69, 125} yields two parameters, the drift velocity $v$ and

---

**FIGURE 3.1** - Nanopore detection of thymine variants in DNA. (a) A microscopic model of a DNA translocation experiment featuring a 3.5 nm diameter SiN nanopore (gray) and a fragment of p-DNA. The DNA molecule is colored according to its nucleotide content: red corresponds to adenine nucleotides (A), green to cytosines (C), orange to guanines (G), blue to 5-hydroxymethyluracil (hmU) and magenta to phosphorylated 5-hydroxymethyluracil (phmU). (b-d) Chemical structures of thymine (T) and its hmU and phmU variants. The modifications are highlighted in blue and magenta. (e) Gel electrophoresis of 101 bp DNA samples (20% PAGE, lanes 1 and 5: 50 bp dsDNA ladder; 2: t-DNA; 3: h-DNA; 4: p-DNA).
The presence of hmU modifications increases the values of both $v$ and $D$, while the mean current blockade decreases (Fig. 3.2d). Considering the added oxygen within the hmU modification, we were surprised to find lower $\Delta I/I_o$ values for that...

**FIGURE 3.2** - Translocation of thymine-modified DNA through a 2.8 nm diameter pore (effective pore thickness = 10 nm). (a) Raw current trace (top) and a concatenated set of events (bottom) obtained for t-DNA (traces filtered at 200 kHz). (b) Scatter plot of $\Delta I/I_o$ vs. $t_d$ for each DNA variant. (c) Normalized dwell time histograms for each DNA sample (event totals displayed). Histograms are fit to a 1D drift-diffusion model (see Table 3.1 for fit parameter details, and Eq. 1 for fitting equation). (d) Fractional current blockade histograms of each DNA variant (Gaussian fits shown).
molecule, which implies that the larger steric footprint of an hmU nucleotide (which is expected to increase the current blockade) is overcompensated by some other effect that reduces the current blockade. Our observations are consistent with a model where a DNA containing hmU modifications has a more compact overall structure than its native T-form analog. Although steric exclusion is the most common mechanism of ionic current blockade, it is possible that changes in hydrophobicity and flexibility of DNA can account for the change in the ionic current blockade, as discussed below.

**TABLE 3.1. Experimental and Simulated Translocation Data for Thymine-Modified dsDNA**

<table>
<thead>
<tr>
<th>Sample</th>
<th>$&lt;t_d&gt;$ (μs)</th>
<th>$v$ (nm/μs)</th>
<th>$D$ (nm$^2$/μs)</th>
<th>$\Delta I / I_o$E</th>
<th>$\Delta I / I_o$S</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-DNA</td>
<td>293 ± 5</td>
<td>0.189 ± 0.003</td>
<td>0.45 ± 0.04</td>
<td>0.516 ± 0.022</td>
<td>0.477 ± 0.008</td>
</tr>
<tr>
<td>h-DNA</td>
<td>159 ± 2</td>
<td>0.325 ± 0.004</td>
<td>0.71 ± 0.07</td>
<td>0.490 ± 0.024</td>
<td>0.454 ± 0.007</td>
</tr>
<tr>
<td>p-DNA</td>
<td>309 ± 4</td>
<td>0.169 ± 0.002</td>
<td>0.53 ± 0.05</td>
<td>0.498 ± 0.023</td>
<td>0.471 ± 0.007</td>
</tr>
</tbody>
</table>

*Left: Experimental mean dwell time $<t_d>$, drift velocity $v$, diffusion coefficient $D$, and fractional current blockade $\Delta I / I_o$ determined by best fits of dwell time (1D drift-diffusion, Eq. 1) and current blockade (Gaussian) histograms. Right: MD simulation results for $\Delta I / I_o$. For determination of uncertainties and details on the MD simulations, see Section 3.6.*

### 3.3 Melting curve analysis of DNA variants

To investigate the relative stability of our DNA variants, we determined the melting temperatures of each sample by collecting melting curves of the three DNA variants (Fig. 3.3a). From the peak of each differential fluorescence curve, we obtained the following melting temperatures for each sample in 0.4 M KCl buffer conditions: $T_{t-DNA} = 87.6 ± 3.0 \degree C$, $T_{h-DNA} = 80.6 ± 1.8 \degree C$, $T_{p-DNA} = 77.6 ± 2.6 \degree C$. The melting curve data for the p-DNA sample is more sensitive to ionic strength than that of h-DNA, as shown in Fig. 3.3b, which is consistent with the added charge of the modified base pair in that construct. The lower melting temperatures of h-
DNA and p-DNA may be explained by the more hydrophilic structure of the grooves as compared with native t-DNA, which was found to be the case for the hmC modifications. However, the faster transport of h-DNA could not be explained directly using hydrophobic/hydrophilic considerations, as it is not clear what impact this parameter has on transport dynamics through a pore.

3.4 Molecular dynamics simulations reveal hmU increases DNA flexibility

To elucidate the structural differences between the modified DNA strands, we performed all-atom MD simulations of 36 bp fragments of the 101 bp DNA constructs in 0.4 M KCl solution and in the absence of a solid-state nanopore (the nucleotide sequence of DNA used in the simulations is provided in Section 3.6.3). Nine independent trajectories (three for each variant) of over 100 ns each were obtained in total. The trajectories were analyzed using the 3DNA program that characterized the equilibrium conformation and structural fluctuations of the DNA constructs in terms of six local intra-base pair parameters (shear, buckle, stretch, propeller, stagger, and opening) and six parameters defining the conformation of two neighboring base pairs (shift, tilt, slide, roll, rise, rise,
and twist). Of these parameters, roll and twist describe bending of the double helix, whereas rise describes longitudinal stretching.

We were primarily interested in the relative flexibility of the DNA constructs. Since the density of hmU modifications in the experimental constructs was very high (i.e., ~60% of base pairs were modified), we predominantly focus here on structural fluctuations averaged over the entire constructs (Fig. 3.4c-d), although the effect of point modifications was also addressed through simulations (Figs. B3, B4). There are two modalities of flexibility in DNA: kinking and bending. During our ~1 µs of simulation, we observed three kinking-like events (described in detail in the Section B.2), which did not allow us to draw a conclusion about the relative frequency of such events among the DNA variants. Thus, we restrict our analysis to the elucidation of the effect of hmU and phmU modifications on near-equilibrium structural

![FIGURE 3.4](image)

**FIGURE 3.4** - Results of molecular dynamics simulations characterizing structural fluctuations of the DNA variants. (a-b) Molecular bond representation of the hmU modification (a) and phmU modification (b) within the h-DNA and p-DNA constructs, respectively. (c) Differences in the standard deviation of the intra-base pair parameters for the hmU- and phmU-modified DNA relative to the values obtained for unmodified t-DNA. The h-DNA construct exhibits increased fluctuations in the average base pair geometry. (d) Differences in the standard deviation of the inter-base pair parameters for the h-DNA and p-DNA constructs relative to t-DNA. Both h-DNA and p-DNA show increased inter-base pair fluctuations suggesting that these molecules are more flexible than t-DNA. The relative difference in standard deviation $\Delta \sigma_x = 100 \times (\sigma_x - \sigma_t)/\sigma_t$, where $\sigma_t$ and $\sigma_x$ are the standard deviations of a parameter in t-DNA and x-DNA, respectively (x is either h or p), averaged over the time series and then over base pairs or base pair steps. The error was calculated by splitting MD trajectories of the same chemical system into thirty blocks and calculating the standard deviation within each block for each base pair or step. The standard error of these 30 estimates was then propagated to reflect averaging over all base pairs or steps.
fluctuations. Figure 3.4a-b shows typical near-equilibrium structures of the hmU and phmU modifications observed in our MD simulations.

The results of our analysis indicate that both hmU and phmU modifications have a pronounced effect on the standard deviations of the intra- and inter-base pair parameters, and hence on the elastic properties of the molecules. Figure 3.4c-d plots $\Delta \sigma$, the relative difference in the standard deviation averaged over base pairs and trajectories, of each inter- and intra-base pair parameter for the two DNA variants with respect to the values observed for unmodified DNA (t-DNA). The h-DNA shows an increase in fluctuations for all of the intra-base pair parameters (Fig. 3.4c), indicating increased flexibility of the DNA upon addition of hydroxyl groups. Similar changes in DNA flexibility were previously observed upon addition of hydroxyl groups to mC-modified DNA. The inter-base pair parameters also show a marked increase in fluctuations, as seen in Fig. 3.4d, with all six parameters showing at least a 4% increase in the standard deviation over t-DNA. Specifically, the roll and twist parameters, which are closely linked to DNA bending, are 4.5% and 7.5% greater than in t-DNA, indicating a greater bending ability of h-DNA. The standard deviation of the rise parameter is 7% greater in h-DNA than in t-DNA, indicating a greater propensity of h-DNA to stretch. Although large increases in the standard deviation of the intra-base pair parameters were not observed for p-DNA, the fluctuations of the inter-base pair parameters for p-DNA were similar to those of h-DNA, suggesting similar flexibility of the two constructs. The latter result is perhaps not surprising, as our p-DNA molecule contained a large number of hmU modifications in addition to phmU modifications. Similar increases of the standard deviations were observed in our MD simulations of DNA constructs containing individual hmU modifications (Fig. B3). In comparison to hmU,
individual phmU modifications reduced the standard deviation of the structural parameters, making the DNA more rigid (Fig. B4).

MD simulations of the DNA constructs in a solid-state nanopore provided further insights into our experimental observations. For these simulations, we constructed four atomic-scale models of the experimental system, each containing a 3.5 nm diameter nanopore in a 7 nm-thick SiN membrane, as shown in Fig. 3.1a (see Section 3.6.3 for simulation details). One of the models contained a nanopore filled with 0.4 M KCl (open pore system), while the other three also contained a 36 bp fragment of the 101 bp construct used in experiment inserted into the pore. Each system was subject to an external electric field producing a 200 mV bias across the membrane resulting in an ionic current that was measured\textsuperscript{189} (see Section B.5 for details). For each variant of DNA, we ran five independent simulations of 120 ns each. To minimize the effect of conformational noise, the DNA constructs were restrained to adopt the same conformation in the nanopore, allowing us to obtain an accurate estimate of the ionic currents. The ionic current data was first averaged in 20 ns blocks, and then the mean current and standard error were computed by averaging over 30 such blocks. The results of the applied field simulations indicated a 4.5% increase in the current when all T were replaced by hmU, which was accompanied by a 5.2% average increase in the number of mobile ions (not bound to the pore walls) within the pore constriction (Fig. 3.5a). When comparing h-DNA with p-DNA, the number of potassium ions is higher in the pore for p-DNA (Fig. 3.5b). This increase, however, is fully explained by the added charge of the phosphorylated base (-2e\textsuperscript{−}), while no change in chloride density in the pore was observed. While simulated fractional current blockades $\Delta I/I_o$ qualitatively agree with the experimental trend h-DNA < p-DNA < t-DNA (see Table 3.1), it is
the combination of the DNA structural dynamics and the ion distribution in the nanopore that accounts for the experimental translocation data.

### 3.5 Discussion

Our combined results suggest that the presence of a hydrophilic hmU modification in place of T enhances DNA flexibility. Thus, our nanopore measurements indicate faster translocation of h-DNA in comparison to t-DNA, while MD simulations indicate that h-DNA is more flexible both laterally and longitudinally than t-DNA. This enhanced flexibility of h-DNA renders its navigation through the steric constraints of a narrow pore more efficient than that of a stiffer molecule, which explains the faster translocation. This situation is in contrast to the behavior expected for large pores and long DNA constructs, where increased flexibility of a molecule is expected to make its translocation through a nanopore slower due to increased polymer entropy. Enhanced longitudinal flexibility of h-DNA and p-DNA could have an additional effect on the ionic current blockades by reducing the effective diameter of the molecules when

*FIGURE 3.5* - MD simulations of ion atmosphere in a solid-state nanopore occupied by DNA variants. (a) The effect of hmU modification on the distribution of K⁺ (*solid line*) and Cl⁻ (*dashed line*) ions in a 3.4 nm diameter nanopore. The local concentration difference between the h-DNA and t-DNA systems (*black lines*) is plotted as a function of the radial distance from the pore axis. The horizontal line passing through zero is a guide to eyes. (b) Same as in panel (a), but for the p-DNA and t-DNA systems. The regions occupied by DNA are indicated by blue rectangles.
they are subject to the stretching force of the electric field in a nanopore. Overall, our study suggests that oxidation of DNA bases can affect its mechanical properties and that such modifications can be studied by nanopore translocation measurements. These findings suggest a role of oxidized DNA bases on the deformability and ionic environment of the grooves of modified DNA, which may modulate the binding of regulatory and repair proteins.

