DNA Labels for Improved Detection and Capture with Solid-State Nanopores

uOttawa

Philipp Karau

Ottawa-Carleton Institute for Physics

University of Ottawa

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

2018

Abstract

Nanopores have emerged as a simple but effective tool to investigate the behavior of polymers in solution. They have shown great potential to simplify expensive and time consuming procedures like DNA sequencing, protein detection, and disease biomarker detection. With the development of in situ fabrication of solid-state nanopores by controlled breakdown (CBD) of a dielectric material, nanomanufacturing of nanopore-based technologies became feasible. However, there are still a lot of challenges to overcome for these applications to become reality. One of the major problems with solid-state nanopores is the rapid passage time of analytes going through the pore, complicating detection and reliable identification of molecules. In this thesis molecular structures are proposed that increase passage times due to increased interactions between analyte and pore wall, and at the same time increase signal amplitude due to increased blockage of the pore. These structures are short, branched DNA molecules that were assembled with built-in modifications and matching sequences to assume their structure. Nanopore experiments reveal that these structurally defined DNA produce higher detection rates than their linear DNA counterparts, making them better candidates for labels in single-molecule sensing experiments.

Statement of Originality

I hereby declare that, to the best of my knowledge, the presented work in this thesis is entirely original. Research was conducted under the supervision of Professor Vincent Tabard-Cossa in the Centre for Interdisciplinary Nanophysics, Department of Physics at the University of Ottawa.

I, Philipp Karau, contributed the large majority of the work presented in this thesis. This includes performing the experiments, analyzing the data, and writing the manuscript and this thesis. Experiments were outlined and discussed together with Professor Vincent Tabard-Cossa.

Chapter 3 comprises a manuscript that will be submitted for publication.

The nanopore fabrication setup was designed by Dylan Gunn and Matthew Waugh. LabView routines for pore fabrication are versions of code by Mathieu Gibeault. Analysis software tools were developed by Kyle Briggs and modified for the purposes of this thesis by me.

Acknowledgements

First, I would like to thank Prof. Vincent Tabard-Cossa from the Department of Physics at the University of Ottawa. He gave me the great opportunity to come to Canada and study in his lab. He supported me not only financially but also scientifically by giving constructive feedback, where necessary, and by pushing this project in the right direction, when needed.

I want to thank all the former and present members of the lab for countless discussions and all the support I got from them; specifically: Samuel Berryman, Kyle Briggs, Dr. Autumn Carlsen, Martin Charron, Dylan Gunn, Simon King, Zachary Roelen, Dr. Daniel Tessier, Matthew Waugh. Additionally I want to thank Eric Beamish, Dr. Ali Najahi Sohi, and Dr. Radin Tahvildari of the Godin lab for useful feedback on matters related to the work presented in this thesis.

I also want to thank Prof. John Pezacki, Dr. Mirka Strmiskova, and Dana Danielson from the Department of Chemistry and Biomolecular Sciences at the University of Ottawa for their collaboration and useful discussion with regards to click reactions.

Finally, I would like to thank my parents, Anke Karau and Jörg Karau, and my sister Nina Karau for their support in all my endeavors and for always being there for me. Last, but definitely not least, I want to specially thank my girlfriend Michelle Mensing for always giving me a goal to work towards.

Contents

List of Figures

List of Tables

Table of Acronyms and Abbreviations

SiN_x	→	Silicon Nitride
KCl	→	Potassium Chloride
LiCl	→	Lithium Chloride
CBD	→	Controlled Breakdown
DNA	→	Deoxyribonucleic acid
ssDNA	→	single stranded DNA
dsDNA	→	double stranded DNA
bp	→	basepair
kbp	→	kilobasepair
nt	→	nucleotide
oligo	→	Oligonucleotide (short strands of single stranded DNA)
Ag/AgCl	→	Silver / Silver Chloride
I-V curve	→	Current vs. Voltage curve
nm	→	Nanometer (10^{-9} m)
PSD	→	Power Spectral Density
RMS	→	Root-mean-square noise
PDMS	→	Polydimethylsiloxane
CUSUM	→	Cumulative Sum algorithm
MOSAIC	→	Modular Single-Molecule Analysis Interface algorithm
ADePT	→	Algorithm for Decision-making after Pilot and Feasibility Trials
nM	→	Nanomolar (10^{-9} moles/L)
SNR	→	Signal-to-Noise Ratio

Chapter 1

Physical Principles

1.1 Nanopores

Nanopores are nanometer sized holes in thin membranes, made of dielectric material (e.g. silicon nitride (SiN_x), glass). They connect two reservoirs of electrolyte solution (e.g. KCl, LiCl). There are two main classes of nanopores. The first one is biological pores. These mainly consist of a protein embedded into a lipid bilayer separating the two reservoirs [1]. Alternatively pores can be designed using DNA origami structures embedded in lipid bilayers or other support structures [2]. Biological pores have the advantage of structural consistency and site directed mutagenesis allowing for precise control over the pore geometry and chemistry. Despite this the life time of biological pores is limited due to the instability of the supporting membrane to pressure differences. These pores also cannot tolerate wide working conditions (pH, temperature, salt concentration of the surrounding medium, voltage). Despite these there is a range of applications for biological pores including DNA sequencing [3].

The second class of pore are solid-state nanopores. These pores give access to a broader range of pore sizes and lengths. They are generally more stable than their biological counterpart and have a broader range of available geometries (asymmetrical pores, nanopipettes, ultra-thin pores in 2D-materials etc.). Their main disadvantage is the fast translocation times for analytes due to a lack of interaction between the pore wall and the analyte requiring more expensive and complicated electronics for sensing. Fabrication of these pores can be achieved in different ways (see section 1.2). The working principle of a nanopore goes back to that of a Coulter counter patented by Wallace H. Coulter in 1953 [4]. A schematic visualization of the principle is given in figure 1.1. From the recorded

current blockages, information about the molecule (e.g. charge, length, conformation, etc.) can be inferred as will be explained in the next sections.

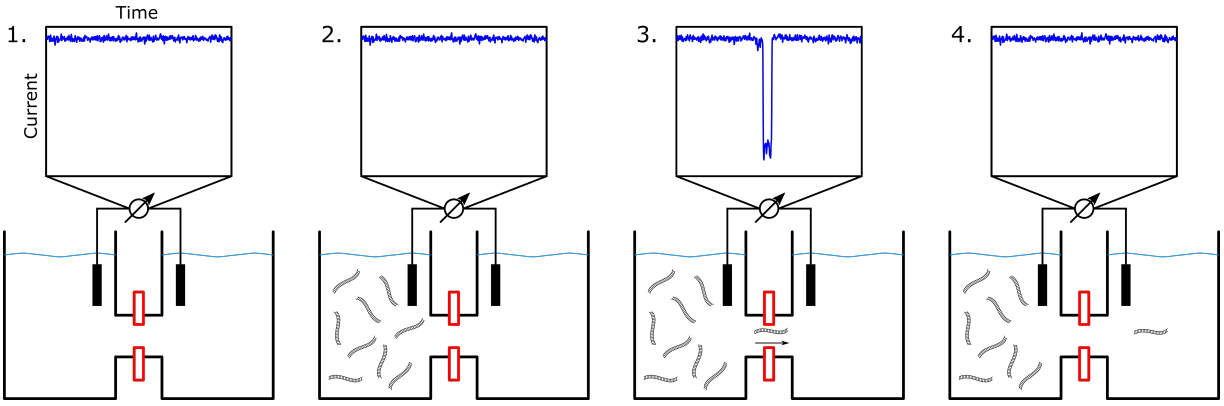

Figure 1.1: *Working principle of the Coulter counter. 1. A small channel (in this case a nanometer sized hole in a membrane) connects two reservoirs with electrolyte solution. 2. A sample of macromolecules (proteins, DNA, etc.) is injected into one of the reservoirs. 3. A constant voltage is applied across the two reservoirs inducing electrophoretic and electroosmotic flow of ions through the channel. If a macromolecule enters the channel it physically blocks a part of the ions flowing through the channel resulting in a drop in the measured ionic current. 4. When the molecule exits the channel the ionic current jumps back to the normal level.*

1.2 Controlled Breakdown

Traditionally nanopores are fabricated using a transmission electron microscope (TEM) or other high energy particle beams [5]. This procedure is very time and work intensive and needs expensive equipment as well as specially trained personnel to perform. In recent years it has been demonstrated that nanopores can also be fabricated by controlled breakdown (CBD) of a thin dielectric membrane [6], allowing for easy and fast access to nanopores without these constraints. The silicon chip containing the free standing SiN_x membrane is mounted in a fluidic cell as shown in figure 1.2. Both reservoirs of the cell are filled with electrolyte solution. A voltage is applied across the membrane via Ag/AgCl electrodes suspended in both reservoirs. It can be observed that, although the membrane consists of a dielectric, there is a current flowing through the membrane. This *leakage current* is likely due to tunneling of charges through defects in the SiN membrane [7]. These charge defects will accumulate over time until the material breaks down in a single spot on the membrane. A sudden and steep increase in the leakage current is observed as the material breaks down, indicating the formation of a pore in the membrane. Since defects

have accumulated in other spots on the membrane, the voltage needs to be switched off immediately after this first breakdown in order to prevent the formation of multiple pores in the membrane. Real-time feedback control with an adaptive threshold is used to detect this spike in leakage current and switch off the voltage (see figure 1.3).

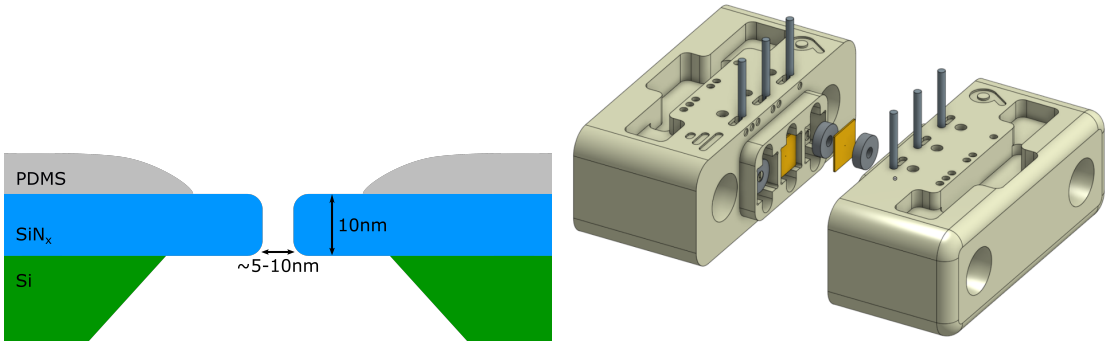

Figure 1.2: *Left: Schematic cross-sectional view of the silicon chip containing the membrane. Right: CAD drawing of the fluidic cell used to hold the membranes and the electrolyte solution. The fluidic cell can mount up to three membranes at a time and each membrane can be separately addressed using custom electrodes. The silicon chip supporting the membrane is drawn in yellow.*

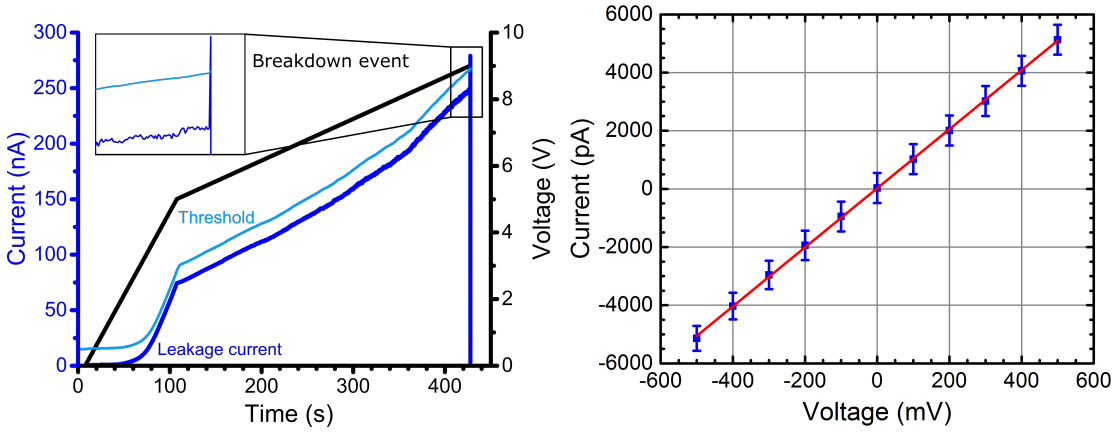

Figure 1.3: *Left: Example current trace of a pore forming. The inset shows a zoom in on the breakdown indicated by the sudden jump in leakage current. The black curve shows the voltage applied to the membrane. Right: I-V curve of a pore. The slope of the linear fit is equal to the pore conductance G, which is used to calculate the pore size.*

Since the nanopore was formed in situ, it is very challenging to use standard imaging techniques (STM, TEM) in order to confirm the size or the shape of the pore. The pore forms randomly anywhere on the membrane [8] and the membrane is usually large ($50\,\mu m \times 50\,\mu m$) compared to the size of the pore in standard membranes, making it hard to locate the pore on the membrane, especially with debris from the fabrication process

(salt crystals, etc.) on the membrane. Various studies have shown that nanopores formed by CBD are approximately cylindrical in shape [9, 10, 11, 12]. Hence, the size of the nanopore can be calculated using a simple model for the conductance of a cylindrical pore in a membrane and accounting for the access resistance [13, 14]. This gives rise to the equation for the conductance of the pore

$$G = \sigma \left[\frac{4t}{\pi d^2} + \frac{1}{d} \right]^{-1}, \tag{1.1}$$

where G is the conductance of the pore, σ is the conductivity of the electrolyte solution, t is the thickness of the pore, and d is the pore diameter assuming cylindrical geometry. This approximation works only in high salt when surface charges on the membrane are effectively screened by the counter ion shell and can be ignored. The conductance is experimentally determined by measuring the current at various voltages and extracting the slope of the linear fit to this I-V curve as shown in figure 1.3.

After fabrication by CBD nanopores are usually 2 to 3 nm in size but for most applications, larger pores are needed. To grow the pore to the required size a series of rectangular voltage pulses is applied to the pore to control the pore size with sub-nanometer precision [15]. This process is called *conditioning*. The conditioning voltage is usually 1/3 or less of the voltage at which the breakdown happened, to prevent formation of additional nanopores during conditioning.

1.3 Sensing

In this study DNA molecules are electrophoretically driven through nanopores. The DNA is injected on the membrane side of the silicon chip and a voltage is applied across the nanopore while the current through the nanopore is monitored. A DNA molecule in the pore physically blocks electrolyte ions from passing through the pore and therefore partially blocks the current through the pore. Successful translocations of DNA through the pore can be seen in the current trace as transient current shifts. The expected blockage for a double strand of DNA can be calculated using

$$\Delta G_{\mathrm{DNA}} = \frac{I_{\mathrm{DNA}}}{V_{\mathrm{bias}}} = A_{\mathrm{DNA}} * \frac{\sigma}{t}, \tag{1.2}$$

where I_{DNA} is the amount of current blocked by the presence of the DNA, V_{bias} is the applied bias voltage, and A_{DNA} is the area the DNA blocks inside the pore. In the simplest approximation DNA is assumed to be a long cylinder with radius $d_{\mathrm{DNA}} = 2.2\,\mathrm{nm}$ giving

$$A_{\mathrm{DNA}} = \frac{\pi d_{\mathrm{DNA}}^2}{4} \approx 3.8\,\mathrm{nm}^2. \tag{1.3}$$

Equation (1.2) can also be used to calculate the thickness of the pore. Due to the fabrication process of the silicon chips the membrane thickness is subject to fluctuations. By measuring the blockage for a piece of DNA and extracting the average blocked current, the thickness can be determined experimentally (see section 2.1).

Substituting equation (1.3) back into equation (1.2) with a membrane thickness of $t = 10\,\mathrm{nm}$ and a solution conductivity of $\sigma = 16\,\mathrm{S/m}$ gives

$$\Delta G_{\mathrm{DNA}} = \frac{I_{\mathrm{DNA}}}{V_{\mathrm{bias}}} = 3.8\,\mathrm{nm}^2 * \frac{16\,\mathrm{S/m}}{10\,\mathrm{nm}} \approx 6\,\mathrm{nS}. \tag{1.4}$$

In this study most experiments are performed at $V_{\mathrm{bias}} = 200\,\mathrm{mV}$, giving $I_{\mathrm{DNA}} = 1.2\,\mathrm{nA}$. This tiny jump in current needs to be detected with high temporal resolution, since passage times of DNA through nanopores are generally very short (on the order of a few 10s of microseconds). Therefore high-bandwidth, low-noise amplifiers are required in order to perform these measurements [16].

For DNA sensing in nanopores a Chimera VC100 current amplifier (Chimera Instruments LLC, Providence, RI) was used. The Chimera has a bandwidth of up to 1 MHz with a sampling rate of 4.166 MHz, meaning that events as short as 2 µs can reliably be detected. At those frequencies the noise performance suffers badly as described in section 1.4 and additional means of improving signal-to-noise ratio (SNR) are needed. A schematic of the final sensing setup is shown in figure 1.4. The whole sensing setup consists of a holder for the fluidic cell including custom electronics for fabrication and a sensing headstage for electronic current readings. Connected to the headstage of the sensing electronics is the data acquisition card of the Chimera VC100, which is then connected to the PC via USB. The system is operated using Matlab (The Mathworks Inc., Natick, MA) scripts on the PC. For pore fabrication a custom LabView (National Instruments) program is used, that runs an automated fabrication protocol to condition pores to the desired size. All measurement data have been filtered using an 8-pole software low-pass Bessel filter with a cut-off at 900 kHz unless otherwise noted.