3.6 Materials and Methods

3.6.1 Nanopore fabrication and uncertainty determination

The fabrication of solid-state nanopore devices follows the procedure described in our previous publication. The specified uncertainties of $\Delta I/I_0$ represent the standard deviations determined from Gaussian fits, and the uncertainties of $t_d$ represent the standard errors of the mean. The uncertainties reported for $v$ and $D$ are calculated using a standard bootstrapping procedure, as explained in earlier work.

3.6.2 DNA sample preparation

The 101 bp sequence used for these studies is displayed in Fig. B1. Within this sequence there are 57 thymine sites that are enzymatically modified to hmU in h-DNA and 10 sites that are additionally modified to phmU (with 47 sites remaining hmU) in p-DNA. The DNA fragments t-DNA and h-DNA were prepared by PCR amplification using dNTP mixtures composed of either canonical nucleotides or a mixture with deoxy-hmUTP fully replacing dTTP. Fragments containing phmU (p-DNA) were prepared by treating the h-DNA with a kinase that specifically phosphorylates the hydroxymethyl moiety of an hmU 5' to a guanine. phmU is ordinarily formed as an intermediate to base hypermodification during the morphogenesis of the bacteriophages SP10 and ΦW14. The kinase catalyzing the formation of phmU from ATP and hmUpG in
polymeric DNA has recently been purified and shown to function *in vitro*. (P. Weigele, manuscript in preparation)

### 3.6.3 Molecular dynamics simulations

All MD simulations were performed using the NAMD2\textsuperscript{107} software package using the CHARMM36\textsuperscript{112} force field for nucleic acids, water, and ions. Ionic interactions included NBFIX corrections;\textsuperscript{113} a custom force field was used to describe the Si\textsubscript{3}N\textsubscript{4}\textsuperscript{21}. The Si\textsubscript{3}N\textsubscript{4} parameters were derived from first principle parameterization\textsuperscript{134} and modified to be compatible with the CHARMM force field. The parameters for the hmU and phmU modifications were obtained using the CHARMM general force field library.\textsuperscript{135}

Periodic boundary conditions were used along with the particle mesh Ewald method for long-range interactions electrostatics.\textsuperscript{108} Local interactions were calculated every time step, and the full electrostatic calculation was performed every three time steps using a multiple time-stepping scheme.\textsuperscript{109} A 2 fs time step was used with RATTLE\textsuperscript{110} and SETTLE\textsuperscript{111} algorithms applied to covalent bonds involving hydrogen atoms in DNA and water, respectively. The van der Waals forces were smoothly cut off starting at 10 Å and were cut off completely at 12 Å. A Nose-Hoover Langevin piston\textsuperscript{115} was used for pressure control in NPT (i.e., constant number of particles, pressure, and temperature) simulations with a period of 400 fs and a damping time scale of 200 fs. The temperature was controlled by a Langevin thermostat acting on the membrane atoms with a damping constant of 1.0 ps\textsuperscript{-1}. In our free solution simulations, the Langevin thermostat acted on water oxygen atoms with a damping constant of 0.5 ps\textsuperscript{-1}. To induce a transmembrane potential, a constant electric field was applied such that \( E = -\Delta V/l_z \), where \( \Delta V \) was the target transmembrane bias and \( l_z \) was the length of the simulation system.
normal to the membrane (the z-axis in our system).\textsuperscript{21, 89} In accordance with a previously established protocol,\textsuperscript{114} each atom at the surface or in the interior of the membrane was harmonically restrained with a spring constant of 695 or 69.5 pN/Å, respectively, to give the membrane material a relative bulk permittivity of 7.5.

The atomic-scale model of the silicon nitride nanopore was built as in our previous work.\textsuperscript{125} The nanopore was made by removing atoms from a 7 nm thick crystalline silicon nitride membrane, so that the final nanopore had an hourglass shape with a minimum diameter of 3.5 nm. The DNA fragments were created using the 3D-DART web server\textsuperscript{136} in a BDNA conformation. The 36 bp DNA fragments were then placed coaxial with the pore axis. Water was added to the open nanopore and DNA/nanopore systems with the solvate plug-in for VMD, and then ions were added to a target KCl concentration of 0.4 M using the autoionize plug-in for VMD. Each DNA/nanopore system was then equilibrated under NPT for 1 ns using a Langevin piston. Following equilibration, the system was simulated in the NVT ensemble (\textit{i.e.}, constant number of particles, volume, and temperature) at a 200 mV transmembrane bias for 55 ns, which was sufficient for the number of ions to reach a constant value in the nanopore. The DNA sequence was chosen to reproduce a 36 bp fragment of the DNA used in the experiment or its variants (listed below).

\textbf{t-DNA:} CCATTCTTCCAAGTAGCTGAGTCTATGTGGATTTTA

\textbf{h-DNA:} CCAHHCHHCCAAGHAGCHGAGHCHAHGHGGAAHHHHA

\textbf{p-DNA:} CCAHHCHHCCAAGHAGCPSGAGHCHAHGPGGAAHHHHA
Each DNA/nanopore system was then copied to make five independent simulations, each subject to an additional 2 ns equilibration in the NPT ensemble. The fifteen systems were simulated for 140 ns each under a 200 mV transmembrane bias in the NVT ensemble. The first 20 ns of each trajectory were discarded to ensure the ions were well equilibrated; the final 120 ns of each simulation were used to measure the blockade current. Each backbone phosphorous atom of the DNA was harmonically restrained to maintain a radial distance of 9.45 Å from the nanopore axis, and to not move along the nanopore. The spring constant of each restraint was chosen to be 6.95 pN/Å based on several short simulations. This choice of restraints allowed the DNA to alter its conformation as much as possible while remaining in the center of the nanopore. The application of such restraints maintained the DNA conformation coaxial with the nanopore while allowing the DNA to rotate about its axis.

The simulations of DNA in free solution were performed starting from the DNA conformations obtained at the end of the nanopore simulations. The molecules were placed into a cubic box, 13.7 nm on each side, and solvated with a 0.4 M KCl solution. Three free solution equilibration trajectories of ~120 ns each were obtained for each of the three DNA variants (nine simulations in total). The simulations were run at a constant pressure of 1 atm and temperature of 295 K as described above. The trajectories were then analyzed using the 3DNA program,\textsuperscript{108} discarding the first 10 ns of each trajectory.

In two of the nine trajectories, a localized break in the DNA structure developed near the end of the simulation. As we were predominantly interested in characterization of near-equilibrium fluctuations of the DNA conformation, the part of the trajectory following the breaks were excluded from the analysis (see further details in Section B.2). The mean value and the
standard deviation of the intra- and inter-base pair parameters were calculated for each base pair (or pair of base pairs). The last 2-3 base pairs at each end of the molecule frayed during the simulations, making the bases intermittently unpaired. As such unpaired bases are not suitable for 3DNA analysis, five terminal base pairs at each end of each molecule were excluded from the analysis. The individual base pair values were averaged over the entire construct (excluding the five terminal base pairs). Doing so was justified by the high density of the modifications, which were present in at least every third base pair.
Chapter 4 – Direct Analysis of Gene Synthesis Reactions Using Solid-State Nanopores

4.1 Introduction

Synthetic genes\textsuperscript{137, 138} and genomes\textsuperscript{139, 140} offer a transformative ability to design and create new cellular functions and entire organisms.\textsuperscript{141} To leverage these programmable DNA elements into their cellular chassis, precise synthetic control of their properties (sequence, length) must be obtained. However, reliable construction of these DNA elements using polymerization,\textsuperscript{142} restriction enzymes,\textsuperscript{143} and/or ligation reactions\textsuperscript{144, 145} can be challenging, especially for long formats. Every gene synthesis method faces the challenge of unwanted side products, with the purity of resultant mixture being a function of the assembly method, DNA sequence composition (length, complexity, repeats), quality of input oligonucleotide building blocks, and other factors.\textsuperscript{146, 147}

To further increase the complexity of producible synthetic genes and better understand their function, new methods are needed to assess and control their output diversity so that unwanted resultant product formation is minimized. Two major classes of deleterious products that arise during gene synthesis reactions are global structural defects (incomplete or misassemblies) and small local defects (point mutations, short insertions, and/or deletions). These errors can be identified using laborious procedures such as electrophoresis, cloning, and sequencing, which consequently become major components of the time and cost of gene synthesis. Moreover, as synthesis approaches continue to move towards miniaturization (employing microarrays\textsuperscript{148} and microfluidic chips\textsuperscript{149}), where reactions are parallelized and costs are lowered, analytical quality-control tools that match smaller scales (\(\mu\text{m} \) sizes and nM/pM concentrations) become increasingly valuable.
Solid-state nanopores\textsuperscript{5} are a new class of sensors that can electronically detect the structure and conformation of single DNA molecules with high throughput (thousands of molecules per minute). A nanopore is a nanometer-scale aperture in a dielectric membrane that separates two ionic solutions. Application of a voltage bias across the membrane induces an ionic current through the nanopore (Fig. 4.1a). As single DNA molecules in the ionic solution electrophoretically move through the pore, they temporarily block ion current through the pore. Transient changes in the ionic current can then be used to infer characteristics of the translocating molecules such as length differences,\textsuperscript{18, 150, 151} bound proteins,\textsuperscript{40, 152} epigenetic modifications,\textsuperscript{7, 153} and defects.\textsuperscript{154} The ability to detect small sample amounts of nucleic acids without amplification or labeling, as well as the potential of using nanopores to isolate a selected type of molecule, makes nanopores an attractive platform for direct analysis of gene synthesis reactions. Herein, we utilize ultrathin solid-state nanopore sensors\textsuperscript{71} to directly detect different types of synthetic gene products. We designed, assembled, and interrogated a set of DNA structures that represent major categories of expected defects (\emph{i.e.}, dsDNA molecules with single-base mismatches, overhangs, flaps, and Holliday junctions) to train our system and depict differentiation of these structures. We then monitor a gene synthesis reaction at different stages and discretize the reaction products into categories of correct and incorrect assemblies based on ensemble bunching of molecular transport signatures. With further refinement of nanopore-based discrimination among correctly and incorrectly assembled structures and real-time on-the-fly analysis, we envision a true molecular selection process in which correct structures can be identified and enriched.
4.2 Rapid detection of small DNA defects

De novo gene synthesis reactions can produce a variety of unwanted products such as small sequence defects as well as structural misassemblies. To enable detection of these defects, we first aimed to show that the translocation signatures of correctly assembled, homogenous DNA molecules would be fundamentally different from those of misassembled synthetic products. First, thin nanopores were fabricated in silicon nitride (SiN) windows as previously described\(^\text{125}\) with diameters of 2.4-2.6 nm and effective thicknesses of 4-8 nm (Fig. 4.1a). Nanopores within this diameter range permit unfolded entry and single current blockade levels during electrophoretic translocation.\(^\text{56}\) The TEM-drilled nanopores were rinsed with hot piranha, cleansed thoroughly with deionized water, vacuum dried, and mounted into a custom PTFE flow cell that was subsequently filled with an ionic solution (0.40 M KCl, 10 mM Tris, 1 mM EDTA, pH 8.0). Application of a voltage bias across the membrane drove an ionic current through the nanopore, which demonstrated Ohmic behavior (Fig. 4.1b). Upon addition of a small concentration of 70 bp dsDNA homoduplexes (50 nM, IDT Technologies) and application of a voltage bias (200 mV), transient dips in

**FIGURE 4.1 - DNA translocation through a solid-state nanopore.** (a) Illustration of a small trans-membrane voltage causing dsDNA to pass through a nanoscale aperture. Inset: Transmission electron microscope image of a small SiN nanopore (scale bar is 2 nm). (b) Current-voltage curve in the range -300 mV to 300 mV for one of the nanopores used in this study. (c) Current trace for 70 bp DNA at an applied voltage of 200 mV, low-pass filtered at 200 kHz. Each deep spike corresponds to a single DNA molecule translocating the nanopore in head-to-tail fashion. (d) Analysis of DNA translocation data extracts three desired quantities: dwell time \((t_d)\), current blockade \((\Delta I)\), and fractional current blockade \((\Delta I/I_o)\).
current were observed, signifying translocation of individual DNA molecules through the nanopore (Fig. 4.1c). DNA capture has been shown to be dominated by electrophoresis due to an exponential increase in the capture rate with voltage,\textsuperscript{93} despite a small negative surface charge present in SiN nanopores, which causes an electroosmotic flow in a direction opposite to DNA translocation.\textsuperscript{155, 156} Using custom analysis software, two key parameters were extracted for each translocation event: the dwell time $t_d$ and the current blockade $\Delta I$ (Fig. 4.1d). Plotting the fractional current blockade ($\Delta I/I_o$) versus $t_d$ for 1,830 events showed a distribution with an average translocation speed (1 bp/μs) consistent with previous studies\textsuperscript{56, 125} and a single current blockade level with mean amplitude of $\Delta I/I_o = 0.79 \pm 0.03$ (Fig. 4.2b). Next, we sought to determine the nanopore’s ability to detect small synthetic errors, such as point deletions or insertions, using one strand of the control oligo of “correct” sequence (70 nt) annealed to a “defective” DNA oligo (71 nt), which was synthesized with a single base insertion, as illustrated in Fig. 4.2a. The defective site was located in the center of the mismatch oligo and is expected to cause a small kink in the short, otherwise stiff dsDNA, which we hypothesized would alter the entry kinetics and enable differentiation. Insertion of a small concentration of the mismatch dsDNA species (70/71 bp) into the nanopore setup and application of a voltage bias as before generated 1,410 events with a highly similar distribution of $\Delta I/I_o$ against $t_d$ to the correctly matched oligos (Fig. 4.2b). One subtle difference observed between the current blockade distributions of the correct and defective DNA was a small, additional population at lower $\Delta I/I_o$, which could be due to the translocation of misassembled DNA molecules or unsuccessful translocations (collisions) caused by the kink. The dwell time distributions yielded no statistically significant discrepancies between the two molecules, and as such, differentiation of single-base defects would be inherently difficult to detect directly.
We validated the difference between the correctly matched and defective species by incubation of both products with a mismatch-binding protein from *Thermus aquaticus* named MutS\(^{157}\) (Taq MutS, 89 kDa), which is a part of the DNA repair pathway in a variety of organisms (Fig. C1). *Taq* MutS possesses an elongated shape (longest dimension about 10-12 nm) and binds to all single base mismatches, insertions, or deletions up to four bp long in double stranded DNA. Previously, we demonstrated the use of *Taq* MutS to discriminate all types of
single base mismatches and insertions/deletions of various lengths (in one case as much as a 50 bp deletion) and gel shift analysis showed a clear shift for the defective products compared to the correctly matched samples (Fig. C1). To further enable detection of small sequence defects on synthetic genes using a solid-state nanopore, the correctly matched 70 bp DNA and mismatch-containing 70/71 bp DNA species were each incubated with the MutS protein (ratio of 2 MutS : 5 DNA) for 20 min at 60°C in 0.4 M KCl, 8 mM MgCl₂ buffer (Fig. 4.2d). Both populations were individually measured using the nanopore setup as before, and the correctly matched dsDNA incubated with MutS demonstrated highly similar \( t_d \) and \( \Delta I/I_0 \) distributions to the same correctly matched dsDNA population without MutS incubation, indicating little to no interaction between the two species (Figs. 4.2e-f). This observation is consistent with gel shift analysis whereby incubation of the correctly matched dsDNA with MutS did not significantly alter correctly assembled product migration, even with increases in concentration of MutS (Fig. C1). In contrast to the correctly matched dsDNA products incubated with MutS, the mismatch products + MutS displayed \( t_d \) distributions that could not be fit to a drift-diffusion model⁶⁸,⁶⁹ and \( \Delta I/I_0 \) distributions did not fit to a single Gaussian distribution (Fig. 4.2e). The current blockade distribution for the mismatch products + MutS showed a similar peak as the 70 bp molecule with an additional shoulder at a lower current blockade, which may be attributed to collisions of the DNA/MutS complex with the pore mouth, or translocation of free oligos and misassembled DNA molecules that also bind MutS. In contrast to the correctly matched dsDNA + MutS, the mismatch dsDNA + MutS also displayed a distorted dwell time distribution that was fit to a single exponential with mean timescale of \( t_d = 829 \pm 60 \) μs, compared to a mean \( t_d \) of 524 ± 47 μs for the correctly matched dsDNA (Figs. 4.2e-f). The increase in dwell times can be attributed to weak DNA-protein interactions that stall the complex inside the pore until
dissociation\textsuperscript{159-161}, which consequently enables the detection of single-base mismatches and enumeration of defective species amongst a population. To ensure that the increase in dwell time we observed is not due to MutS being trapped near the pore, we analyzed the power spectral densities of the open pore current for DNA translocations of each sample, with and without MutS, and found that all the spectra had similar noise characteristics that differed greatly from when MutS clogs the nanopore (Fig. C2).

4.3 Differentiation of structural defects

In addition to small sequence defects, \textit{de novo} gene synthesis reactions generate a range of incomplete or structurally misassembled products. We designed and assembled a set of DNA structures intended to represent the major categories of expected defects: overhangs, flaps, and Holliday junctions (Fig. 4.3a). For ease of construction, we generated populations of each of these types of defects directly from synthetic oligonucleotides (35 nt and 70 nt) by thermal annealing, without any enzymatic steps, and verified their formation using gel electrophoresis (Fig. C3). Each type of population was inserted into the nanopore setup and compared to a correctly paired 70 bp dsDNA reference molecule constructed from the same oligos. Figure 4.3b-c shows representative traces of each type of molecule translocating through the nanopore and their corresponding $t_d$ histograms compared to the 70 bp dsDNA control (see heat map contours of $\Delta I/I_o$ vs. $t_d$ in Fig. C3). The 70 bp dsDNA reference molecules showed relatively low scatter with highly repeatable transport dynamics when translocating through the nanopore, in stark contrast to the defective molecules. The overhang molecule (blue in Fig. 4.3), which is composed of a 70 nt oligo annealed to a 35 nt oligo, forms a double-stranded structure with a single-stranded end. On average, the overhang molecules migrated through the pore faster and with a lower current blockade, which agrees with its smaller average diameter compared with the
dsDNA. As seen in Fig. 4.3c, the dwell time histogram contains two populations, which can be ascribed to unbound oligos (most likely the faster population) and correctly paired overhang complexes. This result is consistent with gel electrophoresis assays, where two populations were observed (Fig. C3) and consisted of an unpaired ssDNA oligo and the correctly assembled overhang structure. The flap molecule (green in Fig. 4.3), which is composed of a 70 nt oligo annealed to a 35 nt oligo and another 70 nt oligo, contains a 70 bp dsDNA body with a 35 nt ssDNA overhang in the center. The flap sample, similarly to the overhang, yielded two distinct

FIGURE 4.3 - Gene synthesis misassemblies have unique translocation profiles. (a) Illustrations of four common products of gene synthesis by polymerase chain assembly: dsDNA (desired product), overhangs, flaps, and Holliday junctions. (b) Concatenated event traces of each product for translocation through a 2.5 nm nanopore (traces low-pass filtered at 100 kHz to show the complex number of current levels for the events). (c) Log-dwell time histograms for each product show distinct transport “fingerprints” for each species (N = number of events).
dwell time populations, which is also in agreement with gel electrophoresis results, whereby free oligos and properly constructed molecules were detected. However, the flap molecules generated on average longer dwell times than the overhang and 70 bp molecules, which was also expected given its added bulk when compared to the nanopore diameter. The last structural defect studied was a synthetically constructed Holliday junction (red in Fig. 4.3), which consists of four dsDNA strands that connect via Watson-Crick base-pairing (Fig. C3) to form a four-armed molecule. Holliday junctions are the primary elements used to create DNA origami structures and typically act as scaffolds for assembly of more complex shapes and structures. Figure 4.3b-c highlights how Holliday junctions yield long dwell time, multi-level translocation events and a large breadth in dwell time, which can be ascribed to extreme bending or breaking of the molecule to transport through the small nanopore. Many translocation events of the Holliday junction sample, as seen in Fig. 4.3b, begin with a moderate current blockade and terminate with a deeper current blockade, indicating the initial capture of one strand of the Holliday junction (moderate Δl) followed by the rupturing of the complex leading to several DNA strands occupying the pore at once (deep Δl). Using these unique translocation signatures, an algorithm could be developed to analyze and differentiate correctly assembled gene synthesis products from various misassembled products.

In addition to discriminating between these samples by dwell time differences, we quantified their capture rates using the time between successive events, or the inter-event time. When fitting inter-event time histograms of each sample to single exponential functions, we found that the normalized capture rate for the 70 bp sample was highest, the overhang and flap samples had slightly lower capture rates, and the Holliday junction sample had by far the lowest capture rate (see Table 4.1 and Fig. C4). Notably, these capture rates were normalized using
concentration units of ng/µl because each sample contains impurities in the form of free oligos or incomplete assemblies as seen in gel electrophoresis assays (Fig. C3). These impurities, which skew the true capture rate of each misassembly, make quantifying the percentage purity of a gene synthesis product (containing a mixture of these misassemblies along with correct assemblies) very difficult. Despite this, recognizing the presence of the correct gene assembly is still possible, as described in detail below.

**TABLE 4.1 - Capture Rates of Gene Synthesis Misassemblies**

<table>
<thead>
<tr>
<th>Sample</th>
<th>( R_c ) (s(^{-1}) (ng/µl(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>70 bp</td>
<td>54.6 ± 0.9</td>
</tr>
<tr>
<td>Overhang</td>
<td>24.0 ± 0.4</td>
</tr>
<tr>
<td>Flap</td>
<td>18.2 ± 0.3</td>
</tr>
<tr>
<td>Holliday</td>
<td>3.4 ± 0.1</td>
</tr>
</tbody>
</table>

**4.4 Discrimination between correctly assembled gene products and misassembled gene products**

As a case study, we synthesized a gene for HIV Protease (HIVPr) from small building blocks (oligonucleotides) using polymerase construction and amplification (PCA), wherein the oligonucleotides themselves act as both template and primers\(^{164}\) (See Section 4.6.2). HIVPr is a protease that is required for the human immunodeficiency virion to remain infectious and sequence variants of the gene can confer resistance to drug inhibitors of the enzyme.\(^{165, 166}\) The synthetic gene products were analyzed on a gel (Fig. 4.4a) and showed both correctly assembled products (524 bp) and incorrectly assembled products (such as oligos, intermediate assemblies of
shorter length, as well as longer fragments). A fraction of the as-synthesized products (6 ng/μl) were analyzed using the nanopore setup and transient blockades in current were observed as

FIGURE 4.4 - Nanopore-based monitoring of gene synthesis products for the HIV-Protease gene (524 bp) with and without additional PCR purification. (a) PAGE assay of the as-synthesized HIVPr synthetic gene and PCR-cleaned HIVPr synthetic gene. The red triangle shows the migration of the correctly assembled product band. The gel shows both correctly assembled products (524 bp) and incorrectly assembled products (such as oligos, intermediate assemblies of shorter length, as well as longer fragments) are present before and after cleanup with PCR, though the fraction of incomplete assemblies was markedly reduced after PCR cleanup. (b) Scatter plot of current blockade ΔI/I₀ and dwell time t_d for reference 500 bp dsDNA products (black), as-synthesized HIVPr synthetic genes (blue) and PCR-cleaned HIVPr synthetic gene products (red). (c) The HIV-Protease gene product as synthesized (blue) shows a wide spread in current blockade (ΔI/I₀), when compared to homogenous products (black). PCR-cleanup of the HIVPr synthetic genes (red) improved the fraction of events that resemble correctly assembled genes, but did not remove many incorrect or misassembled synthesis products. Comparison of the dwell times t_d for the homogenous products compared to the as-synthesized (blue) and PCR-purified samples (red) shows significant variation that can be ascribed to both pure and impure synthetic gene products. Light blue shaded regions depict potential thresholds for selecting correct gene assemblies using a nanopore-based on-the-fly “accept or reject” protocol.
before. Inspection of the distribution of the current blockade versus translocation time showed a significantly large spread of $\Delta I/I_0$ and $t_d$ values compared to homogenous products of similar size (500 bp, Fisher Scientific), indicating the solution contained a range of DNA sizes and structures (blue in Fig. 4.4). This result agrees well with the gel that showed both smaller molecular weight species (assembly intermediates) below the product bands as well as larger products such as dimers and trimers (Fig. 4.4a). In contrast to larger diameter solid-state nanopores, where high translocation speed and DNA self-interaction/folding induce signal scatter that precludes definitive assignment of molecular features such as length, small-diameter nanopores (< 3 nm) produce narrow distributions of $t_d$ values,$^{125}$ which enable accurate identification of errors. The increase in the range of the current blockade distribution can be ascribed to fast oligo and incomplete assembly translocations, collisions with the pore, and multi-level translocation events (Fig. 4.3b), such as flaps or Holliday junctions migrating through the nanopore. Using the signatures of homogenous populations of synthetic defects (overhangs, flaps, and Holliday junctions) and reference dsDNA with similar size (500 bp) as classifiers for detection in the synthetic gene population, direct detection of different types of structural misassemblies could be accomplished. This result is in contrast to gel electrophoresis where complex structures such as overhangs and flaps are difficult to view directly.