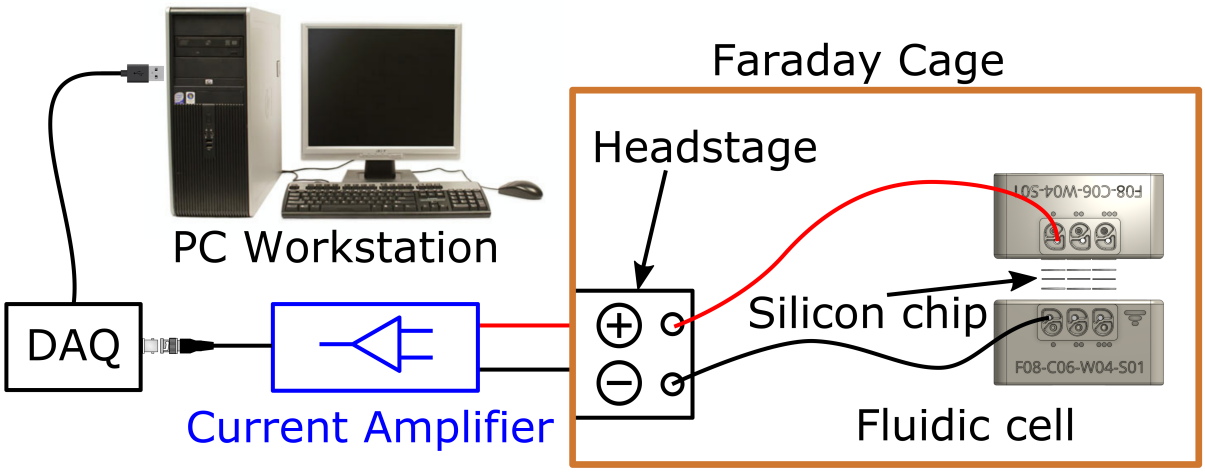

Figure 1.4: *Schematic overview of the sensing setup for nanopore experiments.*

1.4 Noise in Nanopores

Noise is in all physical measurements. For nanopores specifically noise influences the way the measurements are performed. To achieve the best time resolution, ideally, nanopore measurements should be performed at highest possible bandwidth. The dielectric and capacitive noise in a nanopore scales with frequency and therefore increases drastically with higher bandwidth [17]. Finding the right balance between low noise measurements and high time resolution is a challenge to every part of the nanopore setup.

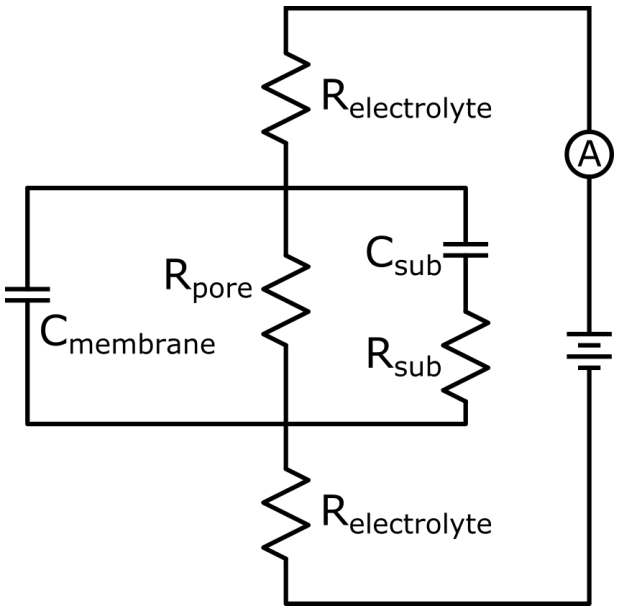

Figure 1.5: *Simplified equivalent circuit of a nanopore setup considering the capacitance and the resistance of the silicon substrate.*

The simplest way to understand noise pickup is to look at the equivalent circuit of a nanopore as shown in figure 1.5 [18]. The circuit shown has various sources of noise each dominating in different ranges of the frequency spectrum. They differ not only in origin but also in frequency behavior [19]. The power spectral density (PSD) plot reveals the dominant sources of noise for each frequency region. The overall shape of the PSD can be fitted to the following polynomial form

$$S_I(f) = \frac{a_0}{f} + a_1 + a_2 f + a_3 f^2 \ \ \mathrm{A}^2/\mathrm{Hz}. \tag{1.5}$$

All the sources are shown in figure 1.6 schematically. As a measure for the actual level of noise in a current trace, the root-mean-square (RMS) noise is used. The RMS is derived from the PSD as the square root of the integral over all frequencies

$$I_{\mathrm{RMS}}(f) = \sqrt{\int_0^B S_I(f) df} \tag{1.6}$$

where B is the bandwidth of the measurement. At all frequencies thermal and shot noise set the lower limit in the noise spectrum.

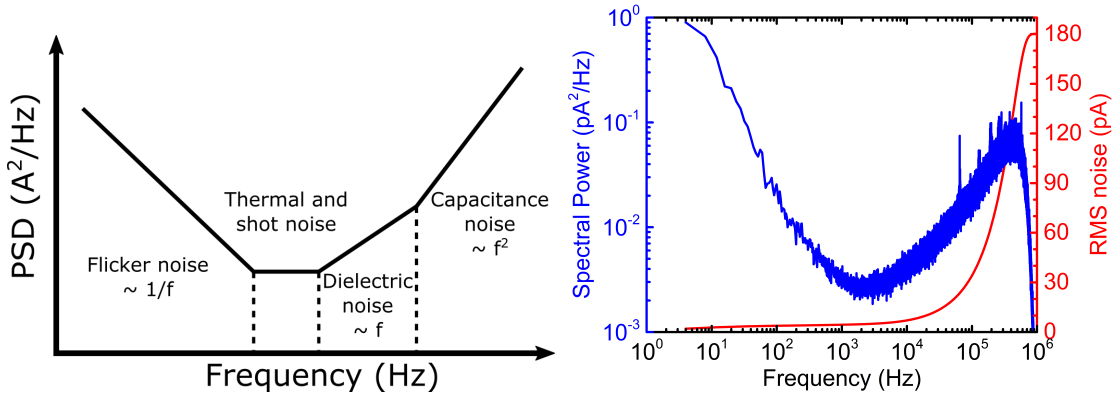

Figure 1.6: *Left: Different sources of noise with their influence on the overall PSD [19] (Scales are log-log). Right: Actual PSD (blue) of a nanopore recorded in 3.6 M LiCl at 200 mV bias voltage. The PSD is generated as the average of 30 discrete Fourier transforms on 1 s long traces of the baseline current. The cutoff frequency of the software low-pass Bessel filter that was used for this PSD was set to 900 kHz (3 dB point). The red curve shows the RMS noise as a function of frequency as shown in equation* (1.6).

These forms of noise are independent of the frequency. The thermal noise is generated by thermal fluctuations of charge carriers in a conductive medium and scales with the

temperature T and resistance R of that medium. The noise current S_{thermal} assumes the form [20, 21]

$$S_{\text{thermal}} = \frac{4kT}{R}. \tag{1.7}$$

Electric current flowing through the nanopore gives rise to another type of noise originating from the random fluctuations in the motion of charge carriers in the medium. This so called shot noise S_{shot} is also independent of frequency but scales with the current flowing through the pore. It takes the form

$$S_{\text{shot}} = 2Iq \tag{1.8}$$

with q being the effective charge of the carriers and I the average current flowing through the pore.

The dominating noise in the low frequency range is called flicker noise S_{flicker}. This noise appears when a voltage is applied to a nanopore and exhibits a $1/f^{\beta}$ characteristic according to

$$S_{\text{flicker}}(f) = \frac{\alpha I^2}{N_c f^{\beta}} \tag{1.9}$$

where α is the Hooge parameter, N_c is the number of charge carriers, and β is an experimental constant that is usually equal to unity but varies under different experimental settings [22]. Flicker noise appears in solid-state pores more pronounced than in biological pores and methods have been developed in order to decrease the amplitude of this type of noise [17, 18]. The origin of this type of noise is still not entirely clear [23].

At high frequencies dielectric and capacitance noise dominate the spectrum. The dielectric part of the noise $S_{\text{dielectric}}$ originates from the fact that the materials comprising the silicon chip, supporting the membrane, are not ideal dielectrics. Therefore they convert a part of the electrical energy into additional thermal noise. This noise scales linearly with frequency in the form

$$S_{\text{dielectric}}(f) = 8\pi kTDC_{\text{chip}}f \tag{1.10}$$

with D being the dielectric loss and C_{chip} being the capacitance of the silicon chip.

At the highest frequencies (>10 kHz) the noise is dominated by capacitance noise $S_{\text{capacitance}}$. This noise scales quadratically with frequency,

$$S_{\text{capacitance}}(f) \propto (C_{\text{total}} * f)^2 \tag{1.11}$$

with C_{total} being the total capacitance of the setup. Especially for high bandwidth measurements (>100 kHz) this noise will dominate the PSD and the overall RMS value. In solid-state nanopores this noise often imposes an upper limit on the bandwidth of the current reading. Because of its dependence on C_{total} this noise can be minimized by reducing the capacitance of the system. This can be done by adding an additional layer of a dielectric material to the supporting chip of the membrane. All chips used in this thesis have been painted with the fast curing silicon elastomer Polydimethylsiloxane (PDMS) as described in [17] in order to achieve the desired reduction in total capacitance. An extensive explanation of the different equivalent circuits and their noise behavior can also be found in [19, 22] and [18].

1.5 Analysis Software

The analysis of the recorded data was performed using the CUSUM+ [24] algorithm which is an improved version of the CUSUM (Cumulative Sum) algorithm [25], as well as MOSAIC (A Modular Single-Molecule Analysis Interface), which was developed by NIST (National Institute of Standards and Technology) [26, 27]. For linear DNA molecules CUSUM was used while for more complex DNA structures the ADePT (algorithm for decision-making after pilot and feasibility trials) module implemented in MOSAIC was used for analysis.

1.5.1 The CUSUM+ algorithm

The CUSUM algorithm was first published by Page [28] as a way to detect changes in a given statistical variable that is observed in an ordered manner. The first application to nanopore experiments was shown by Raillon *et al.* [25]. For this thesis a custom CUSUM+ implementation written by Kyle Briggs was used for analysis.

If a variable follows a known distribution (e.g. Gaussian) the algorithm finds changes in parameters of this distribution and locates the position where these changes happened. For a nanopore experiment the noise of the open pore current follows a Gaussian distribution

[25]. Any change in the value of the current (i.e. any blockage in the current due to the presence of a DNA molecule inside the pore) changes the mean of that Gaussian distribution. This change in the mean (which is called an *event*) can be detected by the CUSUM algorithm. Additionally the start and end times of these shifts in the mean can be determined giving information about the dwell time of the recorded events.

This algorithm is useful because it does not need any information about the system, which makes the fitting very fast. If events become short (less than 5 times the system characteristic relaxation time τ) they do not have enough points to be fitted correctly by this algorithm, which results in the blockage depth of the events not being determined correctly [27]. The relaxation time τ of the system is determined by the capacitance of the membrane and can be influenced by the cutoff frequency of the applied software filter. CUSUM+ is able to fit multi-level events by applying the same algorithm on every level of a given event. It will determine the start and end time for every sub-level of an event first and then fit all of the levels to the same precision. The output after fitting of the events is a database containing all information about the baseline (open pore current) as well as number, depth, and dwell time of every sub-level of the event. Custom Python scripts are then used to post-process, filter, and plot different aspects of the events collectively.

In this thesis CUSUM+ was used for analysis of events generated by long strands of DNA (>1 kbp (kilo basepairs)). For shorter and more complex molecules CUSUM+ is not precise enough and the ADePT module of MOSAIC was applied.

1.5.2 MOSAIC and the ADePT module

MOSAIC is an open source algorithm developed by Arvind Balijepalli from NIST and is made available via GitHub [29]. It uses a variety of different fitting algorithms for nanopore recordings. For this study the ADePT module, implemented in MOSAIC, was used. Detection of events in the current trace works in the same way as for the CUSUM algorithm. The fitting of the event after localizing it uses a different approach. The ADePT algorithm assumes a physical system behind the current values that are recorded. Events are seen as step changes in the current taking the system response into consideration. The current function of an event with N levels then takes the form

$$i(t) = i_0 + \sum_{j=1}^{N} a_j \left(1 - e^{-(t-\mu_j)/\tau_j}\right) H(t-\mu_j) \tag{1.12}$$

where i_0 is the open pore current, H is the Heaviside function with delay μ_j and step height a_j of the jth level, and τ_j the duration of that level [26]. This equation is then used fit to the data. The output of these fits is a similar database to the one generated by CUSUM and the same post-processing scripts are applied for data analyzed with MOSAIC.

The benefit of ADePT is the ability to correctly fit events down to 2.5 times the system relaxation time τ. The trade-off is a longer calculation time for the fit. In this thesis ADePT is used every time short and structured molecules are used since the levels and sub-levels are expected to be very short.

1.6 Structure of DNA

DeoxyriboNucleic Acid (DNA) is a macromolecule that can be found in all living organisms. Its structure allows for easy molecular storage, reading, and replication of genetic information. The structure of DNA was first shown by Watson and Crick [30, 31]. The building blocks of DNA are called nucleotides. These consist of a sugar molecule (2-deoxyribose), a phosphate group, and one of the four organic bases adenine (A), thymine (T), guanine (G), and cytosine (C). Nucleotides are connected via their phosphate groups and sugars to form oligonucleotides (*oligos*). Two oligos can be connected through the formation of hydrogen bonds between the bases in process called hybridization. For hybridization to occur the base sequences of the two strands need to be complementary to one another. In this process an A on one strand always pairs up with a T on the other strand and vice versa, the same applies for C and G. The final product of this process is a double stranded piece of DNA (dsDNA) that assumes a helical structure. The length of a dsDNA molecule is usually given in the number of monomers. A monomer for dsDNA is called a *basepair* (bp, 1000 bp = 1 kbp). Each basepair is approximately 0.34 nm long, meaning a fragment of DNA with N basepairs is roughly $N/3$ nm long.

In electrolytic solution the phosphate groups in the backbone of the DNA are negatively charged and the whole molecule is therefore surrounded by a shell of counter-ions [32] effectively shielding its charge, thus enhancing conformational freedom. Especially strands of DNA significantly longer than the persistence length of DNA ($\gg$170 bp or 50 nm [33]) will be tangled up in free solution. The persistence length is a measure of the stiffness of a polymer and describes the length scale over which the polymer can completely be folded back on itself (all correlations between pieces of the chain are lost over this distance

and the pieces move independently of each other). This leads to capture and, for large enough pores, translocation of the DNA in a folded fashion for the majority of events. Translocation kinetics differ drastically with decreasing DNA length and pore size as DNA folding cannot be seen for DNA lengths close to the persistence length ($< 250\,\mathrm{bp}$) or pores smaller than double the DNA cross-sectional diameter ($< 4\,\mathrm{nm}$).

The high charge of DNA enables electrophoretic sensing and is therefore one of the crucial properties of DNA for nanopore sensing. Additionally DNA is chemically stable under a lot of experimental conditions (wide range of pH, high salt, etc.), allowing to adjust the conditions to fit other aspects of the experimental setup (e.g. pore fabrication, pore stability).

1.7 Theory of Capture

In an electrolyte solution DNA carries a negative charge. This allows for electrophoretic capture in a nanopore experiment. Due to the setup of the nanopore, the electric field is strongest directly at the pore (see figure 1.7), since all of the potential drop happens across the membrane (the resistance of the electrolyte solution is negligible). From there the field radially decays until it becomes negligibly small and charged molecules behave according to regular diffusion [34]. The point at which almost unbiased diffusion crosses over to electrophoresis due to the presence of the applied electric field (the "event-horizon" of the nanopore) is called the *capture radius* r^*. Models show that the capture radius has an approximately hemispherical shape [35] and its size is determined by the solution conductivity, pore size and applied voltage as well as the diffusion coefficient of the analyte.

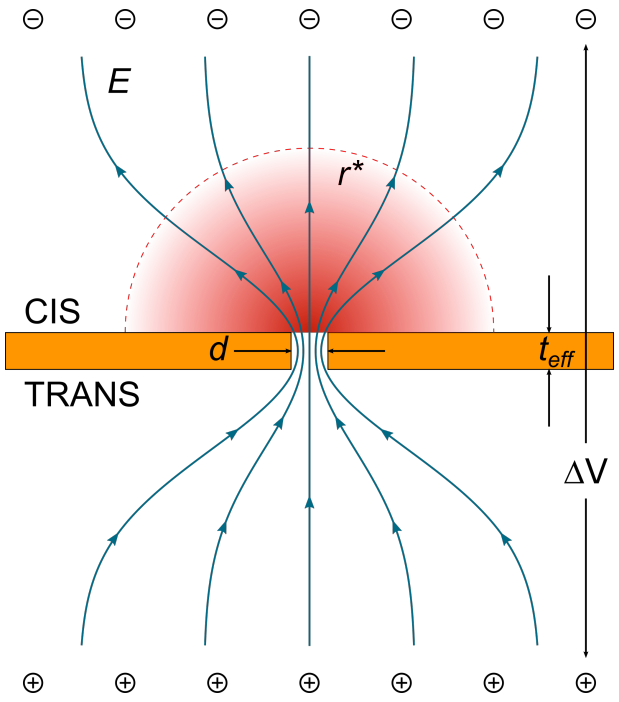

Figure 1.7: *Electric field profile (blue) of a nanopore in a thin insulating membrane and its capture radius (red). Inside the capture radius the electric field is strong enough to drive molecules towards the pore (Directed diffusion). Molecules are added to the* cis *side of the membrane and free diffusion drives them close to the capture radius.[36]*

In the description of capture behavior of biological molecules, two regimes can be distinguished. These regimes are defined by the behavior of the molecule shortly before translocation. If the pore is very small an energetic barrier exists, that prevents translocation on the first attempt if the electric field is not strong enough to pull the molecule into the pore (*barrier limited regime*). For larger pores this energetic barrier becomes smaller and even low bias voltages are enough to translocate the molecule through the pore, meaning that every molecule captured inside the capture radius will directly translocate without delay (*diffusion limited regime*). The final translocation speed and rate are a result of various forces acting on the molecule close to the pore, such as flow through the pore (electrical and fluid flow), electroosmotic flow generated by the presence of surface charges on the walls of the pore, diffusional forces acting on the DNA and the ions of the solution, and electrical forces between the surface charges of the pore and the DNA. A more in-depth discussion on the behavior of polymers entering a pore can be found in [37].