A common strategy to remove unwanted gene synthesis products after assembly reactions is PCR amplification, whereby the full-length product can be selectively amplified with the typical exponential profile of PCR. Incomplete products (which contain only one of the two PCR primer sequences) amplify linearly and the end effect is that the incomplete products are “diluted” in the final population of DNA molecules. However, misassembled DNA containing both primer sites will also amplify exponentially, and this approach does nothing to deal with
small sequence defects such as point mutations. The as-synthesized HIVPr gene was amplified by PCR and a fraction of PCR-cleaned up products (6 ng/µl) was loaded into the nanopore setup and measured as before. In contrast to the wide distribution of current blockades observed for the unpurified products, the purified product distribution (red data in Fig. 4.4) contains a dominant peak that closely resembles the homogenous 500 bp reference distribution as well as other smaller peaks that resemble defective species. In addition to the multi-peak distribution of current blockades for the purified synthetic genes, a large variation in translocation times was also observed. This variation could be ascribed to different types of defective species such as unpaired oligos and assembly dimers and trimmers that were unable to be cleaned by PCR.

Notably, there are two dwell time populations that emerge, which are visible on the gel as the correct HIV-protease gene assembly (524 bp) and a shorter incomplete assembly of ~125 bp (see Fig. 4.4a,c). This ability to differentiate between two DNA lengths using dwell time confirms our study from Section 2.6 where we discriminated between 100 bp and 500 bp DNA in a mixture using small nanopores.

Translocating our gene synthesis products through a nanopore after PCR clean up yields a clear population of properly assembled genes, but notably does not isolate a pure sample of the gene in the trans chamber of the nanopore setup. Using a selection protocol based upon stringent dwell time and current blockade thresholds defined by a control molecule (in this case 500 bp DNA), our nanopore system could be designed to reject any translocating molecule that does not fall within these thresholds. Following this protocol for many translocation events, it would be possible to isolate the correct gene assemblies, which could then be amplified by PCR with as few as a hundred correctly assembled molecules. Even though very short defective molecules may translocate too fast for our electronics to reject, these molecules can be effectively diluted.
by a post-nanopore PCR step, since they would not contain both primer sequences necessary for amplification.

### 4.5 Discussion

*De novo* gene synthesis has tremendous potential to design and develop combinations of new or altered genes, gene regulatory networks and pathways, as well as chromosomes and genomes for new cellular functions. While gene synthesis is an incredibly powerful tool, a significant limitation preventing large-scale and low-cost production has been the removal of unwanted side products. Since gene synthesis techniques are rapidly moving in the direction of array-based synthesis\(^\text{167}\) where oligo costs and concentrations are lower, new tools need to be developed that can operate on these scales.\(^\text{168}\)

An ultrathin, small-diameter nanopore system was employed to detect different types of synthetic gene products and characterize a conventional gene synthesis reaction at different stages (before and after PCR clean up). Using this approach, we discretize the reaction products into categories of correct and incorrect assemblies and speculate that an algorithm could be devised for isolation and amplification of correct gene assemblies. The new method described herein offers a potential solution to provide information on gene synthesis reactions with higher complexity, lower costs, and during the reactions as opposed to after the reaction is complete. This advance may facilitate monitoring of reaction kinetics, real-time tuning, and understanding of factors that drive reaction-to-reaction variations, which will further our ability to engineer gene synthesis and DNA nanotechnology complexity.\(^\text{169}\) Since nanopore systems can easily be integrated with micro- and nanofluidic systems,\(^\text{170}\) which have been used for single-molecule sorting, the possibility of constructing\(^\text{148, 149}\) and purifying molecules from a pool of mostly
defective products in an automated fashion on-chip and with concentrations far below those required for current electrophoresis techniques, makes them attractive options for future research. Thus, using nanopore-based monitoring to enhance gene synthesis product purity can facilitate the fabrication of constructs that perform new cellular functions such as complex sensing, computation, and actuation, which will have a broad impact on biomedicine and biotechnology.

4.6 Materials and Methods

4.6.1 Fabrication & Characterization of Ultrathin Nanopore Devices.

Silicon nitride (SiN) membranes were fabricated using photolithography along with dry and wet etching steps as explained in previous work.\textsuperscript{125} An additional SF\textsubscript{6} plasma-etch step was used to thin the SiN membranes to a total thickness of 15-20 nm for nanopore experiments. All experimental data was collected with a sampling rate of 4.167 MHz using the Chimera Instruments VC100 (New York, NY), which allows the detection of events as fast as 2.5 $\mu$s when low-pass filtering at 200 kHz. Nanopore data collected for the single mismatch molecules (Fig. 4.2), isolated misassembled products (Fig. 4.3), and gene synthesis products (Fig. 4.4) were low-pass filtered at 200, 400, and 300 kHz, respectively, before event analysis. Prior to addition of DNA, pores were deemed fit for translocation experiments by ramping voltage from -300 to 300 mV and observing an Ohmic relationship between current and voltage (\textit{i.e.}, linear IV curve).

4.6.2 Gene Synthesis Reactions.

The HIV protease (HIVpr) synthetic gene was assembled using 22 construction oligos (a table of oligo sequences is provided in Fig. C5) in a 50 $\mu$l PCA reaction using KOD Hot Start polymerase reagents (EMD Millipore/Novagen). Reactions contained 1x KOD buffer, 1.5 mM
MgSO₄, 0.2 mM dNTP, 1 unit KOD polymerase, 0.5 μM HIVpr-B22 primer (C CGCCTCTCCCCCGAGTTCTCCTGCAAAACCCCTCAA), 0.5 μM HIVpr-T1 primer (ATGAATCGCCAAACGTCCGCGTAGAGGCGAAATTAATACGACTCCTATAGGGAG), and 1 μM of each construction oligo. The initial PCA reaction was cycled 30 times (95°C for 20 seconds / 72°C for 30 seconds) followed by a 5-minute extension at 72°C and cooling to 12°C. An aliquot of the reaction was diluted 1:1000 prior to being used in a secondary PCR amplification. The secondary PCR amplification for the HIVpr DNA was performed in a 100 μL reaction using 1 μL of the initial PCA template, 0.5 μM of primers HIVpr-B22 and HIVpr-T1, and Phusion® High-Fidelity PCR Master Mix (New England Biolabs, M0531). This reaction was heated to 98°C for 45 seconds, then cycled 30 times (98°C-45 seconds / 55°C-45 seconds / 72°C-45 seconds), finishing with a 5-minute extension at 72°C and cooling to 4°C. Aliquots of both reactions were purified using the Qiagen Gel Extraction Kit as per manufacturer’s protocol, eluting in water. DNA was quantitated via Nanodrop spectrophotometer (Thermo Scientific). Primers were synthesized at 100 nM scale (IDT), and HPLC purified. Primer annealing for different DNA structures and mismatches was performed by mixing 2 μM of each primer, 25 mM NaCl, 5 mM Tris pH 8.0, 1 mM EDTA, and 8 mM MgCl₂. Solutions were heated to 95°C for 1 minute, followed by decreasing the temperature 5°C each minute to room temperature to anneal the primer pairs. The initial PCA reaction was Sanger sequenced and found to have a defect rate of 0.0018 per base (1 error per 555 bp). Using this error rate, the percentage of defect-free strands could be predicted as $(1-0.0018)^{524} = 39\%$ and double-stranded duplexes to be 15% (or $0.389 \times 0.389$).
4.6.3  *Gel Electrophoresis Assays.*

A 75 ng aliquot of each annealed primer mixture and individual primers was analyzed by electrophoresis on both a 20% TBE-acrylamide gel and a 15% TBE-urea acrylamide gel (Life Technologies). Samples analyzed on TBE-urea gels were heated to 70°C for 3 minutes in 2x TBE-urea sample buffer (Life Technologies) prior to loading. After electrophoresis, gels were rinsed for 30 seconds in water prior to immersion staining for 10 minutes in water + SYBR-Gold DNA stain (1:10,000 dilution; Life Technologies). Gels were rinsed for 30 seconds in water prior to exposure to UV light and photo documentation with Gel Doc camera and software.

4.6.4  *Analysis of Nanopore Data.*

All nanopore data was analyzed using Python, a custom Python-based analysis software designed by Robert Y. Henley in the Wanunu lab (www.github.com/rhenley/Pyth-Ion). Gaussian and 1D drift-diffusion fits of histograms were computed using Igor Pro (Wavemetrics).
Chapter 5 – Detection of G-Quadruplex Structures within Gene Promoter Regions

A portion of this chapter is adapted with permission from Larkin, J., Carson, S., Stoloff, D. H., and Wanunu, M. “Nanopore-Based Analysis of Chemically Modified DNA and Nucleic Acid Drug Targets.” Israel Journal of Chemistry 53.6-7 (2013): 431-441. Copyright 2013 John Wiley and Sons.
5.1 Introduction

Nanopores can be employed to study tertiary structures of DNA outside of the standard Watson-Crick double helix. One structure of particular interest is the G-quadruplex (GQ), which forms in guanine-rich sequences and is most commonly found in telomeres, single-stranded sequences of DNA at the ends of chromosomes, as well as gene promoter regions. The consistent motif of GQ formation is the stacking of G-quartets, which are stabilized by monovalent or bivalent cations, with loops of 1 to 7 bases interconnecting these guanine sections (see Fig. 5.1a-b). When it was found in the mid-1990s that cellular immortalization was catalyzed by telomerases in roughly 85% of cancer cells, scientists sought for methods of inhibiting their activity. Later, it was proposed that G-quadruplexes could also form in genomic DNA sequences and specific structures have been characterized extensively by x-ray diffraction and NMR studies. In the case of both telomeric and genomic sequences it was desired to stabilize G-quadruplexes by binding small molecules in order to repress the activity of DNA polymerases.

Several studies have shown that the formation of GQ in ssDNA can be sensed using α-hemolysin nanopores to determine properties such as folding/unfolding kinetics, cation dependencies, and GQ interactions. Shim et al. demonstrated that a particular DNA oligonucleotide (GGTTGGTGTGGTTGG) can form a GQ for various electrolytic solutions and found that the stability and folding kinetics are salt dependent (see Fig. 5.1c for current traces). One can gain structural traits by other methods such as atomic force microscopy and fluorescence resonance energy transfer (FRET) spectroscopy, but nanopores are preferable since there is no need for the labeling or immobilization of DNA to obtain information. Because
of the number of and types of nucleotides contained in loops that connect the G-quartets, there are as many as 26 theoretical conformations possible for the folding of G-quadruplexes.\(^{186}\)

It had been hypothesized that G-quadruplexes could form within coding regions of dsDNA,\(^ {188}\) which was later confirmed by the Balusubramanian group using fluorescently tagged antibodies to image internal GQ structures within chromosomes.\(^ {189}\) Given the proper conditions, \textit{i.e.} pH, salt concentration, and temperature, it is possible that solid-state nanopores could sense the presence of G-quadruplexes for specific DNA sequences in a duplex DNA setting due to differences in the diameters of double helix DNA and GQ-containing DNA.

Many putative GQs are located upstream from sequences known to express genes, which has indicated that GQ structures could play a major factor in gene expression and additionally a potential drug target to prevent overexpression. Within this chapter, studies of the
insulin-linked polymorphic region (ILPR) and vascular endothelial growth factor (VEGF), both sequences of which are known to have promoter region GQ structures, are performed using small, solid-state nanopores.

5.2 Detection of ILPR G-quadruplex structures in nanopores

The insulin-linked polymorphic region (ILPR) has been shown to affect the transcription rate of the human insulin gene based upon its location 365 bp upstream from the transcription start site.\textsuperscript{190} This promoter region is polymorphic due to the fact that it contains a variable number of a repeating 14 nucleotide sequence (\textit{i.e.}, ACAGGGGTGTGGGG), which has been shown \textit{in vitro} to form GQ structures\textsuperscript{191, 192} that promote the binding of Pur-1, a transcription factor for the expression of insulin.\textsuperscript{193} Furthermore, longer ILPR sequences (\textit{i.e.}, a greater number of 14 nt repeats) have been directly correlated to greater INS transcriptional activity, therefore increasing the expression of insulin.\textsuperscript{194} In combination with this finding, shorter ILPR sequences have been linked to genetic susceptibility to insulin-dependent diabetes mellitus, which spurred researchers to detect GQ structures in ILPR using methods such as circular dichroism,\textsuperscript{195, 196} gel electrophoresis,\textsuperscript{193} nuclear magnetic resonance,\textsuperscript{197} and optical tweezers.\textsuperscript{196} Our objective was to use solid-state nanopores to compare an 87 bp DNA fragment containing the ILPR (see Fig. 5.2a) to a control 100 bp fragment with no ILPR and hence zero probability of GQ formation.