1.8 Label Design

The objective of the thesis is the design of a DNA label that can be used for a nanopore-based counting assay. An ideal label is reliably detected in a nanopore setup, meaning that every event is detected and distinguished from background noise. Furthermore an ideal label will always translocate in a similar fashion giving the event a distinct current signature indicating that every event corresponds to a full translocation. Due to random molecular motion nanopore translocation events show a certain spread in total translocation time as well as blockage depth [38]. This means that a statistically significant number of events need to be recorded in order to correctly analyze and characterize the sample. Therefore nanopore measurements are critically dependent on the rate of translocations (*capture rate* R_C) in order to minimize measurement time. The capture rate depends on various properties of the DNA (length, charge, diffusion coefficient) [37], the pore itself (aspect ratio, surface properties) and outside factors (DNA concentration, salt concentration, applied voltage) [35]. Altogether an ideal label needs to have five main characteristics:

1. Fast capture for short total measurement times
2. Long translocation times to guarantee all events are detected and to reduce bandwidth requirements
3. Deep current blockage (increased SNR) for reliable detection at any given bandwidth
4. Distinguishable and consistent event shapes during translocation for multiplexing purposes
5. Easy fabrication and customizability

Linear dsDNA naturally fulfills some of these conditions. Due to its high charge in high-salt solutions it captures fast. This is especially true for long strands of DNA ($> 10\,\text{kbp}$) [36]. Additionally DNA strands of that length also show translocation times on the order of $100\,\mu\text{s}$ allowing for lower bandwidth measurements ($<100\,\text{kHz}$) therefore reducing noise and increasing SNR. In this length regime DNA will naturally coil itself up. This means it will approach the pore in a variety of different configurations and will be sterically hindered from entering the pore [39]. Therefore it is highly probable that the molecule gets captured by the pore somewhere in the middle of the molecule and translocates in a partly folded fashion. This folding will also be seen in the recorded traces in the form

of multi-level events. Detecting modifications along the dsDNA backbone (proteins, side branches, etc.), that generate additional levels in the translocation events, is therefore made more difficult. It has been shown that adding even a single modification on a long strand of dsDNA greatly complicates analysis by increasing the number of observed event shapes [40]. Furthermore strands of this length cannot simply be modified at sequence specific locations in order to build up DNA structures (as required for the purposes of this study) in a bottom-up fashion, meaning that customizability is greatly reduced and sample preparation becomes more difficult. Overall long strands of dsDNA are a good choice for a label but they lack multiplexing capabilities, due to the variety of event shapes observed for these molecules, and customizability. However, labeling schemes using long strands of DNA have been proposed by Bell *et al.* [41] as well as Plesa *et al.* [42].

Some of the problems with long DNA can be solved by using short strands instead (<150 bp). In this regime dsDNA behaves like a semiflexible rod in free solution. This leads to much less conformational freedom for the molecule and therefore more consistent event shapes. The trade-off that has to be made lies in the shorter translocation times of short strands of DNA requiring higher bandwidth electronics especially if pore sizes become larger removing interactions between the pore wall and the molecule. Higher bandwidth comes at the cost of higher noise (see section 1.4) and therefore reduced SNR setting a lower limit for the length of DNA that can be used. Very short strands of DNA (<100 bp) can easily be synthesized in a bottom-up scheme allowing for complete control of sequence and sequence specific modifications of the DNA.

The approach taken in this thesis is based on the customizability of short strands of DNA. A recent study has already proven the usability of short, branched DNA scaffolds with application to an ATP-detection assay [43]. In order to combat the challenges associated with short strands, a design was chosen that increases the observed blockage depth and therefore increases SNR and allows higher bandwidth measurements. Due to the increased SNR it also allows for the use of larger nanopores, as the use of high bandwidth electronics enables higher temporal resolution to compensate for the faster translocation times of these short strands. To achieve this the molecule was chosen to assume a branched conformation. The backbone and the side branch of the construct need to be structurally simple and rigid in order to not deform during translocation. The linker that connects both parts needs to be flexible but strong such that the side branch can freely move with respect to the backbone to allow easy translocation and minimize any energetic barrier

associated to the translocation but is strong enough to keep the two parts together even with considerable force acting on the link. This also enables the use of fairly small pores down to 4 nm. The first design for a label proposed here uses a 51 bp long dsDNA backbone and a 25 bp long side branch. The 26th base of the backbone is replaced with a DBCO-dUTP (Dibenzylcyclooctyne-deoxyuridinetriphosphate) that forms a covalent bond to the azide modification on one end of the side branch in a process called the *click* reaction. For more detail on this reaction see appendix A. In this approach the side branch can be substituted easily to any label that can be modified with an azide modification (e.g. Polyethylene glycol (PEG), proteins, dyes). PEG modifications have also been tested prior to the full DNA approach, but it was found that the modification was not reliably picked up during experiments. The final structure of the molecule resembles the form of the letter T and will therefore be called *T-shape*.

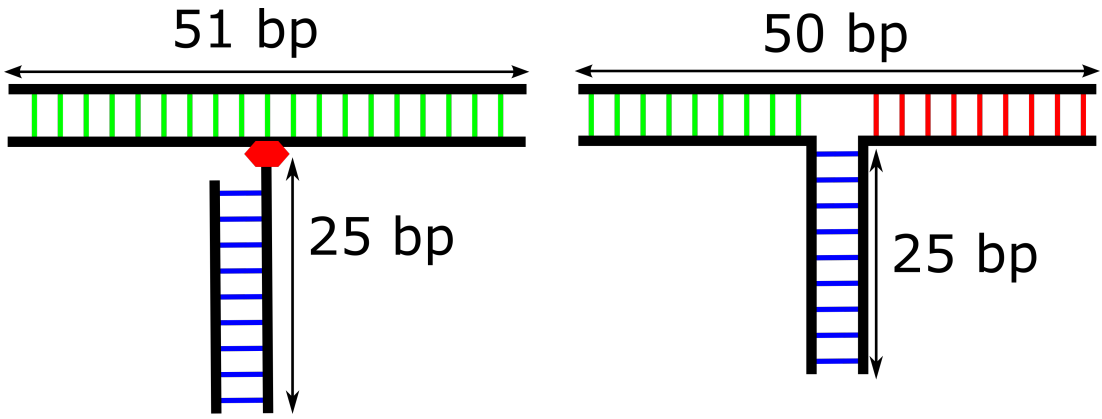

Figure 1.8: *Schematic structure of the two labels described. Left: T-shaped label assembled using the click reaction. The red hexagon marks the position of the linker. Right: Y-shaped label assembled using DNA origami. The colors mark complementary sequences.*

In a second design a similarly structured molecule is assembled in a pure DNA approach using DNA origami [44]. Three 50 nt (nucleotide) long oligos are designed, with their sequences halfway complementary such that the final structure assumes a three branch star like structure similar to the one described above. The structures of both molecules are shown schematically in figure 1.8.

The goal of this thesis is to establish a DNA-based label that shows robust detection in a range of pore sizes. This label is proposed to be used for a biomarker detection assay as described in [45] with nanopore readout. For this purpose the first label proposed is discussed in great detail to show that all requirements are fulfilled as mentioned above. Mainly it needs to be shown that every label that enters the pore is detected by our software

and hardware. This means that SNR needs to be sufficiently high and dwell times need to be in an accessible range in order to correctly characterize the events. This allows for quantification of the actual number of molecules in the solution and enables these labels for use in the mentioned biomarker detection assay. In addition to the proposed labels above the continuation of this project will include the design of more advanced DNA labels by adding more branches to improve on those properties already shown for the T-shaped and Y-shaped labels. This will also enable potential multiplexing capabilities by distinguishing different structures from one another at the nanopore sensing step.

Chapter 2

DNA in Nanopores

2.1 DNA as a molecular ruler

Due to the method of fabrication of the free standing SiN membranes, the thickness of these membranes is subject to fluctuations. The pores used in this thesis have been made in membranes from different batches of production as well as slightly different designs of the chip geometry. Furthermore, while CBD makes to first order cylindrical pores when applied to thin (<10 nm) membranes, some conical structure can not be completely ruled out. An effective membrane thickness is therefore defined to represent the thickness of the pore in the cylindrical approximation. This leads to a variability in the membrane thicknesses and therefore, according to equation (1.2), also in the observed blockage depths of DNA translocating through these pores. Additionally the procedure of fabricating pores *in situ* means that standard imaging methods like Scanning electron microscopy are not available to gain information about the size and shape of the pore. Thus the need for a calibration arises.

The pore size can be determined as described above using equation (1.1). In order to measure the thickness of the membrane we will again use equation (1.2) in combination with a molecule of known geometry and with long enough events for the software to be able to correctly fit the blockage states. In this study linear strands of dsDNA are used for that purpose. A length of 2 kbp was found to be sufficiently long in order to fully resolve every translocation event and fit the levels (due to folding) appropriately. In every pore used in this study a sample of 4 nM 2 kbp dsDNA was injected into the fluidic cell and translocation events were recorded at a voltage of 200 mV or 400 mV, depending on pore size. The recorded current traces were filtered to 250 kHz to increase SNR and

analyzed using the CUSUM algorithm and blockage depths were extracted. A single peak Gaussian has been fitted to the blockage depth histograms and the peak position was used to calculate the thickness (figure 2.1). With the thickness of the pore and the solution conductivity all the blockages recorded for the structured molecules can be converted to an effective area of the pore, that is blocked by the molecule. These effective areas can then be used to compare results across pores and different molecule geometries.

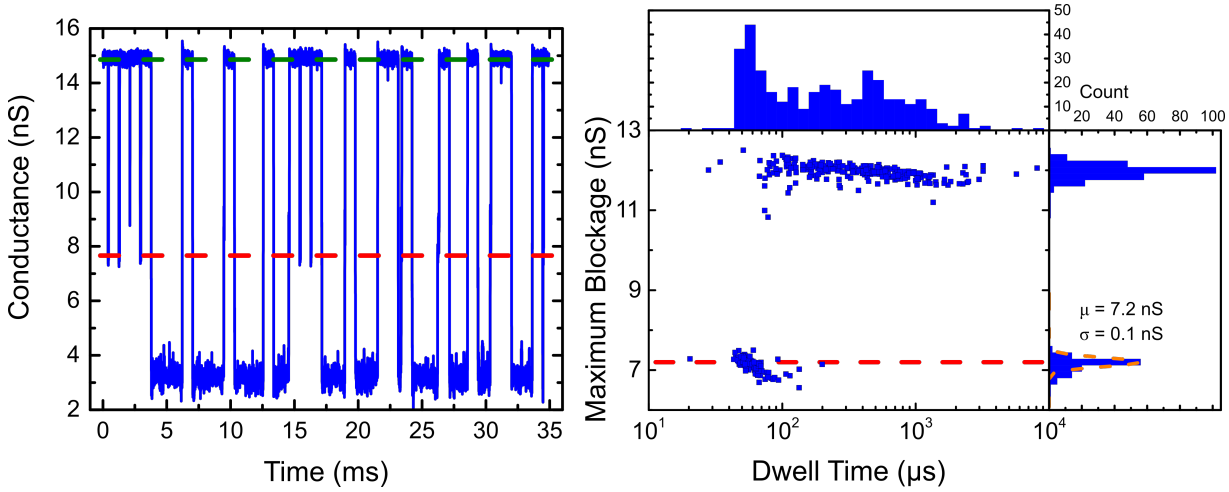

Figure 2.1: *An example thickness calibration from the blockage depth histogram of a 2 kbp dsDNA sample in a 3.7 nm pore at 200 mV. Left: Concatenated current traces of the first 20 events. The green dashed line marks the baseline level and the red dashed line marks the first recorded blockage level. Right: All points plot of the maximum conductance blockage versus the dwell time for every event. On the top and right side are histograms of the corresponding quantity. The orange dashed line represents the Gaussian fit to the first observed peak in the maximum blockage histogram. The peak position of this fit is then used for the calculation of the thickness according to equation* (1.2)*. The red dashed line marks the peak position of the fit for the first level. The secondary deeper blockage level is attributed to DNA folding and is not used for calibration purposes.*

2.2 Capture Rate determination

One of the big advantages of nanopore sensing is the ability to investigate single molecules one at a time. In order to get meaningful statistics on the behavior of the whole ensemble hundreds of events need to be recorded. The rate at which events are recorded in a nanopore experiments is called *detection rate*. The rate at which molecules arrive at the pore is called *capture rate*. Ideally every molecule that passes the pore is also detected by the sensing hardware and software, meaning that capture rate and detection rate are identical. While the detection rate can be improved by outside factors such as bandwidth (temporal

resolution) of the hardware and noise (SNR), the capture rate is an intrinsic parameter, that depends critically on a variety of experimental parameters such as concentration of the analyte, voltage applied, electrolyte solution (conductivity, salt composition), pore size, DNA length, and electrophoretic mobility of the analyte [37, 35, 36].

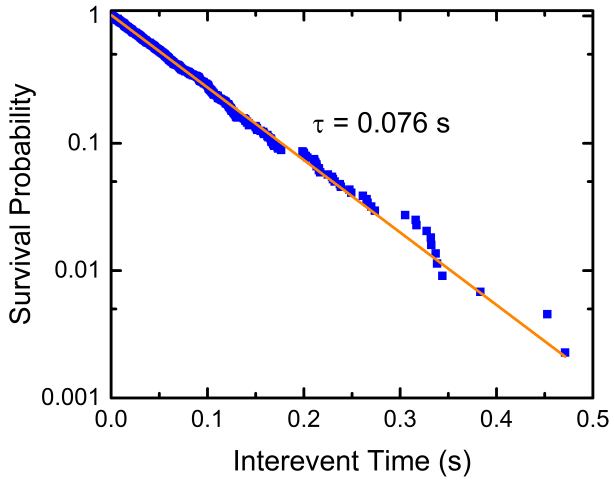

Figure 2.2: *Example of capture rate determination. Scatter plot of the Survival probability versus the inter-event time for a 30 nM 50 bp dsDNA sample in a 7.6 nm pore at 200 mV bias voltage. The orange line represents a single exponential fit to the data with a rise time of* $\tau = 0.076\,\mathrm{s}$ *giving a capture rate of* $R_C = 13.1\,\mathrm{s}^{-1}$.

The capture rate of an experiment can be extracted by calculating the time between consecutive events (interevent time) in the current trace. For every interevent time the survival probability can be calculated as

$$S_i = 1 - \frac{i}{N}, \tag{2.1}$$

where i is the index of the ith smallest interevent time and N is the total number of valid interevent times. An interevent time value is considered valid if there are no threshold crossings detected by the fitting algorithm in between the two events (consecutive events only). Plotting the survival probability against the interevent time reveals a straight line on a log-normal plot indicating a single exponential relationship between the two quantities (see figure 2.2). The rise time τ of this exponential is related to the capture rate $\tau^{-1} = R_C$ according to [36]

$$S(t_{\mathrm{inter}}) \propto \exp \frac{t_{\mathrm{inter}}}{\tau}, \tag{2.2}$$

where t_{inter} is the inter-event time. All capture rates presented in this study are calculated this way.

2.3 Recent Progress by others

Labels similar to the one used in this thesis have been proposed for different applications. Recent approaches to the development of a label that is easily detectable and shows unique translocation characteristics include the design of DNA substructures [41] as well as complete non linear DNA structures [46]. Labeling schemes using proteins [47] or intercalating γPNA [48] have also been proposed. Applications for these labels reach from improving our understanding of polymer dynamics under confinement [42, 49] to detection schemes for reaction products [43, 50, 51]. Advances have also been made in the development of multiplexed nanopore assays [41].

Chapter 3

Capture and Translocation Characteristics of Short Structurally-Defined DNA Labels in Solid-State Nanopores

3.1 Preface

This chapter comprises a stand-alone manuscript, that is going to be submitted for publication. At the time of submission of the final version of this thesis this chapter has undergone the first round of reviews and proposed changes by the reviewers have already been implemented. The large majority of the work for this manuscript was done by me, including performing experiments, data analysis, and writing of the manuscript. The work was performed under the supervision of Professor Vincent Tabard-Cossa. The format of the manuscript has been modified to comply with the formatting of this thesis.

3.2 Abstract

The challenge when employing solid-state nanopores as single-molecule sensors in a given assay is the specificity of the ionic current signal during the translocation of target molecules. Here we present the capture and translocation characteristics of short structurally-defined DNA molecules that could serve as effective surrogate labels in biosensing applications. We produced T-shaped or Y-shaped DNA molecules with a 50 bp double-stranded DNA (dsDNA) backbone and 25 bp dsDNA branch in the middle, as improved labels over short linear DNA fragments. We show that molecular topologies can be distinguished from linear DNA by analyzing ionic current blockades produced as these DNA labels translocate through 10 nm SiN nanopores ranging in size from 4 nm to 10 nm fabricated by controlled breakdown. Event signatures are shown to be a direct result of the structure of the label and lead to an increased signal-to-noise ratio and well resolvable dwell times irrespective of the pore size, compared to short linear dsDNA. These results show that structurally-defined DNA molecules can be robustly detected for a broad range of pore size, and thus represent promising candidates as surrogate labels in a variety of nanopore-based molecular or immuno-assay schemes.