First, we hybridized the ILPR sequence shown in Fig. 5.2a to its complementary oligomer and confirmed its formation by polyacrylamide gel electrophoresis (PAGE) as shown in Fig. 5.2b. Despite having the same length, the two oligos shown in the far left lanes run at significantly different speeds due to the inherent folding nature of guanine-rich oligomers. A
clear band that runs slower than the oligos is seen in lane 3 that corresponds to the properly hybridized 87 bp molecule containing the ILPR. Notably, this single band does have a "shadow" that runs slower indicating the presence of G-quadruplexes, which should take longer to migrate through a gel due to their larger diameter. Building upon our PAGE results, we then demonstrated that the ILPR sequence hampers DNA transport in small nanopores, which is depicted in the current traces of Fig. 5.2c. The control 100 bp molecule (red trace), shows
uniform, reproducible current blockade events, whereas the 87 bp ILPR molecule yields translocations that vary in blockade with many events having characteristically long dwell times and multiple current levels. When binning over 1,000 events of each analyte, which were run on the same 2.6 nm diameter nanopore, we see that ILPR causes a wide spread in current blockade ($\Delta I/I_o$) and a small population of long dwell times ($t_d$) not present in the 100 bp control sample (see Fig. 5.2d). The observations we see in ILPR samples are likely due to two types of anomalous events: 1) collisions (i.e., unsuccessful translocations) or 2) multi-level events, which skew our reported average current blockade. The evidence for collisions is apparent due to the significant fraction of events with fast dwell times and low current blockade as can be seen observed in Fig. 5.2c-d. We hypothesize that the multi-level events are caused by ILPR samples containing GQ that are captured, but become stuck in the pore because the width of the middle of the DNA molecule is increased by the added diameter of the intramolecular GQ. The ILPR DNA sits inside the pore until the force of the electric field unwinds the GQ, which is followed quickly by the translocation of the molecule, represented by a deep current blockade at the end of these types of events. Fitting our dwell time distributions to a single exponential model,\textsuperscript{68} we extracted values of time constant $\tau_c$ for each sample and found that ILPR had a higher value of $\tau_c$ than 100 bp (i.e., 22.6 $\mu$s compared to 13.1 $\mu$s). We theorize that this greater time constant in the ILPR sample reflects the added time required to rupture the GQ structures near or inside the nanopore.

### 5.3 Effect of oxidation on VEGF promoter GQ formation

The human vascular endothelial growth factor (VEGF) is a protein responsible for angiogenesis, or blood vessel production.\textsuperscript{198, 199} The overexpression of the gene encoding for this protein is crucial to tumor growth and survival in carcinomas.\textsuperscript{200} A proximal promoter region to VEGF, a guanine-rich 22 nt section, has been shown to form GQ structures using circular
dichroism (CD), DMS footprinting, and nuclear magnetic resonance. In addition, one in vivo study demonstrated the formation of the VEGF promoter GQ in the context of chromatin and super-coiled dsDNA in footprinting experiments. Building upon these findings, researchers explored the effect of oxidized guanine (G) bases in the destabilization of GQ by replacing G with 8-oxo-7, 8-dihydroguanine (Og) or guanindinohydantoin (Gh) at a loop or core GQ position within the VEGF gene promoter (see Fig. 5.3a-b). When substituting Og or Gh at a loop position, they discovered that VEGF still formed a GQ structure as evidenced by unchanged CD spectra and melting temperatures. When the oxidized base was located at a core GQ position, however, a red shift in CD spectra and significantly lower melting temperatures were observed indicating the presence of a less robust triplex DNA structure. These results indicated that Og or Gh located within the GQ core disables the formation of G-quartets leading to three-stranded triplex structures that are less stable than G-quadruplexes.

Expanding upon these results in collaboration with the Burrows Lab of the University of Utah, we used solid-state nanopores to probe the effect of oxidized guanine bases on the stability of the VEGF promoter sequence. For this experiment we studied one control 75 bp molecule with no putative GQ sequences and five 60 bp variations of VEGF: WT sequence (as shown in Fig. 5.3b), G → Og at loop GQ position (Og12), G → Gh at loop GQ position (Gh12), G → Og at core GQ position (Og14), and G → Gh at core GQ position (Gh14). These six analytes were first characterized by gel electrophoresis (Fig. 5.3c), which yielded a strong band for each VEGF sample located at ~60 bp when compared to a DNA ladder. Every VEGF sample (lanes 3-7 in Fig. 5.3c) had faint bands that ran faster through the gel, which are most likely excess oligomers that were unsuccessfully hybridized (see protocol for hybridization in Section 5.4). Each VEGF sample also had a faint band that migrated slightly slower than the strong band, which indicates
the presence of triplex or GQ structures within these samples. Gel electrophoresis, however, does not have the resolution to detect the difference between triplex and GQ structures, which makes the nanopore a useful tool for this purpose.

Figure 5.3d displays the results of each sample translocating a small silicon nitride nanopore (experimental conditions were $d = 2.4 \text{ nm}$, $V = 200 \text{ mV}$, $0.40 \text{ M KCl}$). The 75 bp control molecule with no probability of GQ formation has a tight, monodisperse population reflective of smooth, easily reproducible transport through the nanopore. Both the Og14 and Gh14, which have oxidized bases at the loop location, have similar contour plot shapes with dwell times ($t_d$) and current blockades ($\Delta I/I_o$) consistent with the 75 bp control molecule. The

**FIGURE 5.3** – Oxidation of guanine disrupts G-quadruplex formation. (a) The chemical structures of the DNA bases guanine (G), 8-oxo-7, 8-dihydroguanine (Og), and guanindinohydantoin (Gh). (b) The 60 nt VEGF promoter sequence studied with the guanine loop position (red) and core position (cyan) highlighted where oxidized bases Og and Gh are substituted. (c) 17% PAGE gel for the following samples: 1) 50 bp DNA ladder, 2) 60 bp control (not used), 3) WT, 4) Og12, 5) Og14, 6) Gh12, 7) Gh14. (d) Contour plots for DNA transport of VEGF samples through a $d = 2.4 \text{ nm}$ nanopore at 200 mV bias. The total number of events $n$ is displayed in each plot.
spread of each dataset, however, is much wider in the oxidized samples with longer dwell times likely caused by the unwinding of triplex DNA structures while pulled through the nanopore. When the oxidized bases are instead at the core GQ position of VEGF (i.e., samples Og12 and Gh12), we observe an even greater shift towards longer dwell times indicative of GQs being formed in these samples. Unfortunately, the WT dataset, which was the last collected on this nanopore, has noticeably faster dwell times than all other samples due to pore expansion before data collection, represented by a drastic increase in open pore current. Due to this increase in pore diameter, the WT sample is unable to be compared properly to the VEGF samples containing oxidized guanines.

5.4 Materials and Methods

5.4.1. Synthesis of modified VEGF oligomers

For details on the synthesis of the DNA oligomers used to form our 60 bp VEGF samples, see previous studies from the Burrows Lab.\textsuperscript{205}

5.4.2. Heating protocol for DNA hybridization

All GQ samples studied in this chapter were formed by the hybridization of oligomers of specified sequences using the following protocol. All oligomers were added to 1xTE in equimolar amounts, heated to 98°C for 5 minutes, then cooled back to room temperature slowly for a minimum time of 5 hr. Once brought back to room temperature, all samples were stored at 4°C until used in nanopore experiments or gel electrophoresis.
Chapter 6 – Conclusions and Future Outlook

This dissertation has shown that ultrathin, small nanopores provide smooth transport for dsDNA fragments of length spanning four orders of magnitude. In nanopores with a diameter of \( d = 2.5 – 3.0 \text{ nm} \), we discovered that both dwell time and current blockade distributions were monodisperse, with events containing a single blockade level. This finding is a vast improvement over past studies using larger nanopores within thicker membranes, which exhibited DNA transport with multiple event populations and multi-leveled translocation events.

In eliminating the DNA self-interaction present in larger nanopores, we can better define the position and speed of DNA within the pore during translocation, which could be employed for more effective studies in DNA mapping and sizing. With the efficacy of this new regime of nanopores established, we re-calibrated the transport dependence of voltage and DNA length using a physically relevant first-passage time model to fit our dwell time distributions. The effectiveness of our pores to distinguish between different DNA lengths was then tested by differentiating 100 bp from 500 bp molecules in a mixture with an accuracy >98%.

As an application of these nanopores, we then explored the effects of a newly discovered epigenetic modification, 5-hydroxymethyluracil (hmU), on dsDNA structure. We found that nanopore transport was faster and blocked less ionic current when all thymine (T) bases were oxidized to hmU, a finding that was further studied in both free solution and nanopore-constricted MD simulations. This result was unexpected since the added hydroxyl group of hmU makes the DNA base bulkier, which demonstrated that steric considerations are not the primary influence in DNA transport characteristics in this case. The MD simulations revealed that the oxidation of T to hmU causes DNA to be more flexible, which leads to a lower free energy
barrier and faster translocation for very short DNA fragments (i.e., ~100 bp). This study measured the impact of hmU on DNA transport, but our nanopore system did not have the resolution necessary for the detection of individual hmU bases. However, future nanopore sequencing systems, which currently ratchet ssDNA through nanopores using a biological motor, will find our results of increased DNA flexibility for this epigenetic base beneficial since the precise location and speed of DNA molecules is crucial in this methodology.

We demonstrated that our nanopores could detect single base mismatches when bound by a MutS repair protein, differentiate between common gene synthesis misassemblies, and detect the presence of a correct HIV gene assembly after PCR cleanup. As noted in Chapter 4, gene synthesis technology is heading toward miniaturization and microfluidic arrays in an effort to maximize the number of reactions performed simultaneously. The major difficulty in gene synthesis reactions is the presence of incorrect assemblies in the final volume that consist of both local defects (e.g., point mutations and indels) and broad scale structural defects (e.g., overhangs and Holliday junctions). The reaction volume is then PCR purified using primers designed to amplify the gene sequence, but many other fragments that contain these same primer sequences will also be incorrectly amplified. Gel electrophoresis is then performed in order to further isolate the correct gene assembly, which is then extracted and spin column purified. The current purification process in gene synthesis adds cost and time, and also limits the number of reactions that can be performed simultaneously. We envision nanopores acting as molecular filters in a microarray chip of gene synthesis reactions with one pore located at the bottom of each well. Once calibrated to a particular gene length, each nanopore could reject incorrect assemblies from passing through based on thresholds in dwell time and current blockade, therefore allowing only correct gene assemblies to translocate into collection chambers. The collection chambers, free of
defective molecules, could then be easily amplified by PCR resulting in a purified gene assembly. The proper incorporation of nanopores into gene synthesis microarrays would cut cost and time normally involved in gel electrophoresis purification and multiple PCR steps, and would solve the issues surrounding sample contamination.

Lastly, we employed nanopores for the detection of G-quadruplexes in the context of duplex DNA and confirmed that oxidized guanines shift the conformation of DNA from a GQ to a triplex DNA structure. In future studies, we aim to collect a complete set of data without pore expansion, which was the major issue in our study of oxidation in VEGF promoter regions. Since GQs must unwind before translocating small diameter nanopores, we also plan to attempt recapture experiments as in previous studies\textsuperscript{150, 207} to compare the folded and unfolded transport of these GQ structures. In early attempts, it has proven difficult to recapture such short DNA molecules (\textit{i.e.}, < 100 bp), so a lengthening of these constructs may be necessary for future studies.

In conclusion, the solid-state nanopore has proven to be a versatile single-molecule method for studying many types of biomolecules with my focus being on double stranded DNA in the context of epigenetic modifications, gene synthesis reactions, and GQ structures. With commercial nanopore devices already available for DNA sequencing,\textsuperscript{208} the future is exciting for this continuously growing field of research.
Appendices

Appendix A - Supporting Material for Chapter 2

A.1 Materials and Methods

A.1.1 Nanopore experiments.

Substrates for our solid-state nanopore membranes were formed by the thermal oxidation of silicon \(<100>\) wafers to generate a ~2.5-\(\mu\)m thick SiO\(_2\) dielectric barrier layer. Following oxidation, low-pressure chemical vapor deposition is used to deposit a 45-nm-thick low-stress silicon nitride layer. Using a combination of photolithography and dry/wet etching steps, a wafer was etched to obtain an array of 5x5 mm\(^2\) chips with freestanding square membranes at their centers (10-50 \(\mu\)m in length), and further thinned to ~25 nm using a controlled reactive ion etch process. A transmission electron microscope (JEOL 2010FEG) at high energy (200 kV) and magnification (1.5 Mx) was then used to “drill” nanopores of measureable shape and diameters in the range of 2-10 nm. Due to various factors we have found that our effective nanopore diameters did not always correspond to the TEM-based diameter, so all quoted pore diameters in the chapter were estimated from ion current measurements using a solitary model in which the effective thickness is 1/3 of the total membrane thickness (see Section A.4).\(^{209,210}\)

Prior to conducting a nanopore experiment the nanopore chips were cleaned by a 10-minute immersion in a fresh heated piranha solution (1:2 mixture of H\(_2\)O\(_2\) and H\(_2\)SO\(_4\)), followed by cooling and a copious rinse with water. The chips were then stored in DI water until an
experiment was carried out. Upon use, the chips were vacuum dried and mounted on a gasket-sealed two-chamber cell, each equipped with an electrode that is connected to a high-bandwidth amplifier. Buffered electrolyte (0.4 M KCl, buffered to pH 7.9 using 10 mM Tris and 1 mM EDTA) was added to both chambers, and the pore was evaluated by conductance measurements and by recording a time-stable DC ion current response for nonzero bias.