3.3 Introduction

Single-molecule electrical sensing using nanopores has the potential to enhance diagnostic technologies by providing ultra-sensitive detection in a format that is amenable to point of care testing [3]. Promising biomarker detection schemes based on solid-state nanopores have been demonstrated in recent years [40, 41, 43, 46, 48, 51, 52, 53, 54, 55, 56]. As in many bioassays, the use of surrogate labels, in place of directly detecting targets of interests, may be advantageous to improve performance [45, 57]. Indeed, many biomolecular sensing applications based on nanopores would benefit from a label that can be detected more accurately and rapidly, as well as be more easily distinguished from other background molecules when translocating through the pore, to improve specificity, sensitivity, and multiplexing levels. An ideal label would be captured rapidly but translocate slowly so as to facilitate fast detection while still allowing for fully resolved electrical signatures [58]. Realizing this requires (i) a fast diffusion and high electrophoretic mobility in the bulk solution for rapid capture and detection of a statistically relevant number of labels in short measurement times; (ii) a reduced velocity when inside the confined geometry of

a nanopore of varying sizes in order to robustly achieve slow passage times compared to the response time of the detection electronics; and (iii) simple synthesis or assembly of an easily functionalizable backbone for versatility and compatibility with different assay designs. Two of these conditions can easily be achieved using short dsDNA. DNA is highly charged, giving it a high electrophoretic mobility in bulk solution [59]. Double-stranded DNA with lengths below 100 bp are also readily available or can easily be produced using standard oligonucleotide synthesis procedures, allowing for specific sequences and a large choice of precisely located functional groups for further attachments to beads or other modifications (e.g. dyes, peptides, antibodies).

The major issue with short dsDNA molecules arises in the dwell time inside the nanopore, especially when pores become large compared to the molecular width of DNA (2.2 nm). Due to the linear rod-like nature of these molecules, interactions with the pore wall are minimal [60]. Combined with high molecular mobility, this leads to extremely short dwell times once the nanopore size is above $\sim$4 nm [39]. These short ($\sim$1 µs) dwell times give rise to a need for very high bandwidth (>1 MHz). Hardware satisfying this requirement are not easily available and require specially designed chips with reduced capacitance, and the signal-to-noise ratio (SNR) suffers badly for high bandwidth measurements [17, 61, 62], making it difficult to detect these short translocation events even if sufficient bandwidth is available.

Several different approaches can be taken in order to deal with these limitations, like changing salt concentration and composition [63], interfacing the pore with a gel medium [64], or using thermophoretic instead of electrophoretic motion [65]. Most commonly, the use of long DNA strands (>1 kbp) will shift translocation times to a more experimentally accessible region ($\sim$100 µs)[60, 66]. Groups have proposed very promising DNA carrier strategies based on M13 scaffolds [40, 48, 51, 67], though this scheme requires hundreds of staple oligos and careful optimization of assembly conditions.

Here, we present an alternative approach, which relies on short strands of DNA, which exploits their ease of synthesis and customizability and also allows for increased dwell time and increased blockage depth (and therefore SNR), while maintaining the advantages of short double strands of DNA in bulk solution [43, 56]. The proposed molecule is a small structure made of dsDNA fragments linked together using either copper-free click chemistry or DNA hybridization such that the resulting structure assumes a "T" or "Y" conformation with branches of equal length. Such structure increase both dwell time,

due to an increase in interactions with the pore wall, and blockage depth, due to a higher fraction of the nanopore being blocked during translocation. Due to the persistence length of dsDNA (∼150 bp) [33], both the backbone and the branch are rigid. The click chemistry approach has the advantage of being sequence specific, allowing for customization for different applications, potentially enabling distinction of the ionic current signatures of different branched structures (e.g. PEG of varying molecular weights) to allow for multiplexing in an assay [40]. Nevertheless, under our operating conditions, the two, all nucleic acid, "T" and "Y" labels of nearly identical size produced indistinguishable electrical signatures when translocating the nanopores and are therefore treated alike below. Indeed, both the "T" click chemistry design, with a side arm covalently linked via a flexible linker, and the "Y" hybridized structure, with a built-in flexible junction, provide rotational freedom at the junction to allow the different arms to reorient during capture and passage [44, 68].

3.4 Methods

3.4.1 Nanopore Fabrication

Nanopores are fabricated using controlled breakdown (CBD)[6], in nominally 10 nm thick SiN membranes supported by a silicon support structure in either 3 mm or 5 mm chip format (Norcada Inc, Edmonton, AB; NT005Z and NBPX5004Z). Before assembly in a custom fluidic flow cell, the membranes are painted with a layer of polydimethylsiloxane (PDMS) to reduce the device capacitance [17]. The chips are cleaned using air plasma for 20 s at 30 W. Fabrication is performed in 1 M KCl buffered with 10 mM $NaHCO_3$ at pH 10 using custom instrumentation controlled with LabView (National Instruments, Austin, TX) software. Following dielectric breakdown, the pore is conditioned to the desired size using a series of voltage pulses (±3.5 V square pulses with 8 s period). After changing the electrolyte solution to 3.6 M LiCl buffered with 10 mM HEPES at pH 8 ($\sigma = 16.1$ S/m), the pore size is extracted from the measured value of the pore conductance, G, determined by IV measurements, applying voltages ranging from −500 mV to 500 mV and performing a linear fit of the resulting curve. Assuming cylindrical geometry and accounting for access resistance [13, 14], the pore diameter can be extracted from the slope of the linear fit using:

$$G = \sigma \left[\frac{4t}{\pi d^2} + \frac{1}{d} \right]^{-1}, \tag{3.1}$$

where σ the solution conductivity, t the effective thickness of the membrane at the pore, and d the effective pore diameter.

3.4.2 DNA Label Assembly

Linear DNA fragments used in this study are purchased commercially (No Limits 50 bp, 1 kbp, 2 kbp, and 4 kbp DNA fragments, Thermo Fisher Scientific, Waltham, MA). DNA click reaction structures are synthesized and assembled by BioSynthesis Inc (Lewisville, TX). The final structure is schematically shown in figure 3.1a.

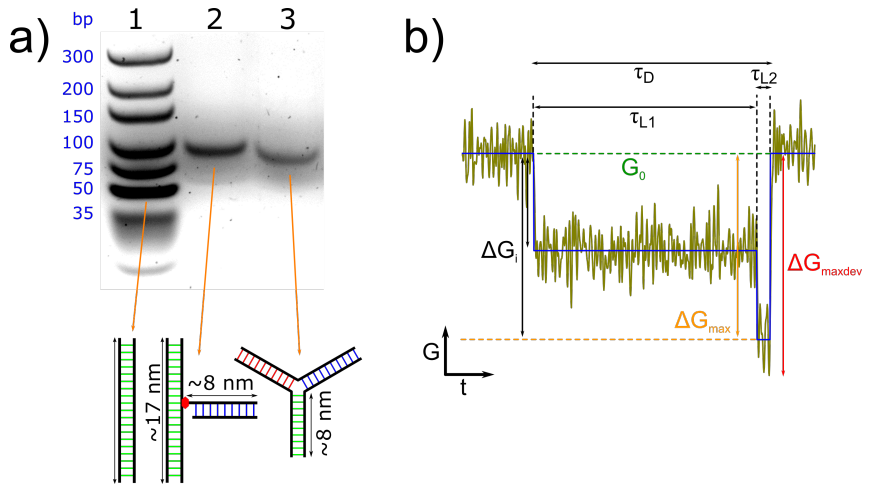

Figure 3.1: *a) 4% agarose gel of all the labels used in this study. Lane 1 shows the GeneRuler Ultra Low Range DNA ladder from Thermo Fisher Scientific. Lane 2, 3 contain the clicked T-shaped, and the hybridized Y-shaped, respectively. The shift in the bands for the T-shaped and Y-shaped molecules indicates the correct assembly of the product (the T-shape is slightly heavier due to the presence of the binding groups and the extra basepair in the backbone). b) Typical event for a Y label highlighting the different characteristics extracted from the current trace and the fit. Open pore conductance is denoted by G_0, blockage Levels by ΔG_i, maximum blockage ΔG_{max}, maximum deviation ΔG_{maxdev}, the level dwell times by τ_{L1} and τ_{L2}, and event dwell time by τ_D.*

DNA hybridization structures are assembled in-house. DNA oligos are mixed in a 1:1 ratio in TE buffer at pH 7.6 and heated to 90 °C for 5 minutes before being cooled down gradually back to room temperature by turning off the heat on the water bath. The resultant mixture is run on a 4% agarose gel to confirm that the hybridization was successful. Reaction mixtures are used without further purification for nanopore experiments. The oligos are purchased from Integrated DNA Technologies (IDT, Coralville, IA). The sequences for all oligos can be found in table 3.1, with the final structure shown schematically in figure 3.1a.

Table 3.1: *Table of all the oligos and their sequences. The U in the Click-T backbone strand marks the position of the DBCO-dUTP that is connected to the side branch 5'azide via click chemistry. The hybridization strands are color coded to indicate the complementary parts of the strands.*

Strand	Length	Sequence
T-shape backbone	51 nt	5′CCT GTG CTG GTT CGC TGC GGC GCC CUG GGG CCG CAG CCC AAG CAG GAA GCA3′
T-shape backbone complement	51 nt	5′TGC TTC CTG CTT GGG CTG CGG CCC CAG GGC GCC GCA GCG AAC CAG CACA GG3′
T-shape side branch	25 nt	5′azide TTT TCG GTT GGT GGG TGT TTG TGT G3′
T-shape side branch complement	25 nt	5′CAC ACA AAC ACC CAC CAA CCG AAA A3′
Origami-T oligo 1	50 nt	5′GTG TCT TGA AGA AGG ACC AGG GAA C GT ATT TGC CTT GTC TGG GAA ATC GT3′
Origami-T oligo 2	50 nt	5′ACG ATT TCC CAG ACA AGG CAA ATA C CC AAC TGG TTG TGG CCT ATC GAA AA3′
Origami-T oligo 3	50 nt	5′TTT TCG ATA GGC CAC AAC CAG TTG G GT TCC CTG GTC CTT CTT CAA GAC AC3′

3.4.3 Nanopore Measurements

Sensing of DNA labels was performed in 3.6 M LiCl buffered with 10 mM HEPES at pH 8 with applied voltages ranging from 200 mV to 400 mV. Samples were introduced on the cis side of the chip (membrane side) with the trans side grounded (etch pit side) and a negative voltage was applied to the cis side. The current recordings were performed using the Chimera VC100 (Chimera Instruments, Providence, RI) with a sampling frequency of 4.16 MHz and a bandwidth of up to 1 MHz. In this study, we show results acquired on 15 pores ranging in size from 2.5 ± 0.1 to (9.4 ± 0.2) nm. The uncertainty is calculated from the uncertainties of the effective length of the membrane as shown in table 3.2 and the standard deviation of the baseline current.

3.4.4 Data Analysis

All the current traces were digitally low-pass Bessel filtered to 900 kHz for analysis. Translocation events in the recorded current traces were located and fitted using the ADEPT algorithm implemented in MOSAIC [26, 27] (`https://pages.nist.gov/mosaic/`). An

event was called when a point in the current trace deviates from the local mean open pore current by more than 6x the local standard deviation. To prevent multiple events with short delays being detected as a single event, the baseline after the event is cropped below the threshold, only for fitting purposes. The events were post-processed using custom software and analysis tools, including a custom implementation of the CUSUM+ algorithm developed by our lab (`https://github.com/shadowk29/CUSUM`). Events with baselines deviating more than 3 standard deviations from the overall baseline level of the experiment are not considered for analysis.

Nanopore conductance changes caused by the translocation of a charged rod-like polymer can, to first order, be described by a simple geometric model [69]:

$$\Delta G = \sigma \frac{A\text{eff}}{t} = \sigma \frac{\pi d_{\text{DNA}}^2}{4t}, \tag{3.2}$$

where d_{DNA} is the diameter of the DNA going through the pore (for all the calculations d_{DNA} is assumed to be 2.2 nm). For unfolded DNA translocation it follows that:

$$A_{\text{DNA}} = \frac{\pi d_{\text{DNA}}^2}{4}, \tag{3.3}$$

where A_{DNA} is the area of the pore occluded by a linear strand of dsDNA. In order to compare different molecules across many pores and voltage values, the blockage depth is transformed into an effective blocked area A_{eff} using the more general version of equation (3.2).

In particular, for T-shaped and Y-shaped DNA molecules, events with sublevels are expected given the structure of this label compared to a linear fragment of DNA. If all the events are well resolved in time, we can confidently fit all levels within an event. However, as passage times get shorter compared to the response time of the system, for example as pores get larger, the resulting sublevels may not be fully resolved anymore and would be missed by the fitting algorithm. Here, we call the deepest "fitted" level for a given event the "maximum blockage". In cases where short sublevels of an event are missed by the fitting algorithm, the maximum blockage will not give an accurate representation of the event signature. Because each sublevel is fitted to a nonlinear curve with 3 parameters and to ensure that the fit will converge, substantially more than 3 data points (>1 µs) are necessary in order to correctly fit the sublevel. Since the response time for the system

is on the order of a microsecond, for events with short sublevels lasting less than a few microseconds, it is useful to instead look at the single furthest data point away from the local open pore current. We call this point the "maximum deviation" of an event. The maximum deviation metric thus measures the deepest sublevel of any event that is detected by the hardware, plus contribution of noise. Because of this, the maximum deviation is very sensitive to the noise level of the ionic current recording and will change based on the software filter settings used for analysis. Therefore, for all data shown here, events are fitted, and the maximum blockage is detected with a digital lowpass filter set at 900 kHz, the maximum deviation is extracted from the raw data, at the full 1 MHz bandwidth.

Both maximum blockage and maximum deviation are converted into an effective area as described above in order to most effectively compare results between experiments performed on different pores.

3.5 Results & Discussion

3.5.1 Linear DNA results

In order to compare blockage depths between structured DNA labels and standard linear DNA we used a relatively long dsDNA fragments as molecular rulers to measure the effective length of the pore and more accurately determine its diameter. The contour length of these DNA fragments (1, 2, or 4 kbp) allowed for easy fitting of the blockage depth since the dwell times were more than long enough for events to be fully resolved by our hardware and software. This control was performed at the start of every experiment prior to injecting the short, structured DNA molecules. Given a dsDNA diameter of 2.2 nm, we used equation (3.2) to determine the effective thickness of the pore from the mean of the observed first blockage level, which is representative of the unfolded polymer conformation. Table 3.2 summarizes the effective membrane thickness and pore diameters used in this study.

Table 3.2: *Characteristics of the pores used in this study. Listed are only pores for which long DNA has been used as a molecular ruler. The mean blockage level is extracted from a single peak Gaussian fit to the maximum blockage depth histogram with the uncertainty representing the standard deviation extracted from the same fit. The effective pore length is calculated from the mean blockage using equation* (3.2)*. The baseline is the mean of current level averaged over 0.5 s of data surrounding each event with the uncertainty representing the standard deviation at 250 kHz. All current traces have been low-pass Bessel filtered to 250 kHz prior to analysis. The pore diameter is calculated using equation* (3.1)*.*

Pore ID	Device and Batch	Length of DNA (kbp)	Applied Voltage (mV)	Number of events detected	Effective Pore Size, d_{pore} (nm)	Effective Pore Length, t_{eff} (nm)	Mean Baseline Level (nS)	Mean Blockage Level (nS)
P1	D1-B1	1	200	96	4.2 ± 0.1	5.4 ± 0.3	25.3 ± 0.4	11.0 ± 0.5
P2	D1-B1	1	400	409	3.6 ± 0.1	10.0 ± 0.3	12.8 ± 0.4	6.2 ± 0.2
P3	D2-B1	2	200	175	6.7 ± 0.1	11.4 ± 0.4	34.4 ± 0.4	5.4 ± 0.2
P4	D2-B1	2	200	186	5.0 ± 0.1	10.8 ± 0.2	21.6 ± 0.2	5.7 ± 0.1
P5	D2-B1	2	200	268	8.1 ± 0.1	12.3 ± 0.3	44.9 ± 0.2	5.0 ± 0.1
P6	D2-B1	2	200	613	7.2 ± 0.1	10.3 ± 0.2	41.5 ± 0.2	6.0 ± 0.2
P7	D3-B1	4	400	110	4.1 ± 0.1	10.5 ± 0.2	15.6 ± 0.2	5.9 ± 0.1
P8	D2-B2	2	400	183	3.9 ± 0.1	9.4 ± 0.2	15.7 ± 0.1	6.5 ± 0.1
P9	D2-B2	2	400	438	5.5 ± 0.1	13.9 ± 0.2	21.4 ± 0.1	4.5 ± 0.1
P10	D3-B1	2	400	529	7.4 ± 0.1	13.4 ± 0.1	36.0 ± 0.1	4.60 ± 0.03
P11	D3-B1	2	400	381	3.7 ± 0.1	8.6 ± 0.1	14.9 ± 0.1	7.2 ± 0.1
P14	D1-B2	0.05	300	515	2.4 ± 0.1	9.7 ± 0.1	6.7 ± 0.2	6.5 ± 0.1
P15	D1-B2	0.05	300	133	2.7 ± 0.1	8.8 ± 0.2	8.8 ± 0.1	7.4 ± 0.2

We first studied the translocation characteristics of short 50 bp dsDNA linear fragments, using the nanopores presented in table 3.2, with a range in size from 2.4 to 7.7 nm, to determine their performance as a label. Experiments were conducted by injecting a solution containing 50 nM DNA in 3.6 M LiCl pH 8 on one side of the nanopore. Figure 3.2a presents the observed maximum blockage averaged over >100 events per dataset. The horizontal lines indicate the expected blockage levels for unfolded (solid line) and folded (dashed line) DNA translocation calculated based on equation (3.3). For pores $\lesssim$4 nm, where the pore diameter is comparable to the size of the 50 bp dsDNA label, the maximum blockage is in good agreement with the expected blockage for a single linear dsDNA molecule (figure 3.2a, figure 3.2b). As the nanopore diameter increases, the maximum

blockage depth appears to be decreasing, since events become too short to be reliably detected and the fitting becomes increasingly difficult. For short events, the fitting algorithm will underestimate the actual blockage level of an event and will likely completely miss an event making quantification extremely challenging.

Detection of 50 bp dsDNA labels in larger pores (>4 nm) becomes increasingly difficult as the majority of the events are happening too fast to be fully resolved. This leads to the need for higher bandwidth electronics or reduced parasitic capacitance [16, 61]. As bandwidth is increased, however, the overall SNR is reduced, making it harder to detect these molecules. This trend is apparent by the drop of the observed mean maximum blockage with increasing pore diameter in figure 3.2a. This drop is not observed for longer strands of linear dsDNA (≥1 kbp), where dwell times are sufficiently long to enable correct sizing of pore dimensions. To highlight this, an example dataset is plotted in figure 3.2c for 50 bp dsDNA translocating a 7.7 nm pore. There is a single population visible, as expected for these labels. The blockage is significantly smaller than the calculated level for unfolded DNA translocation (solid grey line). The raw current trace for the 50 bp sample in this plot had to be filtered with a cutoff at 250 kHz to improve SNR and enable detection of these events. In the unfiltered case (900 kHz), no events were picked up by the fitting algorithm since their signals provided insufficient SNR for detection at high bandwidth. This leads to the assumption that a majority of the events are not picked up even at 250 kHz cutoff. These missed events also lead to an observed reduction in effective detection rate, i.e. fewer molecules are observed per unit time with increasing pore size, while without this finite bandwidth of our detection electronics larger pores are expected to provide a larger capture rate, R_C, seeing as $R_C \sim d_{\text{pore}}^2/t_{\text{eff}}$ [37, 36]. This puts a strong constraint on pore size (2.4 to ∼3 nm) if linear 50 bp dsDNA fragments are to be used as a label in a nanopore-based quantitative assay. Because the diameter of a solid-state nanopore often shifts slightly with time and voltage application, operating conditions requiring a very small range of pore sizes are undesirable. We propose it is preferable to find a way to increase SNR and mean passage times by using short, structured molecules as discussed below in order to relax the pore size constraints.

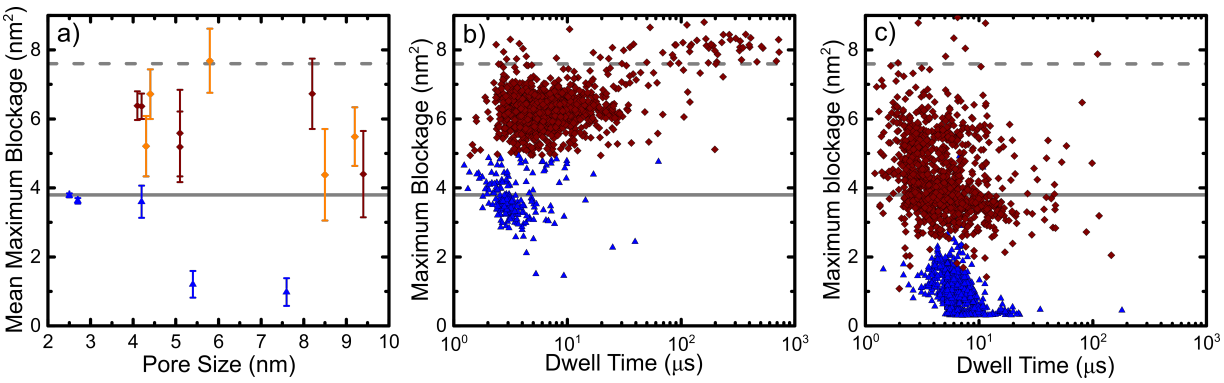

Figure 3.2: *a) Comparison between 50 bp linear dsDNA (blue triangles) and the T-shaped (dark red diamonds) and Y-shaped (orange diamonds) label. Points represent single peak Gaussian fits of the maximum blockage histograms (with N = 515, 133, 178, 187, 993 for the 50 bp measurements and N = 607, 1137, 933, 175, 693, 495, 622, 835, 842, 1037 for the T-shaped and Y-shaped measurements) where the error bars represent the standard deviation. Grey lines indicate calculated areas (assuming diameter of the DNA as 2.2 nm) for unfolded (straight line) dsDNA single and folded (dashed line) translocation. b) Example plots of the maximum blockage vs. dwell time for 50 bp dsDNA and T-shaped label in a 4.0 nm pore. Raw traces have been filtered with a cutoff at 500 kHz for analysis. c) Example plots of the maximum blockage vs. dwell time for 50 bp dsDNA in a 7.7 nm pore and the T-shaped label in a 9.4 nm pore. For analysis, the raw trace for the 50 bp sample was filtered with a cutoff at 250 kHz and the T-shaped data at 500 kHz.*

3.5.2 T-shape Results

Short T-shaped and Y-shaped DNA molecules (figure 3.1a) were tested in the range of pore sizes shown in table 3.3. Experiments were performed by injecting a solution of 50 nM T-shaped or Y-shaped molecule in 3.6 M LiCl solution buffered at pH 8 on the membrane side of the fluidic cell after flushing out the linear dsDNA to achieve a clean baseline. The baseline was considered clean when no translocation events were observed during 2 minutes of current recording.

Table 3.3: *Table of all experiments using T-shaped and Y-shaped molecules. The effective pore diameter is calculated using equation* (3.1) *from the mean baseline level. The uncertainty in the baseline level represents the standard deviation of the current noise. The maximum blockage is extracted from a single peak Gaussian fit to the maximum blockage histogram after converting it to an effective area according to equation* (3.2)*. The uncertainty represents the standard deviation of the single peak fit. For P2, a single peak fit is not applicable as described in the text.*

Pore ID	Applied Voltage (mV)	Sample	Effective Pore Diameter (nm)	Effective Pore Length (nm)	Mean Baseline Level (nS)	Mean Maximum Blockage (nm^2)
P1	200	Click-T	4.1 ± 0.4	5.4 ± 0.3	24.8 ± 1.8	6.4 ± 0.4
P2	400	Click-T	3.5 ± 0.2	10.0 ± 0.3	12.1 ± 0.7	-
P3	200	Click-T	9.4 ± 0.2	11.4 ± 0.3	59.3 ± 0.7	4.4 ± 1.3
P5	200	Click-T	8.2 ± 0.2	12.3 ± 0.3	48.9 ± 0.9	6.7 ± 1.0
P5	200	Origami-T	8.5 ± 0.2	12.3 ± 0.3	48.1 ± 0.9	4.4 ± 1.3
P7	200	Origami-T	4.4 ± 0.4	10.3 ± 0.2	17.5 ± 1.5	6.7 ± 0.7
P7	400	Origami-T	4.3 ± 0.2	10.3 ± 0.2	17.0 ± 0.7	5.2 ± 0.9
P10	200	Origami-T	9.2 ± 0.1	13.4 ± 0.1	52.1 ± 0.7	5.5 ± 0.8
P11	200	Click-T	4.2 ± 0.3	8.5 ± 0.1	18.7 ± 1.1	6.4 ± 0.4
P12	200	Click-T	5.1 ± 0.4	5.6 ± 0.4	34.0 ± 2.0	5.6 ± 1.3
P13	200	Click-T	5.1 ± 0.3	7.6 ± 0.5	28.5 ± 1.3	5.2 ± 1.0

Similar to regular dsDNA, the general behavior of the T-shaped and Y-shaped DNA can be predicted by simple geometric considerations. In pores larger than the label, we expect translocation signatures to possess two levels, and for both levels to line up with the unfolded and folded levels observed for long linear dsDNA rulers. In the case of linear dsDNA, the secondary level is attributed to events with folded conformation during translocation, whereas for the structured DNA labels used here, all translocations should exhibit a deep level comparable to the folded DNA ruler level since a full translocation must at some point have at least two strands of DNA in the pore at the same time. Since the T-shaped and Y-shaped molecules consist of dsDNA strands shorter than their persistence length ($\sim$170 bp for dsDNA)[59], each branch can be regarded as a rigid rod for a simple geometric approximation. As a result, we do not expect to see folding of the branches, implying that any complexity in the observed event signatures will be a direct result of the structure of these molecules.

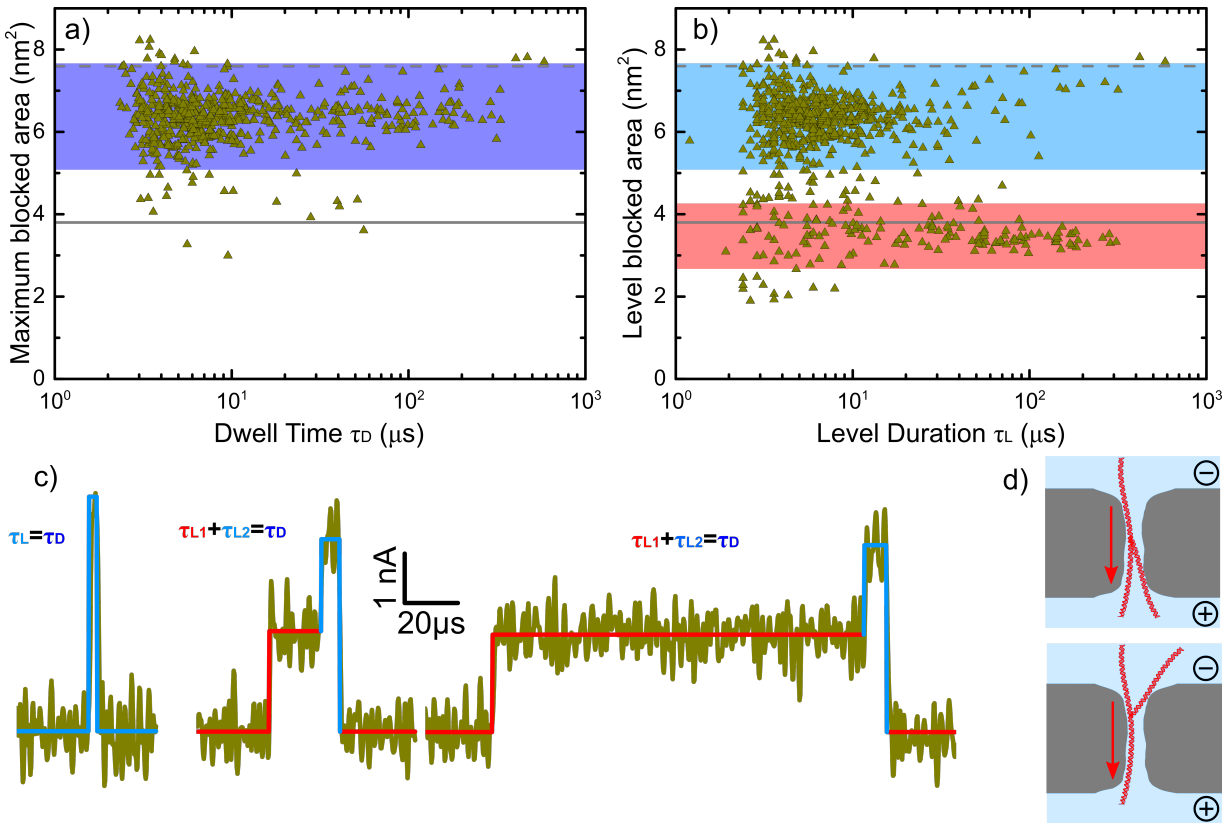

Figure 3.3: *a) Maximum blockage vs. total dwell time for the T-shaped in a 4.0 nm pore at 200 mV. The blue band indicates the mean maximum blockage with the upper and lower bound marking three standard deviations, respectively. b) Blockage for each individual sublevel of all the fitted events vs. the dwell time of that level for the same dataset as in a). In this plot one event can have multiple points depending on the number of fitted levels. The color scheme for the bands is matched to the colors in c) to highlight the type of sublevel c) Sample traces including fits showing the most common event types. Short single level events (left) are the most common type of event (∼ 80%) and are generally deeper than the first blockage state (light blue fit and band in b)). Two-level event types are characterized by a longer dwell time that is caused by the primary level (red fit and band in b)). Colors of fits and legend match the colors of the bands in b) and a). d) Schematic representations of the two proposed possible translocation modes. The top image is the suggested mode for deep single-level events with short dwell times. The bottom image is the suggested mode for longer dwell times and multi-level events.*

Figure 3.3 shows the expected behavior for the T-shaped molecules in pores >4 nm. All events contain sublevels deeper than the unfolded level for linear dsDNA, represented by a horizontal solid line in figure 3.3a. The observed maximum blockage amplitudes do not line up with either of the two quantized levels that we measured from linear DNA ruler experiments, but instead lie in between these unfolded (solid line) and folded (dash line) levels. Figure 3.3b shows the blockage amplitude of all sub-levels with each event, analyzed

by the ADEPT algorithm implemented in MOSAIC (see Data Analysis section). Roughly $\frac{2}{3}$ of the events recorded are short, single-level events deeper than the linear dsDNA level. The remaining events are multi-level events with two distinct blockage levels that can be fitted by ADEPT. The first of these levels lines up with the measured single file blockage level for linear dsDNA on this pore. The secondary level represents a deeper blockage amplitude, implying that more of the pore is blocked, but is below the level of the measured folded blockage for the linear dsDNA ruler (figure 3.3b). We also observe that these primary blockage levels have a significantly broader dwell time distribution, reaching passage times $>100\,\mu s$, while the deeper blockages are much shorter, with passage times often $\lesssim 10\,\mu s$. Dwell times are increased over linear dsDNA translocation, which are expected to be on the order of $\sim 1\,\mu s$ [58, 70], due to increased pore wall interaction. We hypothesize that the longer times observed for the primary blockage levels are a result of the steric hindrance for the branched structure to thread through the pore (figure 3.3d bottom), since this effect vanishes for larger pores. Figure 3.3c shows representative current signatures of these events highlighting the fact that the primary blockage levels are responsible for the long passage times observed for these molecules. It also shows that the secondary levels are generally short, even if a primary level cannot be observed.

These observations can easily be explained by the structure of the molecule. In order for the molecule to successfully translocate, the two strands have to align, since the pore is not big enough for the 25 bp side branch to translocate in a perpendicular configuration ($\sim 9\,nm$). For pores with $d_{pore} > 4\,nm$, we expect to see two different event types. If the molecule is captured by the two strands first (figure 3.3d top), the resulting events should be relatively short and deep (deeper than the linear DNA level). The 25 bp tail of the molecule is most likely translocating too quickly to be detected, since passage is not hindered at that point. On the other hand, when the label is first captured by one of the 25 bp branches (figure 3.3d bottom), the side strand must reorient and align with the pore axis. This results in extended passage times and two distinguishable blockage levels (figure 3.3c), corresponding to one 25 bp strand translocating through the pore (first longer level) before the other strands are finally pulled in (second shorter level). Interestingly, this deeper secondary blockage level is shallower than the observed level for folded linear DNA. This may be due to the thickness of our membrane compared to the label size and to different edge effects between the two molecules.

The translocation characteristics of T-shaped dsDNA labels as a function of pore size are shown in figure 3.4. For pores of relatively small size ($\sim$4 – 5 nm), a long tail in the dwell time distribution is observed when the label is captured by one of the 25 bp strands, as in figure 3.3c. We attribute this to increased interactions between the molecule and the pore walls. The representative event shape for these long events is shown again in figure 3.4b. This indicates that passage of the label is sterically hindered by the side strand orientation. For pores of larger sizes (>9 nm), these interactions will become less prevalent, resulting in a loss of the long tail in the dwell time distribution, which is indeed observed in figure 3.4a. Additionally, the secondary deep blockage level is not present in a large portion of these events. Blockage levels instead line up with the observed level for unfolded linear dsDNA translocation. Looking at the event traces, it is apparent that while the secondary deep level is detected by the sensing hardware, it is not fitted by the fitting algorithm due to its short duration (<2 µs). To verify this, and to better distinguish these T-shaped and Y-shaped labels from linear dsDNA, we characterize these events using the maximum deviation instead of the maximum blockage.