**A.1.2 Molecular dynamics simulations.**

A custom tclforces script was used to produce DNA displacement. Each phosphorous atom of DNA was harmonically restrained ($k_z = 69.5$ or $6.9 \text{ pN/Å}$) to an individual template particle. The $z$ coordinates of the template particles were synchronously changed according to the target DNA velocity. The initial coordinates of the template particles were that of B-form DNA. In addition to the above pulling restraints, the phosphorous atoms of the DNA were harmonically restrained ($k_{rad} = 69.5 \text{ pN/Å}$) to a cylindrical shell coaxial with the pore. Under such restraints, the DNA fragment could rotate about its axis but not stretch or move away from the center of the pore. In each production simulation, the DNA was pulled 3.4 nm through the nanopore. Three pulling velocities were used: 0.1133 nm/ns, 0.0567 nm/ns, and 0.0283 nm/ns, requiring 30, 60, and 120 ns simulation, respectively. Five independent simulations were carried out for each pulling velocity; five additional simulations were done for the 4.5 nm nanopore, where thermal fluctuations were more significant. The instantaneous force applied to all phosphorous atoms was recorded every 200 fs. The time average of the instantaneous force was used to determine the diffusion constant.
A.1.3 Bootstrapping and fit optimization.

We estimated $D$ and $v$ using a numerical maximum-likelihood procedure. We first eliminated any dwell time data points that were caused by fast collision events or falsely detected by our analysis software by setting a minimum and maximum threshold for dwell time. After these limits were incorporated, standard errors for $D$ and $v$ were estimated by bootstrapping. For each experimental dataset, we created 10,000 bootstrap samples by resampling from the original data with replacement. The estimates of $D$ and $v$ generated from these resampled datasets formed bootstrap distributions for $D$ and $v$. The distributions were approximately Gaussian, so we report the standard deviation of each bootstrap distribution as a standard error for the corresponding parameter estimate.

A.2 Effect of low-pass filter frequencies on event detection

Many nanopore experiments are conducted with high salt conditions (i.e., $>1$ M KCl or other electrolytic salt), which boosts the ionic current and reduces the relative amplitude of thermal and capacitive noise.\textsuperscript{211, 212} Since our experiments use an ionic concentration of 0.4 M KCl we need to take great care in deciding what low-pass filter to apply when analyzing our experimental data. Many factors, such as desired time resolution and signal noise, must be taken into consideration when determining what filtering should be done on nanopore current traces. The low-pass filter selected can also change the number and types of events detected based on the current amplitude $\Delta I$ of each respective event. As a demonstration of this effect, in Fig. A1 we have plotted a sample 2-second current trace from an experiment for 100 bp after applying a filter of 10 kHz (black) and 200 kHz (green). All translocation events, signified by deeper
current blockades, are detected for both filters, but a few collision events are only detected by the 10 kHz filter (*designated by red ovals*). Since these collisions are masked by the noise when filtering at 200 kHz, we effectively sift out undesirable events by simply increasing our filter frequency without the need of additional manual analysis.

**FIGURE A1** - Effect of signal bandwidth on masking DNA collision spikes. A two-second example current trace for 100 bp DNA (*V* = 200 mV, *d* = 2.9 nm) is low-pass filtered to 10 kHz (*black*) and 200 kHz (*green*). As seen by the red ovals, collisions of DNA with the pore, which yielded low-amplitude spikes, are missed by our event analysis routine when performed on data low-pass filtered at 200 kHz.
A.3 Time stability and pore-to-pore reproducibility of DNA translocation

**FIGURE A2** - A representative, continuous 40-second current trace for 500 bp transport at $V = 200$ mV through a pore of $d = 3.0$ nm (raw data downsampled to 500 kHz and low-pass filtered at 200 kHz). The vast majority of events show great uniformity in both dwell time $t_d$ and current blockade $\Delta I$. 
A.4 Determination of pore diameter and thickness based on open and blocked pore ionic current ($I_o$ and $I_b$)

To estimate the size of any nanopore (diameter and thickness), we used the following nanopore conductance model:

$$ I = \sigma V \left( \frac{4b_{\text{eff}}}{\pi d^2} + \frac{1}{d} \right) $$  \hspace{1cm} (A1)

The first term inside the parentheses represents the classical geometrical contribution to the resistance due to the length and diameter of the nanopore ($R \sim b_{\text{eff}} / d^2$). The second term is known as the access resistance, which dominates as $b_{\text{eff}}$ approaches zero and prevents the resistance from becoming zero in this limit.\cite{213,214} We first analyzed our current traces to determine the open pore current $I_o$ and the blocked pore current $I_b$ for a given pore. With these
two quantities, along with the known applied voltage $V$ and the solution conductance $\sigma$ (~50 mS/cm for 0.4 M KCl), it is possible to solve this conductance model numerically to estimate the pore diameter $d$ and the effective membrane thickness $b_{\text{eff}}$ for any experiment. Since the nanopores that we fabricated by TEM drilling are hourglass shaped, it has been determined experimentally that it is most accurate to model the membrane using one-third of the total thickness, also known as the effective thickness $b_{\text{eff}}$.\textsuperscript{209, 210} Fitting our current data to this model also served as a useful confirmation of the diameter we estimated from our TEM images. Any pore diameter cited in this paper is the value determined experimentally using Eq. A1.

### A.5 Determination of fit percentage using the 1D drift-diffusion model

When using the 1D drift-diffusion model described in the main text, we constrained the fit to a specific thickness $b$, which was calculated using Eq. A1 and the known DNA length ($b = b_{\text{eff}} + L_C$). Using this thickness constraint we obtained values for $D$ and $v$ when fitting our model to each dwell time distribution. The percentage of dwell time data that is beneath the fitting function was determined by an integration method, which estimates what fraction of our data is contained by the fit. This method was assisted by using an automated integration function in Igor Pro (WaveMetrics, Inc., Portland, OR), which computes the cumulative area integral of both the dwell time histogram and the fitting function over the entire time range of the data. By aligning these two cumulative area curves on the same plot, the percentage of events that agree with our model can be determined. This is accomplished by finding the dwell time where the experimental curve deviates from the fit curve, then dividing the area corresponding to this dwell time by the total area of the experimental histogram, which yields a fractional value representing the fit percentage. When using this method we also made certain to exclude events that occur at fast time scales outside of the fit.
A.6 Derivation of axial diffusion coefficient $D_A$

When extracting diffusion coefficients from our dwell time distributions, we relate this to a bulk value of $D_o$ for 75 bp, or one-half of a persistence length for dsDNA. Approximating DNA as a short rod-like molecule, the bulk diffusion coefficient $D_o$ can be broken down into its axial and radial components, $D_A$ and $D_R$. Since our model is one-dimensional and we are interested in the DNA translocation trajectory, only the axial component of the diffusion $D_A$ is relevant. By ignoring end-effects of the molecule, the axial component can be approximated as $D_A \sim 2D_R$.\(^{215}\) Using this relationship and the fact that $D_o = (D_A + 2D_R)/3$,\(^{76}\) we solve for $D_A$ to obtain $D_A = 3/2 \cdot D_o$. For notational simplicity, $D_A$ is referred to as $D$ in the main article.

A.7 Finite-element simulations of DNA nanopore translocation

As shown in Fig. A4, we modeled our experimental setup as two cylindrical compartments, each with a diameter of 12 µm and a height of 10 µm, connected by an hourglass shaped nanopore. By defining our geometry in this way, we have assumed that our membrane is a perfect electrical insulator, which is an acceptable simplification for materials such as SiN$_x$. The two important user-defined parameters for the nanopore geometry are the diameter $d$ and the effective membrane thickness $b_{eff}$, which defines the total membrane thickness (i.e., $3b_{eff}$) and is used to construct the hourglass shape of the pore. The diameter of the pore opening was defined as $2.5d$, which yielded an hourglass shape reasonably consistent with TEM imaging previously reported.\(^{210}\) Whenever DNA was added to the geometry of the simulation (Fig. 2.5$b,d$) it was modeled as an insulating cylinder with a diameter of 2.2 nm and height of 80 nm. We defined our mesh to be much finer inside and in the vicinity of the nanopore (meshing elements as small as 0.1 nm), and then used normal meshing settings for the rest of the geometry (elements $> 1$...
We also added boundary layers to our mesh at the edges of the nanopore, which will dampen any adverse effects of the sharp edges at the pore/compartment interface. In order to reduce the simulation time, all simulations were computed in two dimensions with a symmetry axis centered inside the pore and perpendicular to the membrane surface. Any results reported were obtained using a temperature of 21°C, a KCl concentration of 0.4 M, and an applied voltage of 200 mV.

All results were steady-state solutions of the Poisson-Nernst-Planck equations, which involve three different physics models in COMSOL. The first and simplest model to implement is the Poisson equation,

\[ \nabla^2 V = \frac{-\rho}{\varepsilon_0 \varepsilon_r} \]  

(A2)
with \( \varepsilon \) as the relative permittivity of water (which we take to be 80) and \( \rho \) as the summed charge density of K\(^+\) and Cl\(^-\) ions. This equation couples \( V \) with the concentration of ions \( C_i \) since the space charge density is \( \rho = F(C_K - C_{Cl}) \), where \( F \) is the Faraday constant. The boundary condition used in this model is zero charge on all surfaces with the exception of setting \( V = 0 \) V at the top of the cis compartment and \( V = 200 \) mV at the bottom of the trans compartment. The second physics model used is the Navier-Stokes equation,

\[
\rho_w (u \cdot \nabla)u = \eta \nabla^2 u - \nabla p + F
\]

where \( \rho_w \) is the density of water, \( u \) is the fluid velocity, \( \eta \) is the viscosity of water, \( p \) is the pressure, and \( F \) is the sum of all external forces per unit volume. The only external force that is significant in our experiment is the volume electrostatic force due to ionic charge, \( F(r, z) = -\nabla V \cdot F(C_K - C_{Cl}) \). A no slip boundary condition was enforced in this model at all surfaces. The final physics model employed is a simplified Nernst-Planck equation,

\[
\nabla \cdot (D \nabla C_i + z_i \mu F C_i \nabla V) - u \cdot \nabla C_i = 0
\]

where \( D \) is the diffusion coefficient (which we estimate as \( 2 \times 10^{-9} \) cm\(^2\)/s for K\(^+\) and Cl\(^-\)), \( z_i \) is the charge number of each ionic species, and \( \mu \) is the electrophoretic mobility defined using the Einstein-Smoluchowski relation \( \mu = D / k_b T \). This equation has been simplified by assuming that our fluid is incompressible (\( \nabla \cdot u = 0 \)) and enforcing a steady-state solution (\( \partial C_i / \partial t = 0 \)). The boundary conditions given for this model are a salt concentration of 0.4 M KCl at the top cap of the cis chamber and bottom cap of the trans chamber, and no flux of ions at any surface in the...
geometry. This model incorporates all forms of motion for the ionic species in a nanopore experiment: diffusion, electro-migration, and convection.