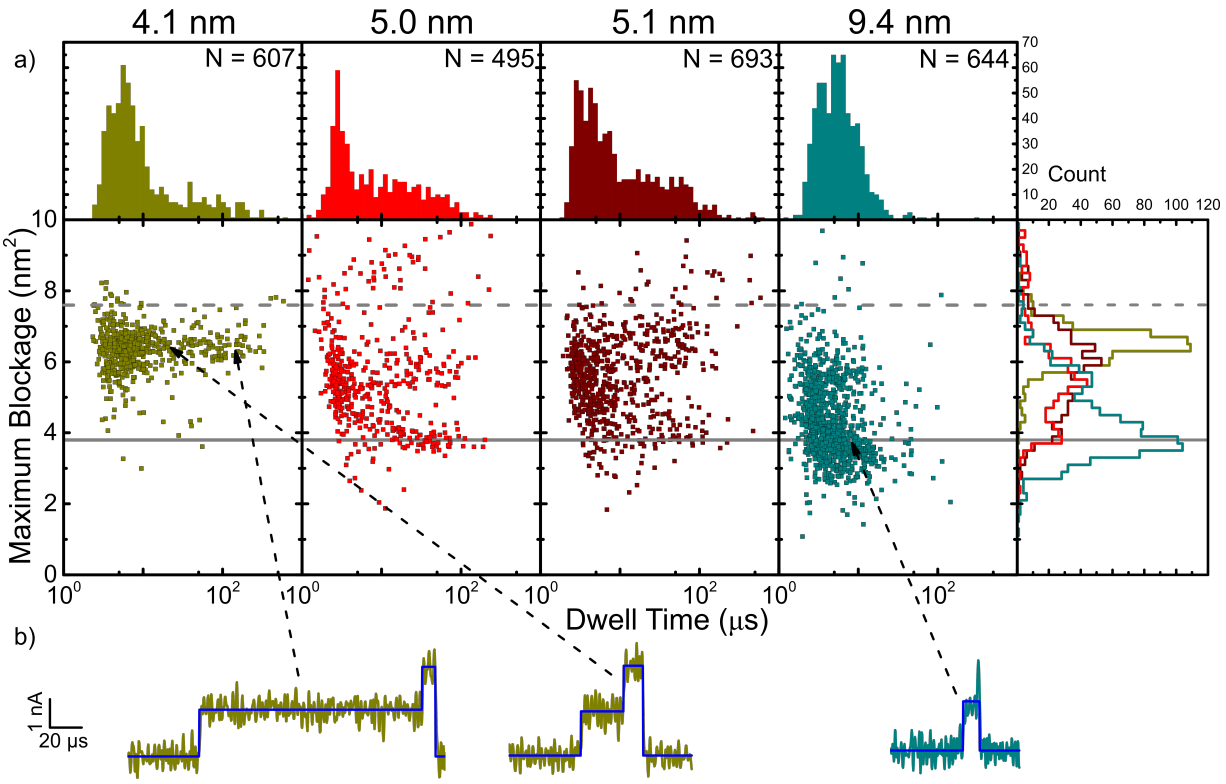

Figure 3.4: *a) Scatter plots and histograms of the maximum blockage vs. total dwell time for pores of increasing size. The grey lines indicate expected blockage levels for linear dsDNA translocating in single file (solid) and folded (dash) fashion. b) Typical ionic current event signatures observed in these experiments. The colors match those from a), indicating the dataset that the event was taken from.*

If the deep blockage is being picked up by the sensing hardware but missed during analysis, we would expect the maximum deviation to only show a single level, ideally lining up with the observed folded level of linear DNA, plus a few standard deviations of the local baseline current. This is indeed observed in figure 3.5. The maximum deviation is extracted from the raw current data at 1 MHz and the mean maximum deviation is obtained from a Gaussian fit to the effective blocked area histogram. Even for the largest pores studied, the mean maximum deviation recovers the deep secondary blockage level and lines up with the folded linear DNA level. This is also indicative of the fact that we are not missing any levels in the current traces. Virtually all translocating molecules are thus detected, since if events were missed due to bandwidth limitations of the sensing hardware, the secondary peak in the maximum deviation histogram would have a long tail toward smaller blockages, indicating that some events were not reaching the full blockage.

At small pore sizes, $\lesssim$4 nm, we also observed a difference in the structure of the events. Generally, events are longer and instead of a step change in current at the start of the

event, the current changes slowly and almost linearly to the maximum blockage level before abruptly jumping back to the baseline current level. Additionally, in these experiments we see a significant number of events that will not reach the expected full blockage level of the unfolded linear dsDNA ruler, which also have shorter dwell times. These events are most likely not full translocations, but rather molecules which transiently interact with the pore and subsequently escape back into solution, likely to be recaptured in a different configuration, until successful translocation. Since the fitting algorithm used is not optimized for events of this shape, analysis is difficult and erroneous sublevels might arise due to the noise in the current recording. Results from pores of these sizes (e.g. P2) are not shown here.

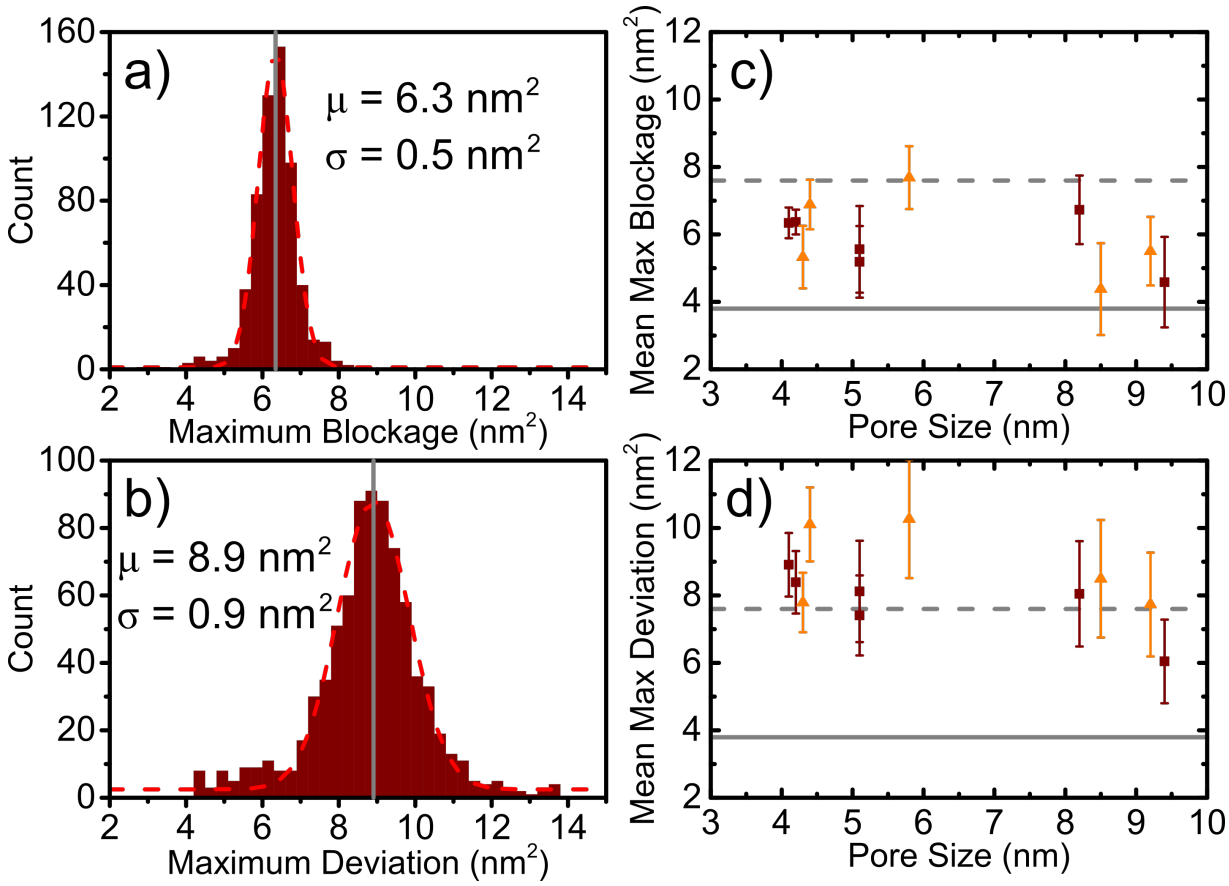

Figure 3.5: *a) Histogram of the maximum effective area blockage extracted from fits using MOSAIC for the T-shaped label in a 4.0 nm pore at 200 mV. A Gaussian curve (red dotted line) was fitted to the histogram and the peak position (grey line) and standard deviation was extracted. b) Same as a) for the maximum deviation. c) Extracted maximum blockage peak positions from the Gaussian fit as a function of pore size. The error bars indicate the standard deviation. Red squares and orange triangles represent the T-shaped and Y-shaped labels respectively. The grey lines indicate expected blockage levels for linear dsDNA single file (solid) and double file translocations (dash). d) same as c) for the maximum deviation.*

We note that nanopores are more susceptible to clogging due to the structure of the T-shaped or Y-shaped molecule compared to linear DNA. This is especially true for smaller pores (stronger pore wall interactions) and at low voltages (weaker electrophoretic forces). The measurement time (i.e. the number of translocations for a given pore) is therefore limited by the pore size and the operating conditions. Generally, pores clogged after translocating ∼500 molecules, representing ∼15 minutes when injecting 50 nM of sample, when the pores were $\lesssim$6 nm. In larger pores, measurement times of more than 1 hour can be achieved and clogged pores are more easily cleared by reversing the voltage polarity. Nevertheless, the number of events detected was sufficient to carry out analysis for all pores presented here.

Furthermore, the recording electronics that was used in these experiments (VC100, Chimera Instruments) imposed a limitation on the maximum pore size for a given choice of electrolyte solution. Since larger pores have a higher baseline current, the voltage has to be adjusted in order to avoid saturating the measurement electronics (±20 nA). For the solution conductivity and pore length used in this study, the upper boundary of accessible pore sizes is ∼10 nm. Translocation kinetics in larger pores are not expected to change from the observed behavior, since most pore wall interactions are already minimized but this would remain to be verified.

Since T-shaped and Y-shaped labels can be more reliably detected in larger pores than short linear dsDNA labels, we can expect the effective detection rate (capture rate of detected molecules) of these structurally defined dsDNA labels to be larger than for short linear dsDNA. Here we show that this increase can be as large as a factor of 10 and thus help decrease significantly the measurement time needed to acquire a statistically relevant number of events. The detection rate for a 50 bp dsDNA label under conditions that guarantee its detection (2.7 nm pore at 200 mV) was measured to be $(0.82 \pm 0.06)\,\mathrm{s}^{-1}$ at a 30 nM concentration, giving a normalized detection rate of $(0.027 \pm 0.002)\,\mathrm{s}^{-1}/\mathrm{nM}$. In contrast, the detection rate for the Y-shaped DNA label was measured at $(11.2 \pm 0.1)\,\mathrm{s}^{-1}$ at a 50 nM concentration for a 8.5 nm pore at 200 mV, thus giving a normalized detection rate of $(0.224 \pm 0.002)\,\mathrm{s}^{-1}/\mathrm{nM}$ (see figure 3.6). This observed ∼7x increase in the effective detection rate can be explained by simple geometric arguments. Since the electrostatic field profile and thus the capture rate of a cylindrical pore, R_C, is proportional to the area of the pore, $\propto d_{\mathrm{pore}}^2$, and inversely proportional to the effective membrane thickness, t_{eff}, the ratio of the capture rates for these two pores should be proportional to $R_{C2}/R_{C1} \sim$

$(d^2_{\text{pore2}} t_{\text{eff1}})/(t_{\text{eff2}} d^2_{\text{pore1}}) \approx 7$. This theoretical factor is in close agreement with the observed increased in detection rate, considering that the actual pores may deviate from a pure cylindrical geometry. In principle, the observed capture rate for the T-shaped and Y-shaped labels could be increased further (up to ∼10x) by using a lower salt concentration solution or a salt gradient [36, 46], permitting higher voltages to be applied without saturating the measurements electronics or extending the electrostatic capture radius.

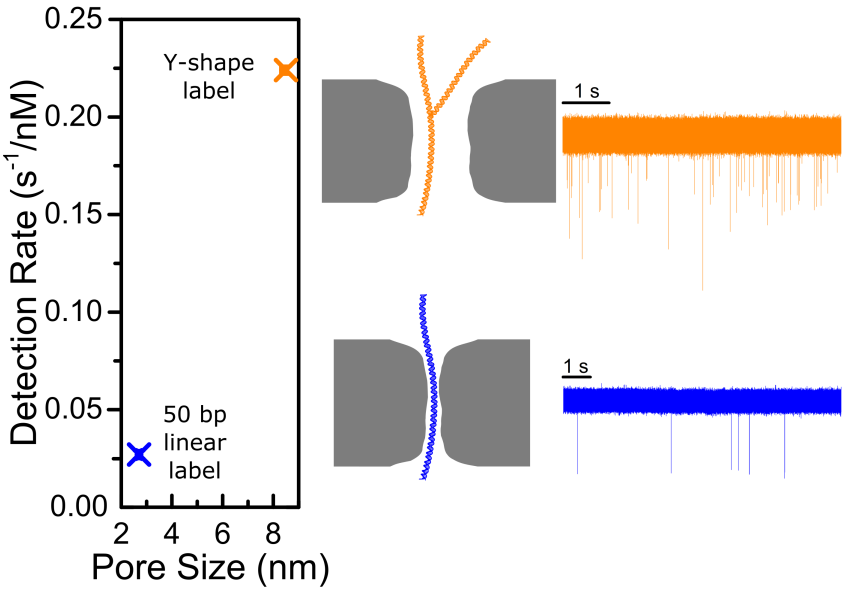

Figure 3.6: *Capture rate of 30 nM 50 bp dsDNA in a 2.7 nm pore (blue) at 200 mV and 50 nM Y-shaped in an 8.5 nm pore (orange) at 200 mV. Both measurements are done in 3.6 M LiCl at pH 8. On the right are schematics of the translocation process and sample traces of the ionic current to highlight the difference in detection rate, recorded on the Axoptach 200B with 100 kHz low-pass filter (blue) and on the Chimera VC100 with 500 kHz low-pass filter (orange).*

3.6 Conclusion

We have shown that a DNA label can be designed, with a small enough structure to reduce entropic complexity (no folding) while still generating deep and consistent blockages in nanopores of different sizes. Short linear dsDNA of the same length scale translocate too fast to be reliably detected under the same conditions. The events generated by the structurally-defined DNA label give rise to richer, readily distinguishable electronic signatures over short linear dsDNA events, that can easily be explained using simple geometric arguments. The consistency of the event depth makes this molecule a good candidate for a

label in generic single-molecule counting assays. Further work is required to ascertain the multiplexing capability of higher-order structures of these short-branched DNA labels.

3.6.1 Funding Sources

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), and by the province of Ontario through their Early Researcher Award to VTC.

3.6.2 Acknowledgments

The authors would like to thank Prof. John Pezacki and Dr. Mirka Strmiskova for valuable discussion regarding click chemistry, and Prof. Charles M. Schroeder for help with DNA nanostructures.

Chapter 4

Additional work

4.1 Holliday junctions as labels

It was presented that T-shaped molecules improve the overall capture and detectability of small molecules in a nanopore experiment. In order to improve even further on this aspect, different molecule geometries are also being considered. The most obvious and simple extension of a T-shaped molecule is a 4-branch star made of DNA. These structures are named *Holliday junctions* after Robin Holliday, who proposed these structures as a mechanism for gene conversion in fungi [71]. In this study Holliday junctions are assembled in a similar fashion to the already described Origami T-shaped molecules, as described in section 1.8, with a fourth oligo hybridized to the existing three (see figure 4.1 and table 4.1).

Table 4.1: *Table of all the oligos and their sequences for the assembly of the Holliday junction. Thea strands are color coded to indicate the complementary parts of the strands.*

Strand	Length	Sequence
Holliday oligo 1	50 nt	5′GTG TCT TGA AGA AGG ACC AGG GAA C GT ATT TGC CTT GTC TGG GAA ATC GT3′
Holliday oligo 2	50 nt	5′ACG ATT TCC CAG ACA AGG CAA ATA C GG ACT TAT CGT TTG CTA ATA CTA CA3′
Holliday oligo 3	50 nt	5′TGT AGT ATT AGC AAA CGA TAA GTC C CC AAC TGG TTG TGG CCT ATC GAA AA3′
Holliday oligo 4	50 nt	5′TTT TCG ATA GGC CAC AAC CAG TTG G GT TCC CTG GTC CTT CTT CAA GAC AC3′

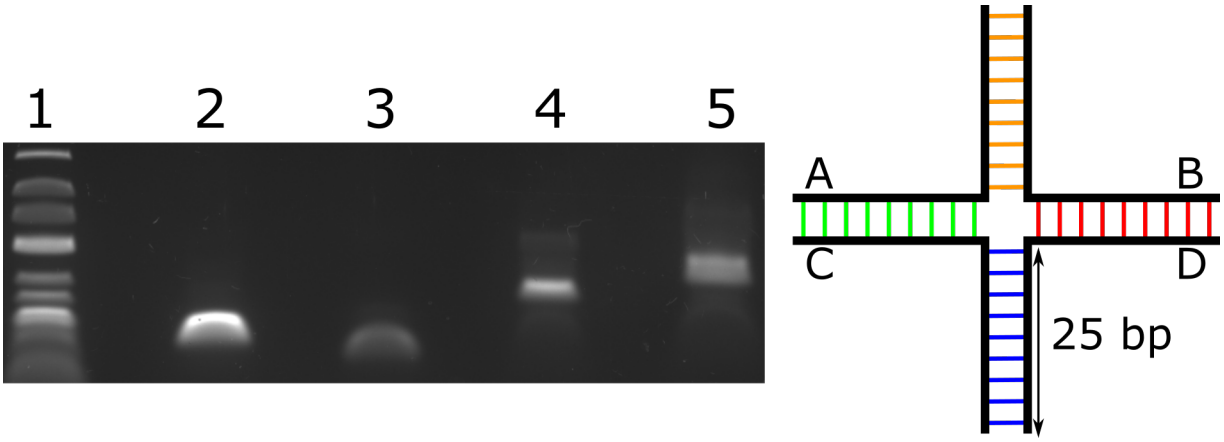

Figure 4.1: *Left: 2% agarose gel of the dsDNA 4-branched star (Holliday junction). The lanes are assigned as follows: 1) Thermo Scientific GeneRuler Low Range DNA Ladder; 2) 50 bp dsDNA; 3) 50 nt ssDNA; 4) Click T-shape; 5) Fully assembled Holliday junction. The shift in the bands towards higher apparent weights is attributed to the structure of these molecules and indicates the correct assembly. Right: Schematic structure of the Holliday junction. The structure consists of four 50 nt oligos (A, B, C, D). Each branch of the star has the same length of 25 bp.*

Due to its importance in genetics the structure and dynamics of Holliday junctions have been studied intensely in recent years [44, 72, 73]. Furthermore Holliday junctions have been proposed for the assembly of more complex DNA structures [74, 75] and have briefly been discussed in a nanopore setting [50].

The Holliday junction has two main conformations and, in free solution, is able to switch back and forth between them. The first conformation is an extended planar cross shaped conformation with all 4 strands forming 90° angles between them. In the second conformation the two opposite strands (A and D, or B and C in figure 4.1) are linearly extended while the other two respectively are folded back on themselves, to make an H like structure [76]. In this conformation the opposing strands can be aligned parallel or antiparallel. For nanopore experiments it is therefore expected to observe three distinct levels for Holliday junction translocation. The three levels are also expected to fall onto the expected blockage for 1, 2, and 3 strands of linear dsDNA blocking the pore at the same time, respectively.

If the molecules are captured in the H like conformation, events are expected to show a single level with the blockage depth being double the linear dsDNA blockage level since two aligned strands of dsDNA are blocking the pore at all times during translocation (*Type 2*). For capture in the open cross conformation events are expected to take a more complex structure. If the molecule is captured at one strand of the four, the other three strands

are dragging behind and the resulting event shape shows a single linear dsDNA blockage first and a triple linear dsDNA blockage after, when all three remaining strands are being pulled through the pore at the same time (*Type 31*). These types of events can potentially, depending on the size of the nanopore and the applied voltage, sterically interact with the sides of the pore and stall translocation, resulting in longer translocation times. For the rare occasion of capturing 3 strands simultaneously it is expected to see an event with a triple linear dsDNA blockage level first and a single linear dsDNA blockage level after (*Type 13*).

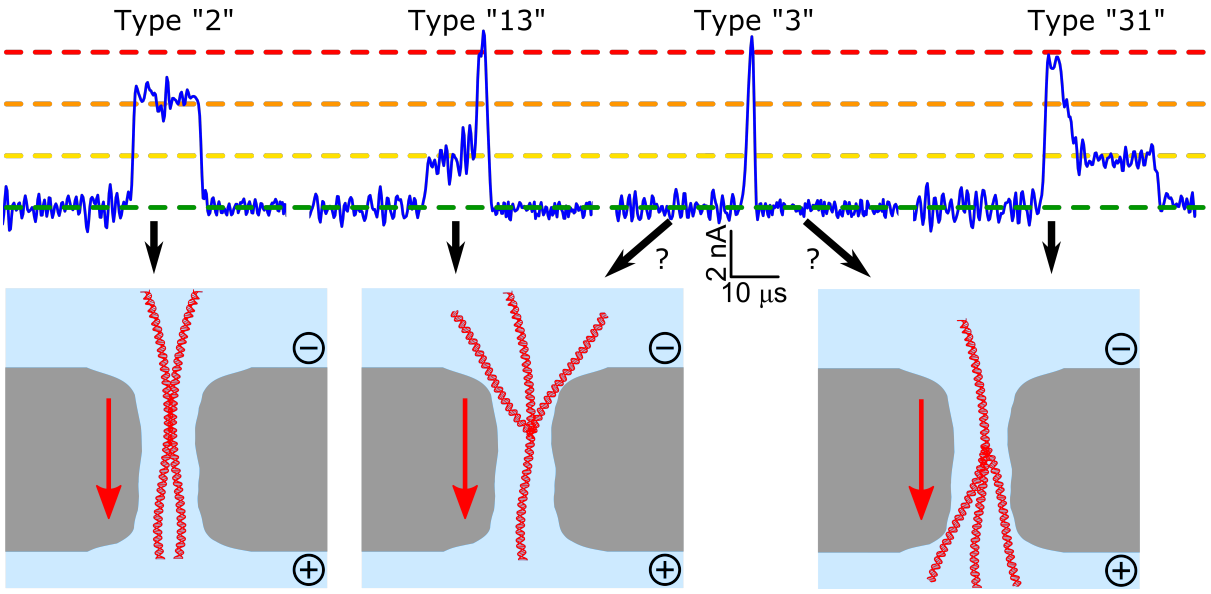

Figure 4.2: *Observed translocation modes for the Holliday junction in a 5.8 nm nanopore with 400 mV bias voltage in 3.6 M LiCl at pH 8. Top: Current traces of the different observed events filtered to 900 kHz. The dashed lines mark the baseline level (green) and the expected linear dsDNA single (yellow), double (orange), and triple (red) blockages. Bottom: Schematics of the interpreted molecular orientation of the Holliday junction based on the observed event shapes and structural considerations.*

Analogous to the explanation for the T-shaped translocations in chapter 3 the secondary single linear dsDNA blockage is likely happening too fast to be reliably detected with the software and hardware used for these experiments, which is especially likely when the molecule gets captured with three strands first. Therefore these types of events will most likely be represented by a single triple linear dsDNA blockage level (*Type 3*). Capture of 2 strands at the same time will result in the molecule being forced into the H like structure described above and cannot be distinguished from it in a nanopore experiment. An overview over possible conformations and resulting event shapes can be seen in figure 4.2.

For larger nanopores or increased voltage, the percentage of Type 3 events is expected to increase compared to compared to Type 13 and Type 31, as steric hindrances are less prevalent. In every case, the maximum blockage for translocation events of Holliday junctions will be at least twice as deep as the linear dsDNA blockage level, increasing SNR for all experimental conditions applied here.

Collective statistics show that these levels are indeed observed, as seen in figure 4.3. The single linear dsDNA blockage (yellow) disappears, when looking at the maximum blockage only, indicating that every event that shows a single level blockage always shows a deeper blockage, too. This behavior was also seen for the T-shaped molecules in chapter 3 and reconfirms the correct assembly of the majority of the molecules.

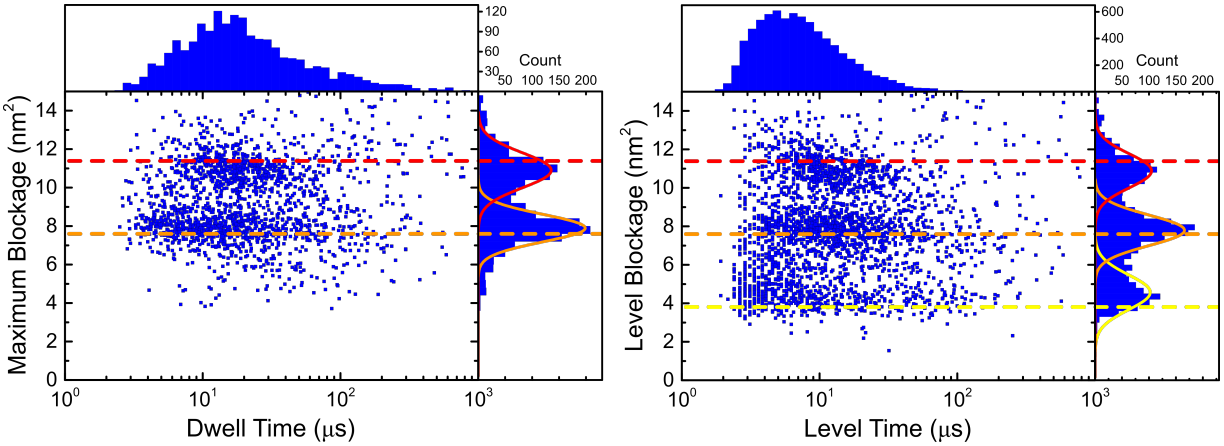

Figure 4.3: *Translocation characteristics for Holliday junctions in a 5.8 nm pore with 200 mV bias voltage in 3.6 M LiCl at pH 8. Left: Maximum blockage vs. dwell time of every event recorded ($N = 1800$). The two peaks in the maximum blockage depth histogram are fitted to Gaussians with $\mu_1 = 7.9\,\mathrm{nm}^2$, $\sigma_1 = 0.7\,\mathrm{nm}^2$, $\mu_2 = 10.9\,\mathrm{nm}^2$, $\sigma_2 = 0.9\,\mathrm{nm}^2$, respectively. Right: Blockage Level vs. level time for every sublevel detected in the same dataset. Three peaks are fitted to the blockage level histogram with $\mu_1 = 4.4\,\mathrm{nm}^2$, $\sigma_1 = 0.9\,\mathrm{nm}^2$, $\mu_2 = 7.1\,\mathrm{nm}^2$, $\sigma_2 = 0.7\,\mathrm{nm}^2$, $\mu_3 = 9.8\,\mathrm{nm}^2$, $\sigma_3 = 1.0\,\mathrm{nm}^2$. The dashed lines mark the expected blockage for the single (yellow), double (orange), and triple (red) linear dsDNA levels.*

Over the range of pore sizes studied here, the Holliday junction presents therefore a clear improvement to the presented T-shape by further increasing blockage depths, and therefore SNR, as indicated in figure 4.4. It also achieves this increase in SNR without further reducing passage times and allows for better detection even when trying to fit the event. Event shapes are still distorted due to the short passage times, but detection for counting purposes is made easier.

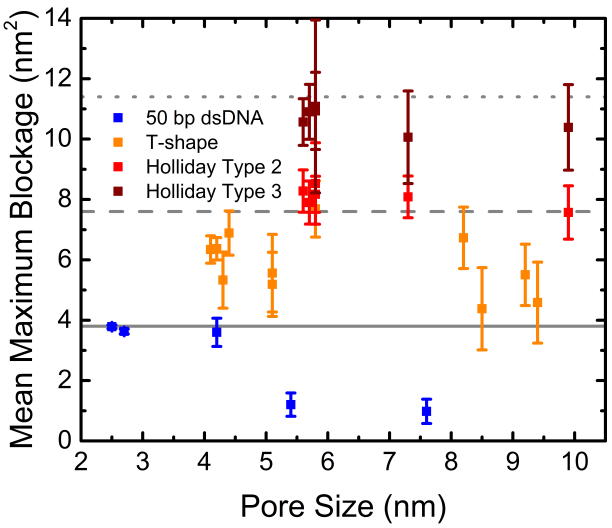

Figure 4.4: *Mean Maximum Blockage vs. Pore Size for all the different molecules tested in this study. The grey lines indicate the blockage levels for single (solid), double (dashed), and triple(dotted) linear dsDNA.*

4.2 Multiplexing with T-shaped molecules

It was presented that structured molecules offer an increased SNR compared to linear strands of dsDNA. Another important aspect of the proposed label is the multiplexing capability of the label. To test this in a proof of principle, an experiment was performed to test the ability to distinguish 50 bp dsDNA and the clicked T-shape. For this purpose a mixture of 24 nM 50 bp dsDNA and 24 nM T-shaped structure was analyzed on a 4.2 nm nanopore at 200 mV in 3.6 M LiCl. At this pore size the linear dsDNA should still be able to be picked up with the full blockage by the sensing hardware and software, while the T-shape only shows a single peak at a deeper blockage.

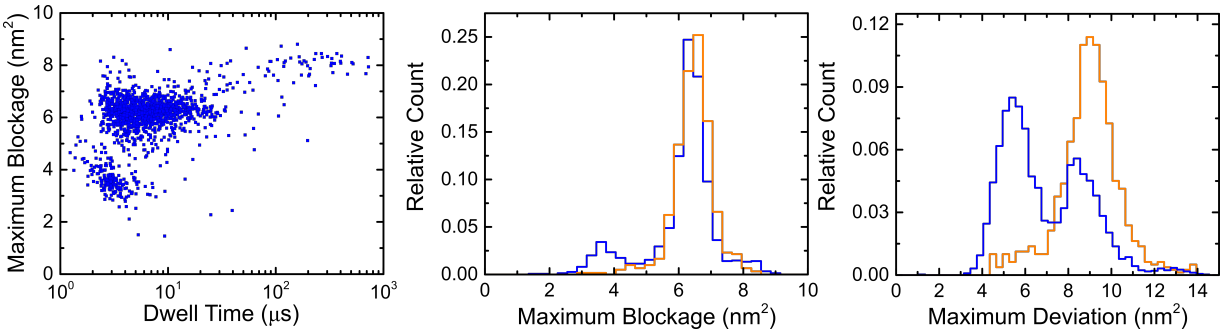

Figure 4.5: *Proof of concept multiplexing with 50 bp dsDNA and T-shaped molecules in a 4.2 nm pore in 3.6 M LiCl at 200 mV. Left: Scatter plot of the maximum blockage vs. dwell time for a mixed sample of 24 nM 50 bp dsDNA and 24 nM Click-T molecule. Center: Histogram of the maximum blockage for the mixed sample (blue) and a control sample (orange) of only 50 nM T-shaped molecules. The control sample is the same as pore P1 from chapter 3. Right: Histogram of the maximum deviation for the same data sets.*

The results are summarized by figure 4.5. Two populations are indeed observed in the maximum blockage vs. dwell time scatter plot and the histograms (blue). The mean of the maximum blockage depth of both populations also agrees with previous experiments on 50 bp dsDNA and T-shape in similar sized pores (see figure 4.4). As expected the 50 bp peak lines up with the single dsDNA blockage level, observed for linear dsDNA, while the T-shape blockage lines up with the deeper blockage level observed in chapter 3. Therefore, separation can easily be achieved using the blockage depth. It should be noted, while it is possible to separate the two populations based on blockage depth, that the majority of 50 bp events are going too fast to be detected reliably and therefore quantization is not yet achievable.

Overall it is possible to separate translocations of linear strands of 50 bp dsDNA from T-shape dsDNA molecules based on blockage depth, an important step for multiplexing applications. However, correct quantization of both species under the same experimental condition is still a challenging task and is subject of on-going work in the lab. A 2-plex assay using 50 bp dsDNA and the T-shape as labels would therefore require a very precisely tuned pore size $\sim$4 nm for it to work.

Chapter 5

Discussion and Outlook

The objective of this thesis was the development of small and simple labels for nanopore experiments with applications in single-molecule counting based assays. Three different labels consisting of short strands of dsDNA have been proposed and tested in this study. The structure of these labels resulted in a direct increase in the observed SNR, which allows higher bandwidth measurements (which comes at the cost of higher noise) and therefore better temporal resolution. This, in turn, allows for the use of larger pores even for short molecules, that usually translocate too fast to be reliably detected under those conditions. Event structures observed for these molecules are more complex, than those of their linear counter-parts, but are still kept simple enough that simple geometric and energetic considerations suffice to describe and explain the observed behavior. Multiplexing capabilities of the T-shaped molecules have been shown in comparison to linear 50 bp dsDNA, but further testing is needed to fully explore the potential of these structures.

All the labels presented here can easily be synthesized in a bottom-up fashion and all the parts needed to assemble the shown structures are commercially available, giving easy customization for future projects and makes this approach flexible for a range of potential applications. Future research will focus on building similar structures to keep improving SNR in nanopore measurements and also build up an arsenal of molecules for potential multiplexing assays.

In addition to the proposed use of these structures as reliable labels for single-molecule counting based assays, they can easily be integrated in standard DNA origami schemes and scaled up to longer molecules and larger structures. Thus these structures can be applied as an alternative to already existing labeling schemes such as the one proposed by Bell *et al.*[41]. In combination with high bandwidth electronics and short pores for increased

SNR the labeling scheme, presented here, has the potential to measure velocity fluctuations during translocation of DNA through a solid-state pore. The ability to measure the velocity of the DNA as it passes through the pore is an important step to enable controlled motion through the pore, one of the main hurdles to overcome on the way to solid-state nanopore DNA sequencing applications [42, 49, 70].

Appendix A

Click Reaction Mechanism

The label is supposed to generate multi-level events in the current trace that can be detected by the sensing setup. To allow for easy upscaling of the label design to longer strands with multiple side branches, the designed branch needs to be sequence specific in order to know its exact position on the DNA strand. Furthermore the branch needs to have a certain stiffness to be detectable even in the presence of noise. The proposed branch for this thesis was therefore chosen to be a short (25 bp) double stranded piece of DNA. To form a covalent bond between the branch and the main DNA strand the Copper(I)-Catalysed Azide Alkyne Cycloaddition (CuAAC, "click" chemistry) was chosen. In order to perform this reaction the backbone strand of DNA has one of the middle bases replaced with a DBCO-dUTP while the branch strand is modified with an azide group at one of the ends. Under addition of Cu^{+}-ions these groups will react to form a covalent bond between the two parts of the structure. The basic reaction scheme can be seen in figure A.1. The modified DNA strands were provided by Baseclick GmbH (Neuried, Germany) and the reaction protocol was adapted from their website [77].

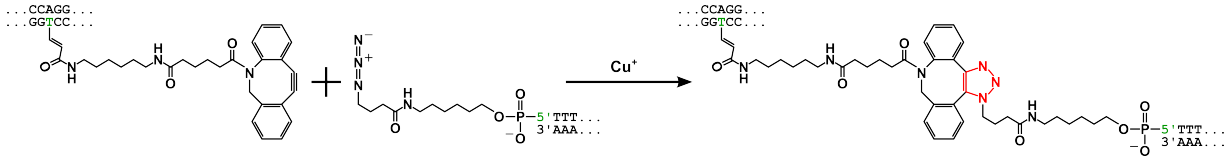

Figure A.1: *Schematic of click reaction*

It was attempted to perform the reaction in house and build the structure directly in the lab, but due to missing equipment specifically for clean up and purification of the reaction mix, it was decided to go with a commercially purchased fully assembled product synthesized by BioSynthesis Inc.