A.8 Agarose gel of DNA samples

FIGURE A5 - An agarose gel of DNA lengths 1 kbp – 20 kbp. To ensure that our long DNA samples ($N > 1$ kbp) are pure, we prepared a 1% agarose gel with a 1 kb ladder as a control and ran it for 100 minutes at 100 V. After staining the gel with ethidium bromide for ~30 min, it was rinsed with water, then imaged. Each sample is labeled in the gel as follows: (1) 1 kb DNA ladder, (2) 900 bp (not used in experiments), (3) 1 kbp, (4) 3.5 kbp, (5) 6 kbp, (6) 10 kbp, (7) 20 kbp. As observed in the gel results, each DNA sample shows one clear band that corresponds to the correct length according to the 1 kb ladder.
A.9 Transport Time vs. DNA Length Studies

**Figure A6** - Continuous current traces for 100 bp, 1 kbp, 3.5 kbp, 10 kbp and 20 kbp DNA for nanopores with \( d = 2.8-3.0 \) nm. The mean open pore current ranges from 0.5-0.7 nA due to small variances in membrane thickness. Despite the open pore current being stable, it would occasionally increase \((i.e., < 10\%)\) over the course of an experiment due to solution evaporation or pore expansion, which required the vast majority of datasets to be collected in experiments of less than 20 minutes in duration. Longer DNA lengths had lower sample concentrations to avoid the clogging of our pores due to the longer transport times required for translocation.
Figure A7 - Current traces of concatenated, consecutive events for the data shown in Fig. A6 overlaid with analysis fits obtained by OpenNanopore software. The traces show that we detect a vast majority of translocations with a single current blockade level with the occasional fast collision or multi-level translocation.
Figure A8 - (a) Scatter plots of $\Delta I/I_0$ vs. $t_d$ for 11 DNA fragments through different nanopores, all in the range of $d = 2.8-3.0$ nm. Single event populations are seen for $N < 6$ kbp, while additional populations are seen for DNA lengths $N > 6$ kbp. The fast ($t_d < 100$ µs) events for long DNA are attributed to DNA collisions with the pore, caused by an increased barrier for a long DNA end to find the pore mouth. Further, we speculate that the additional slow population that forms for $N > 6$ kbp is due to DNA shearing or some other disruptive process by the nanopore. Contamination was ruled out by observing a single band in the gel electrophoresis of these fragments (see Fig. A5), and we do not have any proof for our claim of DNA shearing (apart from the faster timescale of these events than expected). Since this intermediate population contains a very small minority of the total events, we claim that the slowest population corresponds to DNA translocations. (b) A plot of $\Delta I/I_0$ vs. $N$ for the pores used in (a), demonstrating the similar diameters used in the experiments. When fitting to a horizontal line, we find a mean $\langle \Delta I/I_0 \rangle$ value of 0.57. (c) A plot of $v$ vs. $N$ shows a decreasing velocity with contour length. However, in the plot of $D$ vs. $N$ no definite trend is discernable. The anomalously high value of $D$ for $N = 20$ kbp ($D = 270$ nm$^2$ µs$^{-1}$) is not shown in the plot in order to show the trend for most of the other DNA lengths.
### TABLE A1 - Dwell Time vs. DNA Length Data

<table>
<thead>
<tr>
<th>$N$ (bp)</th>
<th>$n_{total}$</th>
<th>$t_d$ (± err)</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>3,315</td>
<td>17 ± 1 µs</td>
</tr>
<tr>
<td>50</td>
<td>2,121</td>
<td>49 ± 2 µs</td>
</tr>
<tr>
<td>100</td>
<td>911</td>
<td>135 ± 8 µs</td>
</tr>
<tr>
<td>200</td>
<td>679</td>
<td>153 ± 4 µs</td>
</tr>
<tr>
<td>250</td>
<td>806</td>
<td>169 ± 3 µs</td>
</tr>
<tr>
<td>500</td>
<td>2,593</td>
<td>890 ± 20 µs</td>
</tr>
<tr>
<td>1,000</td>
<td>552</td>
<td>3.62 ± 0.14 ms</td>
</tr>
<tr>
<td>3,500</td>
<td>590</td>
<td>8.51 ± 0.25 ms</td>
</tr>
<tr>
<td>6,000</td>
<td>283</td>
<td>39.1 ± 2.2 ms</td>
</tr>
<tr>
<td>10,000</td>
<td>424</td>
<td>67 ± 2 ms</td>
</tr>
<tr>
<td>20,000</td>
<td>419</td>
<td>80 ± 4 ms</td>
</tr>
</tbody>
</table>
B.1 Supplementary figures

**FIGURE B1** - Sequence and restriction map of the amplicon used in this study.
FIGURE B2 - Nanopore translocation dataset for thymine-modified DNA samples obtained using a nanopore smaller in diameter than the one reported in the main text (Fig. 3.2). Dwell time and fractional current blockade histograms for modified DNA measured using a 2.6 nm diameter nanopore ($V = 200$ mV, $0.4$ M KCl, pH 7.9, $T = 21^\circ$C). All trends observed in this set of measurements correlate with the data reported in the main text. The order of data collection was t-DNA, h-DNA, p-DNA, which differs from the dataset shown in Fig. 3.2 (t-DNA, p-DNA, h-DNA).
FIGURE B3 - The effect of point hmU modifications on simulated DNA. Differences in the standard deviation of the intra-base pair (left) and inter-base pair (right) parameters for the hmU-modified DNA relative to the values obtained for h-DNA. In contrast to the simulations reported in the main text Figure 3.4, here we analyze free-equilibration trajectories of DNA constructs containing point hmU mutations (see Section B.3 below). The relative difference in standard deviation

\[ \Delta \sigma = 100 \times \left( \frac{\sigma_h - \sigma_t}{\sigma_t} \right) \]

where \( \sigma_h \) and \( \sigma_t \) are the standard deviations of a parameter in h-DNA and t-DNA, respectively, averaged over the time series, and then over base pairs or base pair steps containing the point hmU modifications. The error was calculated by splitting MD trajectories of the same chemical system into thirty blocks and calculating the standard deviation within each block for each base pair or step. The standard error of these 30 estimates was then propagated to reflect averaging over all base pairs or steps containing point hmU modifications.

FIGURE B4 - The effect of point phmU modifications on simulated DNA flexibility. Differences in the standard deviation of the intra-base pair (left) and inter-base pair (right) parameters for the phmU-modified DNA relative to the values obtained for h-DNA. The relative difference in standard deviation

\[ \Delta \sigma = 100 \times \left( \frac{\sigma_p - \sigma_h}{\sigma_h} \right) \]

where \( \sigma_p \) and \( \sigma_h \) are the standard deviations of a parameter in p-DNA and h-DNA, respectively, averaged over the time series, and then over base pairs or base pair steps containing the phmU modifications. The error was calculated by splitting MD trajectories of the same chemical system into thirty blocks and calculating the standard deviation within each block for each base pair or step. The standard error of these 30 estimates was then propagated to reflect averaging over all base pairs or steps containing phmU modifications.
B.2 Abnormal conformations of dsDNA encountered during MD simulations

During the free equilibration simulations of DNA variants, DNA was seen to assume abnormal conformations in three separate instances. Two of these abnormal conformations occurred during a single simulation of p-DNA, and another occurred during a simulation of t-DNA.

In one of the three simulations of the t-DNA system, the DNA backbone assumed a sharp, 90-degree bend at base 8 (T·A) and 9 (C·G) after ~ 90 ns of equilibration. The DNA maintained this abnormal conformation for the remaining 30 ns of the equilibration and the base pairing pattern was not broken. The presence of the kink considerably increased the twist (steps 8, 9), propeller (base 8), and slide (steps 7, 8, 9) parameters of the base pairs involved. As this abnormal conformation was not observed in the previous microsecond simulations of canonical DNA, we assumed it to be a rare event and excluded this part of the equilibration trajectory from our analysis.

During the free equilibration simulation of p-DNA, two abnormal conformations occurred at two different sites on the molecule. Base pair 25 (A·hmU base pair) was no longer hydrogen bonded after $t = 89$ ns, which resulted in a staggered stacking of the bases. The rupture widened so that base 26 (phmU·A base pair) was no longer hydrogen bonded after $t = 104$ ns. Another DNA base pair ruptured in the same p-DNA simulation at $t = 85$ ns at base 14 (A·hmU base pair), but the hydrogen bonds reformed by $t = 95$ ns. At the kinks, DNA was no longer in a near-equilibrium configuration. Furthermore, the base pair pattern was broken, which is incompatible with the geometry assumed by the DNA conformation analysis algorithm. Hence,
we excluded from our analysis the last 35 ns of the free equilibration trajectory of the p-DNA system where the two kinking events were observed.

**B.3 Simulations of DNA variants containing point hmU modifications**

To measure the structural effect of point modifications of T into hmU, we created four copies of the t-DNA system, and in each we modified one T (at position 22, 24, 25 and 26) into hmU. We then simulated these singly modified constructs in free solution for 120 ns each. Following that, we measured the structural parameters for only the modified base pairs or steps (Fig. B3). The increased fluctuations of both intra- and inter-base pair parameters (with the exception of shear which shows no change) suggests a greater local flexibility of an hmU-containing DNA fragment in comparison to unmodified DNA.

**B.4 Determination of local flexibility of DNA for individual phmU modifications**

Since the phmU modifications were not as dense as hmU modifications in the constructs used for our free equilibrium simulations, we were able to use our original free equilibration trajectories to determine the structural effects of point modification of hmU into phmU. From the simulations of p-DNA in free solution, we measured the structural parameters of the modified bases, and compared them to the corresponding parameters measured in the h-DNA construct. The results of this analysis (Fig. B4) indicate smaller fluctuations of all intra-base pair parameters in the phmU-containing DNA fragments in comparison to the hmU-containing fragments. The effect of phmU modification on the fluctuations of the inter-base pair parameters was mixed: slide, roll, and twist increased whereas shift, tilt, and rise decreased. We interpret these results as an indication of an overall decreased flexibility of phmU-containing DNA in comparison to DNA carrying only hmU modifications.
B.5 Determination of ionic current and molarity from MD simulations

To compute the ionic current blockade, the open nanopore system (which had no DNA in it) was first simulated for 45 ns at a 200 mV transmembrane voltage. Following that, five independent simulations of 60 ns were performed using different random seeds to measure the open pore current.

For all nanopore systems, the instantaneous ionic current was computed as

$$I(t) = \frac{1}{\Delta L_z} \sum_i q_i [z_i(t + \Delta t) - z_i(t)],$$

where $q_i$ and $z_i$ were the charge and z-coordinate of ion $i$, respectively, and $\Delta t$ was 10 ps. The current was measured in a 2 nm thick slab (i.e., $L_z = 2$ nm) centered on the midplane of the pore. The average current was calculated by taking the mean of the current series. To calculate the error in the current values, the current data was split into 20 ns blocks from which a mean value was computed, and then the mean and standard error from these blocks were calculated.

Molarity was calculated by counting the average number of K$^+$ or Cl$^-$ ions in a 3 nm tall cylindrical shell centered on the pore axis and then dividing by the volume of the shell. The ion molarity difference between the h- or p-DNA and t-DNA systems, Fig. 3.5, was computed by subtracting the local ion molarity in the unmodified DNA system from the local ion molarity in the modified DNA system.
Appendix C - Supporting Material for Chapter 4

**FIGURE C1** - Gel electrophoresis reveals high affinity binding of MutS to DNA mismatches. (a) Sequences of 70 bp “correctly matched” oligos and 70/71 bp “mismatch” duplex. The single-base mismatch is highlighted in yellow and located 35 bp from either end. (b) Gel shift electrophoresis assay of the correctly matched and mismatch products incubated with various amounts of MutS. The red triangle indicates the migration of the 70 bp and 70/71 bp product bands and the green triangle indicates the migration of the MutS + dsDNA complex. As the concentration of MutS was increased, the intensity of the complex band also increased, indicating a significant fraction of the mismatch products were bound to the MutS. In contrast, little to no binding is observed for the correctly matched products.
FIGURE C2 - Power spectral density (PSD) plots based on ~ 500 ms of open pore current data for each sample of 70 bp (pink), 70/71 bp (light blue), 70 bp + MutS (red), 70/71 bp + MutS (blue), and MutS only (green) at 200 mV in small silicon nitride nanopores. In the case of DNA incubated with MutS, each PSD had a notable spike at a frequency of ~ 70,000 Hz not present in the experiment without MutS, which is due to environmental noise from electronics. We also produced a PSD plot for when the MutS protein clogs a small nanopore in the absence of DNA (green), which was accompanied by a 61% reduction in the pore conductance (i.e., 720 pA to 280 pA). This PSD has noticeably higher noise at all frequencies, which indicates that MutS is not stuck to the pore in our DNA/MutS experiments.
FIGURE C3 - Gel and nanopore analysis of common gene synthesis misassemblies. (a) Sequences of 35 nt and 70 nt oligos used to create different types of defective structures (overhang, flap and Holliday junction). Seven oligos were used and bases paired with other oligos are color-coded. (b) Gel shift electrophoresis assay of the different assembled products and oligos. (c) Heat map contour plots of $\Delta I/I_0$ vs log $t_d$ for different assembled products translocating through a nanopore.
FIGURE C4 - Inter-event time histograms of gene synthesis misassemblies. Fitting inter-event time histograms to single exponential functions allows the extraction of capture rate $R_C$ for each sample as displayed in the Table 4.1. All capture rates were normalized for sample concentration as measured by spectrophotometer. The Holliday junction was found to have a capture rate an order of magnitude lower than the 70 bp sample, and the overhang and flap molecules had comparable $R_C$. 
<table>
<thead>
<tr>
<th>Oligo Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIVprNOGAP-i1</td>
<td>ATGAATCGCCAACGTCCGGGAGTGGGGAATTATAACGACTCCTTTaggAG</td>
</tr>
<tr>
<td>HIVprNOGAP-b2</td>
<td>ATTTCTAGAGGGAACCGTTGGTGCTCCCCTAGTAGTGAAGCTGATTATTTTCG</td>
</tr>
<tr>
<td>HIVprNOGAP-i3</td>
<td>ACCAGAACGGTTTCCCTCAGAAAAATTTGTGTTTAACTTTAAGAAGGAGATATACCATG</td>
</tr>
<tr>
<td>HIVprNOGAP-b4</td>
<td>AGGTTAACCTCGGCCACGACATGTTATCTCCTTCTTAAAGTTAAACAAATT</td>
</tr>
<tr>
<td>HIVprNOGAP-i5</td>
<td>TCTGGTCCGCAGGTTTTGCTGTGGGACGGTTCGCGTGG</td>
</tr>
<tr>
<td>HIVprNOGAP-b6</td>
<td>AGCTGACCACAATTTTGGATGTAACCGGAGCCGGACGCTGCCAC</td>
</tr>
<tr>
<td>HIVprNOGAP-i7</td>
<td>TTAATCAAATTTGTTGTCAGCTCAAAAAGACGGCTGCTGGACAC</td>
</tr>
<tr>
<td>HIVprNOGAP-b8</td>
<td>CGGTGTCGTCCCGACCCGTGTCACAGCCGCTTCTTTT</td>
</tr>
<tr>
<td>HIVprNOGAP-i9</td>
<td>GGGTGCCGACGACCCGTTCTGGAAGAGATGTCTCTCGG</td>
</tr>
<tr>
<td>HIVprNOGAP-b10</td>
<td>ATCAATTTGGGCTTCAACGCCGAGAGACATCTCTCTCAAGAA</td>
</tr>
<tr>
<td>HIVprNOGAP-i11</td>
<td>GTGCGTGGAGACCCGAAAATGATGCTGCGGTATCGCAGGCTTTC</td>
</tr>
<tr>
<td>HIVprNOGAP-b12</td>
<td>TCTGCTGGATCGAGCAACTTTTAGAAGACCGACAGATACGGC</td>
</tr>
<tr>
<td>HIVprNOGAP-l13</td>
<td>ATCAAAGGTTGGTCAAGCCAAGATCCTAGATCGAAATCTGCGCC</td>
</tr>
<tr>
<td>HIVprNOGAP-b14</td>
<td>AGTACCCGTACCGATGCTTTTGTGCGCGATATTTTCTGAAGGA</td>
</tr>
<tr>
<td>HIVprNOGAP-l15</td>
<td>ACAAGGCGATCGGTACCGTACTGTTGTCGCGACCCCGGT</td>
</tr>
<tr>
<td>HIVprNOGAP-b16</td>
<td>GGAGGTTGGAGCCCGATATGTTAAACCGGGTGCGAACCAAC</td>
</tr>
<tr>
<td>HIVprNOGAP-l17</td>
<td>TAACATCATCGTGGCAAACCTCTCTCAACCCAAGATCGGTTGCC</td>
</tr>
<tr>
<td>HIVprNOGAP-b18</td>
<td>TGGCAACACCCGAATTTCAAGGGTGCAACCGATCGGTTGA</td>
</tr>
<tr>
<td>HIVprNOGAP-l19</td>
<td>CTGAACTTCCGGTTGGCCACGGTTGCTGTAATACTAGCATAACC</td>
</tr>
<tr>
<td>HIVprNOGAP-b20</td>
<td>GACCCGGTTTAAAGGCACCACCGAGGGTTATGCTGTTTAATACAGCAAC</td>
</tr>
<tr>
<td>HIVprNOGAP-l21</td>
<td>CTTGGGCGCTTCAACCGGTCTTTAGGGGTTTTTTCGTTGAAGGAG</td>
</tr>
<tr>
<td>HIVprNOGAP-b22</td>
<td>CGCCCTCTCCCGCGAGTTCCTTCTTCACGAAAAACCCTCAA</td>
</tr>
</tbody>
</table>

**FIGURE C5** - Oligo pool information for HIVPr gene assembly. List of 22 construction oligos with names and sequences used to construct the HIVPr gene.
Publications

The following articles were authored during my dissertation studies.

Peer-reviewed Articles:


Review Articles:


Permissions

The following documents are permissions from publishers to use article main texts, figures, and tables for this dissertation. No permission was necessary from ACS for reuse in a dissertation.
ELSEVIER LICENSE TERMS AND CONDITIONS

Jan 11, 2016

This is an Agreement between Spencer Carson ("You") and Elsevier ("Elsevier"). It consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier
Elsevier Limited
The Boulevard, Langford Lane
Kidlington, Oxford, OX5 1GB, UK

Registered Company Number
1982084

Customer name
Spencer Carson

Customer address
64 Dartmouth St
BELMONT, MA 02478

License number
3785941181893

License date
Jan 08, 2016

Licensed content publisher
Elsevier

Licensed content publication
Biophysical Journal

Licensed content title
Smooth DNA Transport through a Narrowed Pore Geometry

Licensed content author
Spencer Carson, James Wilson, Aleksei Aksimentiev, Meni Wanunu

Licensed content date
27 January 2015

Licensed content volume number
108

Licensed content issue number
2

Number of pages
1

Start Page
331a

End Page
None

Type of Use
reuse in a thesis/dissertation

Portion
full article

Format
both print and electronic

Are you the author of this Elsevier article?
Yes

Will you be translating?
No

Order reference number
carsonthesisdissertation

Title of your thesis/dissertation
Regulating DNA Transport through Solid-State Nanopores to Detect Base Modifications and Analyze Gene Synthesis Reactions

Expected completion date
Feb 2016

Estimated size (number of pages)

https://s100.copyright.com/MyAccount/viewPrintableLicenseDetails?ref=058b722a-5caee-d3ea-be7f-19d9d19845b2
INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgment: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
   "Reprinted from Publication title, Vol./Edition number, Author(s), Title of article / title of chapter. Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]," Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article. Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion.
discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. Posting licensed content on any Website: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the home page of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu. Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. For journal authors: the following clauses are applicable in addition to the above:

Preprints:
A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.). Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies.

Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement

- after the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:
discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. Posting licensed content on any Website: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com. Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu. Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. For journal authors: the following clauses are applicable in addition to the above:

Preprints:
A preprint is an author’s own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.). Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
directly by providing copies to their students or to research collaborators for their personal use

- after the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:
- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (JPA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJA's as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and coursework programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. For book authors the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institution's repository.

19. Thesis/Dissertation: if your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for ProQuest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJA's as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the
licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.
CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee. Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. Other Conditions:

v1.8

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.
OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Jan 11, 2016

This is an Agreement between Spencer Carson ("You") and Oxford University Press ("Oxford University Press"). It consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

<table>
<thead>
<tr>
<th>License Number</th>
<th>3784270021186</th>
</tr>
</thead>
<tbody>
<tr>
<td>License date</td>
<td>Jan 08, 2016</td>
</tr>
<tr>
<td>Licensed Content</td>
<td>Oxford University Press</td>
</tr>
<tr>
<td>Publisher</td>
<td>Nucleic Acids Research</td>
</tr>
<tr>
<td>Licensed Content</td>
<td>Hydroxymethyluracil modifications enhance the flexibility and hydrophilicity of double-stranded DNA:</td>
</tr>
<tr>
<td>Publication</td>
<td></td>
</tr>
<tr>
<td>Licensed Content Title</td>
<td></td>
</tr>
<tr>
<td>Licensed Content Author</td>
<td>Spencer Carson, James Wilson, Aleksei Aksimentiev, Peter R. Weigele, Meni Warunu</td>
</tr>
<tr>
<td>Licensed Content Date</td>
<td>11/17/2015</td>
</tr>
<tr>
<td>Type of Use</td>
<td>Thesis/Dissertation</td>
</tr>
<tr>
<td>Institution name</td>
<td>None</td>
</tr>
<tr>
<td>Title of your work</td>
<td>Regulating DNA Transport through Solid-State Nanopores to Detect Base Modifications and Analyze Gene Synthesis Reactions</td>
</tr>
<tr>
<td>Publisher of your work</td>
<td>n/a</td>
</tr>
<tr>
<td>Expected publication date</td>
<td>Feb 2016</td>
</tr>
<tr>
<td>Permissions cost</td>
<td>0.00 USD</td>
</tr>
<tr>
<td>Value added tax</td>
<td>0.00 USD</td>
</tr>
<tr>
<td>Total/Total</td>
<td>0.00 USD</td>
</tr>
</tbody>
</table>

Terms and Conditions

STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL

1. Use of the material is restricted to the type of use specified in your order details.
2. This permission covers the use of the material in the English language in the following territory: world. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12.
3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.
4. No alteration, omission or addition is made to the material without our written consent. Permission must be re-cleared with Oxford University Press if/when you decide to reprint.
5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society must be included in the credit line.
6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from

https://e102.copyright.com/MyAccount/viewPrintableLicenseDetails?ref=021433a-670a-48f2-b8db-fb83dd2007f1d  

Page 1 of 2

108
7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.
8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.
9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.
10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC’s Billing and Payment terms and conditions.
11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
12. Other Terms and Conditions:
v1.4

Questions? customer.cares@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.
This Agreement between Spencer Carson ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

<table>
<thead>
<tr>
<th>License Number</th>
<th>3786101472356</th>
</tr>
</thead>
<tbody>
<tr>
<td>License date</td>
<td>Jan 11, 2016</td>
</tr>
<tr>
<td>Licensed Content Publisher</td>
<td>John Wiley and Sons</td>
</tr>
<tr>
<td>Licensed Content Publication</td>
<td>Israel Journal of Chemistry</td>
</tr>
<tr>
<td>Licensed Content Title</td>
<td>Nanopore-Based Analysis of Chemically Modified DNA and Nucleic Acid Drug Targets</td>
</tr>
<tr>
<td>Licensed Content Author</td>
<td>Joseph Larkin, Spencer Carson, Daniel H. Stoloff, Meni Wanunu</td>
</tr>
<tr>
<td>Licensed Content Date</td>
<td>Jun 5, 2013</td>
</tr>
<tr>
<td>Pages</td>
<td>11</td>
</tr>
<tr>
<td>Type of use</td>
<td>Dissertation/Thesis</td>
</tr>
<tr>
<td>Requestor type</td>
<td>Author of this Wiley article</td>
</tr>
<tr>
<td>Format</td>
<td>Print and electronic</td>
</tr>
<tr>
<td>Portion</td>
<td>Full article</td>
</tr>
<tr>
<td>Will you be translating?</td>
<td>No</td>
</tr>
<tr>
<td>Order reference number</td>
<td>carsonthesis</td>
</tr>
<tr>
<td>Title of your thesis / dissertation</td>
<td>Regulating DNA Transport through Solid-State Nanopores to Detect Base Modifications and Analyze Gene Synthesis Reactions</td>
</tr>
<tr>
<td>Expected completion date</td>
<td>Feb 2016</td>
</tr>
<tr>
<td>Expected size (number of pages)</td>
<td>120</td>
</tr>
<tr>
<td>Requestor Location</td>
<td>Spencer Carson</td>
</tr>
<tr>
<td></td>
<td>64 Dartmouth St</td>
</tr>
</tbody>
</table>

BELMONT, MA 02478
United States
Attn: Spencer Carson

Billing Type
Invoice

Billing Address
Spencer Carson
64 Dartmouth St

BELMONT, MA 02478
TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.rightslink.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, and any CONTENT (PDF or image file) purchased as part of your order, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,
You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY
FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York.
York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

**WILEY OPEN ACCESS TERMS AND CONDITIONS**

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

**The Creative Commons Attribution License**

The [Creative Commons Attribution License (CC-BY)](https://creativecommons.org/licenses/by/4.0/) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-commercial reuse.

**Creative Commons Attribution Non-Commercial License**

The [Creative Commons Attribution Non-Commercial (CC-BY-NC) License](https://creativecommons.org/licenses/by-nc/4.0/) permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. (see below)

**Creative Commons Attribution-Non-Commercial-NoDerivs License**

The [Creative Commons Attribution-Non-Commercial-NoDerivs License (CC-BY-NC-ND)](https://creativecommons.org/licenses/by-nc-nd/4.0/) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

**Use by commercial "for-profit" organizations**

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee. Further details can be found on Wiley Online Library [http://olabout.wiley.com/WileyCDA/Section/id-410895.html](http://olabout.wiley.com/WileyCDA/Section/id-410895.html)

*Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.
References