References

[1] Hagan Bayley and P. S. Cremer. Stochastic Sensors Inspired by Biology. *Nature*, 413(September):226–230, 2001.

[2] Nicholas A W Bell and Ulrich F. Keyser. Nanopores formed by DNA origami: A review. *FEBS Letters*, 588(19):3564–3570, 2014.

[3] Hagan Bayley. Nanopore sequencing: From imagination to reality. *Clinical Chemistry*, 61(1):25–31, 2015.

[4] Wallace H Coulter. Means for counting particles suspended in a fluid, October 20 1953. US Patent 2,656,508.

[5] A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker. Fabrication of solid-state nanopores with single-nanometre precision. *Nature Materials*, 2(8):537–540, 2003.

[6] Harold Kwok, Kyle Briggs, and Vincent Tabard-Cossa. Nanopore fabrication by controlled dielectric breakdown. *PLoS ONE*, 9(3):e92880, 2014.

[7] I Polishchuk, King Tsu-Jae, and Chenming Hu. Physical origin of SILC and noisy breakdown in very thin silicon nitride gate dielectric. *Device Research Conference Notre Dame, In, USA. Conference Digest*, (1999):20–21, 2001.

[8] Adam Zrehen, Tal Gilboa, and Amit Meller. Real-Time Visualization and Sub-Diffraction Limit Localization of Nanometer-Scale Pore Formation by Dielectric Breakdown. *Nanoscale*, 9:16437–16445, 2017.

[9] Itaru Yanagi, Rena Akahori, Toshiyuki Hatano, and Ken-ichi Takeda. Fabricating nanopores with diameters of sub-1 nm to 3 nm using multilevel pulse-voltage injection. *Scientific Reports*, 4(5000):1–7, 2014.

[10] Y. M. Nuwan, D. Y. Bandara, Buddini Iroshika Karawdeniya, and Jason R. Dwyer. Real-time Profiling of Solid-State Nanopores During Solution-Phase Nanofabrication. *ACS Applied Materials & Interfaces*, 8(44):30583–30589, 2016.

[11] Min Jun Kim, Ben McNally, Kazuyoshi Murata, and Amit Meller. Characteristics of solid-state nanometre pores fabricated using a transmission electron microscope. *Nanotechnology*, 18(205302):1–5, 2007.

[12] Meng-yue Wu, Ralph M M Smeets, Mathijs Zandbergen, Ulrike Ziese, Philip E Batson, Nynke H Dekker, and Cees Dekker. Control of Shape and Material Composition of Solid-State Nanopores. *Nano Letters*, 9(1):479–484, 2009.

[13] Stefan W Kowalczyk, Alexander Y Grosberg, Yitzhak Rabin, and Cees Dekker. Modeling the conductance and DNA blockade of solid-state nanopores. *Nanotechnology*, 22(315101):1–5, 2011.

[14] Igor Vodyanoy and Sergey M. Bezrukov. Sizing of an ion pore by access resistance measurements. *Biophysical Journal*, 62:10–11, 1992.

[15] Eric Beamish, Harold Kwok, Vincent Tabard-Cossa, and Michel Godin. Precise control of the size and noise of solid-state nanopores using high electric fields. *Nanotechnology*, 23(405301):1–7, 2012.

[16] Adrian Balan, Bartholomeus Machielse, David Niedzwiecki, Jianxun Lin, Peijie Ong, Rebecca Engelke, Kenneth L. Shepard, and Marija Drndic. Improving signal-to-noise performance for DNA translocation in solid-state nanopores at MHz bandwidths. *Nano Letters*, 14(12):7215–7220, 2014.

[17] Vincent Tabard-Cossa, Dhruti Trivedi, Matthew Wiggin, Nahid N. Jetha, and Andre Marziali. Noise analysis and reduction in solid-state nanopores. *Nanotechnology*, 18(305505):1–6, 2007.

[18] V Dimitrov, U Mirsaidov, D Wang, T Sorsch, W Mansfield, J Miner, F Klemens, R Cirelli, S Yemenicioglu, and G Timp. Nanopores in solid-state membranes engineered for single molecule detection. *Nanotechnology*, 21(065502):1–11, 2010.

[19] Vincent Tabard-Cossa. Instrumentation for Low-Noise High-Bandwidth Nanopore Recording. In *Engineered Nanopores for Bioanalytical Applications: A Volume in Micro and Nano Technologies*, pages 59–93. Elsevier Inc., New York City, 2013.

[20] J. B. Johnson. Thermal Agitation of Electricity in Conductors. *Physical Review*, 32(1):97–109, 1928.

[21] H. Nyquist. Thermal agitation of electric charge in conductors. *Physical Review*, 32(1):110–113, 1928.

[22] Chenyu Wen, Shuangshuang Zeng, Kai Arstila, Timo Sajavaara, Yu Zhu, Zhen Zhang, and Shi-Li Zhang. Generalized Noise Study of Solid-State Nanopores at Low Frequencies. *ACS Sensors*, 2:300–307, 2017.

[23] Hadi Arjmandi-Tash, Liubov A. Belyaeva, and Grégory F. Schneider. Single molecule detection with graphene and other two-dimensional materials: nanopores and beyond. *Chemical Society Reviews*, 45:476–493, 2015.

[24] Kyle Briggs. GitHub Page for CUSUM+ implementation. `https://github.com/shadowk29/CUSUM`, 2014. Accessed: 2018-01-23.

[25] C Raillon, P Granjon, M Graf, L J Steinbock, and A Radenovic. Fast and automatic processing of multi-level events in nanopore translocation experiments. *Nanoscale*, 4(16):4916–4924, 2012.

[26] Arvind Balijepalli, Jessica Ettedgui, Andrew T Cornio, Joseph W F Robertson, Kin P Cheung, John J Kasianowicz, and Canute Vaz. Quantifying Short-Lived Events in Multistate Ionic Current Measurements. *ACS Nano*, 8(2):1547–1553, 2015.

[27] Jacob H Forstater, Kyle Briggs, Joseph W F Robertson, Jessica Ettedgui, Olivier Marie-Rose, Canute Vaz, John J Kasianowicz, Vincent Tabard-Cossa, and Arvind Balijepalli. MOSAIC: A Modular Single Molecule Analysis Interface for Decoding Multi-state Nanopore Data. *Analytical Chemistry*, 88:11900–11907, 2016.

[28] E. S. Page. Continuous Inspection Schemes. *Biometrika*, 41(1):100–115, 1954.

[29] Arvind Balijepalli, Kyle Briggs, Jacob Forstater, Joseph Robertson, and Canute Vaz. MOSAIC: A Modular Single Molecule Analysis Interface. `https://pages.nist.gov/mosaic/`, 2012. Accessed: 2017-12-01.

[30] James Watson and Francis Crick. Molecular structure of nucleic acids. *Nature*, 171(4356):737–738, 1953.

[31] J. D. Watson and F. H. C. Crick. The Structure of DNA. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 18, pages 123–131, Cold Spring Harbor, 1953.

[32] John A. Schellman and Dirk Stigter. Electrical double layer, zeta potential, and electrophoretic charge of double-stranded DNA. *Biopolymers*, 16(7):1415–1434, 1977.

[33] D. Porschke. Persistence length and bending dynamics of DNA from electrooptical measurements at high salt concentrations. *Biophysical Chemistry*, 40:169–179, 1991.

[34] Stijn van Dorp, Ulrich F. Keyser, Nynke H. Dekker, Cees Dekker, and Serge G. Lemay. Origin of the electrophoretic force on DNA in solid-state nanopores. *Nature Physics*, 5(5):347–351, 2009.

[35] Alexander Y. Grosberg and Yitzhak Rabin. DNA capture into a nanopore: Interplay of diffusion and electrohydrodynamics. *Journal of Chemical Physics*, 133(165102):1–15, 2010.

[36] Meni Wanunu, Will Morrison, Yitzhak Rabin, Alexander Y Grosberg, and Amit Meller. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. *Nature nanotechnology*, 5(2):160–165, 2010.

[37] M. Muthukumar. Theory of capture rate in polymer translocation. *Journal of Chemical Physics*, 132(195101):1–10, 2010.

[38] Bo Lu, Fernando Albertorio, David P. Hoogerheide, and Jene A. Golovchenko. Origins and consequences of velocity fluctuations during DNA passage through a nanopore. *Biophysical Journal*, 101(1):70–79, 2011.

[39] Meni Wanunu, Jason Sutin, Ben McNally, Andrew Chow, and Amit Meller. DNA translocation governed by interactions with solid-state nanopores. *Biophysical journal*, 95(10):4716–4725, 2008.

[40] Trevor J. Morin, Tyler Shropshire, Xu Liu, Kyle Briggs, Cindy Huynh, Vincent Tabard-Cossa, Hongyun Wang, and William B. Dunbar. Nanopore-Based Target Sequence Detection. *PloS one*, 11(5):1–21, 2016.

[41] Nicholas A. W. Bell and Ulrich F Keyser. Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. *Nature nanotechnology*, 11:645–651, 2016.

[42] Calin Plesa, Nick Van Loo, Philip Ketterer, Hendrik Dietz, and Cees Dekker. Velocity of DNA during translocation through a solid-state nanopore. *Nano Letters*, 15(1):732–737, 2015.

[43] Eric Beamish, Vincent Tabard-Cossa, and Michel Godin. Identifying structure in short DNA scaffolds using solid-state nanopores. *ACS Sensors*, 2(12):1814–1820, 2017.

[44] Nadrian C. Seeman. Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, 99(2):237–247, 1982.

[45] Shana O. Kelley, Chad A. Mirkin, David R. Walt, Rustem F. Ismagilov, Mehmet Toner, and Edward H. Sargent. Advancing the speed, sensitivity and accuracy of biomolecular detection using multi-length-scale engineering. *Nature Nanotechnology*, 9(12):969–980, 2014.

[46] Mohammad Amin Alibakhshi, Justin R. Halman, James Wilson, Aleksei Aksimentiev, Kirill A. Afonin, and Meni Wanunu. Picomolar Fingerprinting of Nucleic Acid Nanoparticles Using Solid-State Nanopores. *ACS Nano*, 11(10):9701–9710, 2017.

[47] Jae Seok Yu, Min Cheol Lim, Duyen Thi Ngoc Huynh, Hyung Jun Kim, Hyun Mi Kim, Young Rok Kim, and Ki Bum Kim. Identifying the location of a single protein along the DNA strand using solid-state nanopores. *ACS Nano*, 9(5):5289–5298, 2015.

[48] Alon Singer, Srinivas Rapireddy, Danith H. Ly, and Amit Meller. Electronic barcoding of a viral gene at the single-molecule level. *Nano Letters*, 12(3):1722–1728, 2012.

[49] Nicholas A W Bell and Ulrich F Keyser. Direct measurements reveal non-Markovian fluctuations of DNA threading through a solid-state nanopore. *arXiv*, arXiv:1607:1–5, 2016.

[50] Spencer Carson, Scott T. Wick, Peter A. Carr, Meni Wanunu, and Carlos A. Aguilar. Direct Analysis of Gene Synthesis Reactions Using Solid-State Nanopores. *ACS Nano*, 9(12):12417–12424, 2015.

[51] Jinglin Kong, Nicholas A. W. Bell, and Ulrich F. Keyser. Quantifying Nanomolar Protein Concentrations Using Designed DNA Carriers and Solid-State Nanopores. *Nano Letters*, 16(6):3557–3562, 2016.

[52] Yuechuan Zhang, Yanling Chen, Yongqi Fu, Cuifeng Ying, Yanxiao Feng, Qimeng Huang, Chao Wang, De Sheng Pei, and Deqiang Wang. Monitoring tetracycline through a solid-state nanopore sensor. *Scientific Reports*, 6:4–10, 2016.

[53] Meni Wanunu, Jason Sutin, and Amit Meller. DNA Profiling Using Solid-State Nanopores : Detection of DNA-Binding Molecules. *Nano Letters*, 9(10):3498–3502, 2009.

[54] Yao Lin, Yi-Lun Ying, Xin Shi, Shao-Chuang Liu, and Yi-Tao Long. Direct sensing of cancer biomarkers in clinical samples with a designed nanopore. *Chem. Commun.*, 53(84):11564–11567, 2017.

[55] Calin Plesa, Justus W. Ruitenberg, Menno J. Witteveen, and Cees Dekker. Detection of individual proteins bound along DNA using solid-state nanopores. *Nano Letters*, 15(5):3153–3158, 2015.

[56] Fanny Wang, Osama K Zahid, Brandi Swain, Derek Parsonage, Thomas Hollis, Scott Harvey, Dr. Fred Perrino, Rahul M Kohli, Ethan W Taylor, and Adam R. Hall. Solid-state nanopore analysis of diverse DNA base modifications using a modular enzymatic labeling process. *Nano Letters*, page acs.nanolett.7b03911, 2017.

[57] David R. Walt. Optical methods for single molecule detection and analysis. *Analytical Chemistry*, 85(3):1258–1263, 2013.

[58] Spencer Carson, James Wilson, Aleksei Aksimentiev, and Meni Wanunu. Smooth DNA transport through a narrowed pore geometry. *Biophysical Journal*, 107(10):2381–2393, 2014.

[59] Earle Stellwagen and Nancy C Stellwagen. Determining the electrophoretic mobility and translational diffusion coefficients of DNA molecules in free solution. *Electrophoresis*, 23:2794–2803, 2002.

[60] Arnold J. Storm, Cornelis Storm, Jianghua Chen, Henny Zandbergen, Jean Francois Joanny, and Cees Dekker. Fast DNA translocation through a solid-state nanopore. *Nano Letters*, 5(7):1193–1197, 2005.

[61] Siddharth Shekar, David J Niedzwiecki, Chen-Chi Chien, Peijie Ong, Daniel Adam Fleischer, Jianxun Lin, Jacob K. Rosenstein, Marija Drndic, and Kenneth L. Shepard. Measurement of DNA translocation dynamics in a solid-state nanopore at 100-ns temporal resolution. *Nano Letters*, 16(7):4483–4489, 2016.

[62] Jacob K. Rosenstein, Meni Wanunu, Christopher A. Merchant, Marija Drndic, and Kenneth L. Shepard. Integrated nanopore sensing platform with sub-microsecond temporal resolution. *Nature Methods*, 9(5):487–492, 2012.

[63] Stefan W. Kowalczyk, David B. Wells, Aleksei Aksimentiev, and Cees Dekker. Slowing down DNA translocation through a nanopore in lithium chloride. *Nano Letters*, 12(2):1038–1044, 2012.

[64] Matthew Waugh, Autumn Carlsen, David Sean, Gary W. Slater, Kyle Briggs, Harold Kwok, and Vincent Tabard-Cossa. Interfacing solid-state nanopores with gel media to slow DNA translocations. *Electrophoresis*, 36(15):1759–1767, 2015.

[65] Yuhui He, Makusu Tsutsui, Ralph H. Scheicher, Fan Bai, Masateru Taniguchi, and Tomoji Kawai. Thermophoretic manipulation of DNA translocation through nanopores. *ACS Nano*, 7(1):538–546, 2013.

[66] Jiali Li and David S Talaga. The distribution of DNA translocation times in solid-state nanopores. *Journal of physics. Condensed matter: an Institute of Physics journal*, 22(454129):1–8, 2010.

[67] Jasmine Y.Y. Sze, Aleksandar P. Ivanov, Anthony E.G. Cass, and Joshua B. Edel. Single molecule multiplexed nanopore protein screening in human serum using aptamer modified DNA carriers. *Nature Communications*, 8(1):1–10, 2017.

[68] Amanda B. Marciel, Danielle J. Mai, and Charles M. Schroeder. Template-directed synthesis of structurally defined branched polymers. *Macromolecules*, 48(5):1296–1303, 2015.

[69] Ralph M M Smeets, Ulrich F. Keyser, Diego Krapf, Meng Yue Wu, H. Nynke, Cees Dekker, and Nynke H. Dekker. Salt-dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett.*, 6(1):89–95, 2006.

[70] Kyle Briggs, Gregory Madejski, Martin Magill, Konstantinos Kastritis, Hendrick W de Haan, James L McGrath, and Vincent Tabard-Cossa. DNA Translocations Through Nanopores Under Nanoscale Pre-Confinement. *Nano Letters*, 18(2):660–668, 2018.

[71] Robin Holliday. A mechanism for gene conversion in fungi. *Genetics Research*, 5(1964):282–304, 1964.

[72] David M. J. Lilley. Structures of helical junctions in nucleic acids. *Quarterly Reviews of Biophysics*, 33(2):109–159, 2000.

[73] B F Eichman, J M Vargason, B H Mooers, and P S Ho. The Holliday junction in an inverted repeat DNA sequence: sequence effects on the structure of four-way junctions. *Proceedings of the National Academy of Sciences of the United States of America*, 97(8):3971–3976, 2000.

[74] N C Seeman and N R Kallenbach. DNA branched junctions. *Annual review of biophysics and biomolecular structure*, 23:53–86, 1994.

[75] Di Liu, Gang Chen, Usman Akhter, Timothy M. Cronin, and Yossi Weizmann. Creating complex molecular topologies by configuring DNA four-way junctions. *Nature Chemistry*, 8(July):1–8, 2016.

[76] Jin Yu, Taekjip Ha, and Klaus Schulten. Conformational model of the Holliday junction transition deduced from molecular dynamics simulations. *Nucleic Acids Research*, 32(22):6683–6695, 2004.

[77] Baseclick Homepage. `https://www.baseclick.eu/technology.php?ID=13`. Accessed: 2018-01-26.