Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System

Alexandra Helen Dias-Lalcaca

Department of Chemistry

Imperial College London

A thesis submitted in partial fulfilment of the requirement for the degree of Doctor of Philosophy

ABSTRACT

The continued development of diagnostic and therapeutic techniques for biological samples requires robust single-molecule detection techniques. Protein assays that are currently commonly used, such as ELISA, do not fit the specificity and sensitivity requirements, and also involve significant sample processing. In short, they do not provide what is needed for biomarker qualification and quantification.

One technique that has the potential to offer improved single-molecule detection is nanopore sensing. Nanopore sensing uses electrical voltage to drive molecules across a small pore, one that is either biological or solid state. For this research, as solid state pore at the end of a pipette, a nanopipette, is used. In this technique, the current is continuously monitored and as specific molecules, usually DNAs or proteins, cross through the pore (translocate) the current changes. As nanopore sensing is a single-molecule technique, it easily fulfils the sensitivity requirements. Unfortunately, on its own, it is not able to select for specific molecules, and this limitation, in addition to difficulties that are encountered with protein translocations, leads to the necessity of DNA carriers. DNA carriers are typically double-stranded DNAs that has been modified to bind specifically to certain proteins. The translocations of these DNA tethered protein molecules then look significantly different compared to DNA on its own, such that it is possible to select for them specifically. The distinguishing figure for these events is typically another peak inside these events, or rather a subpeak. This research investigates two DNA carriers, a DNA dendrimer and a DNA plasmid carrier.

The DNA dendrimer has great potential due to its customisability and ability to perform multiplexed sensing. To form the dendrimer, Y-shaped DNAs, each made of three oligonucleotides, was combined in stoichiometric ratios. For the first generation (G1) dendrimer, four Y-shaped DNAs were combined. This G1 could have three protein binding sites or be combined with another six Y-shaped DNAs to form a second generation (G2) with six protein binding sites. However, unfortunately, at the low concentrations required for the nanopore, the binding of the oligonucleotides was not strong enough. Many of the dendrimer structures simply fell apart or could not stay in a reliable shape with the required salt conditions. Therefore, there were few translocation events recorded and it was established that the dendrimers were not feasible DNA carriers.

The plasmid, like the dendrimer, is completely customisable and has the potential to perform multiplexed protein sensing. A 10 Kbp plasmid was modified such that a short single-stranded section was replaced with a sequence including both the original sequence as well as a biotin. Theoretically, any aptamer could be used in place of the biotin, but the biotin-streptavidin bond was used for the proof of concept. The plasmid was also linearised.

Successful modification of the plasmid so that it could bind to streptavidin was confirmed. Several protein binding curves with both monovalent streptavidin and quadrivalent streptavidin were performed, and events with subpeaks were regularly identified. Additionally, a binding curve for a biotinylated phosphatase bound to quadrivalent streptavidin was performed. This allowed for an investigation of a sandwich-like assay as it moved through the nanopore. While there was a significant difference in the binding curve for this sandwich-like assay, individual events did not have as many differences as was expected.

This plasmid carrier has a lot of potential, as it overcomes the issue of introducing selectivity into the nanopore platform without diminishing its sensitivity. The long term goal for this plasmid carrier is for it to be used to identify specific proteins, ideally with multiplexed sensing capabilities, at extremely low concentrations.

DECLARATION OF ORIGINALITY

I, the author, declare that this thesis presents work that is original and was performed by myself; otherwise it is clearly cited and acknowledged. This work has not previously been submitted in any form to satisfy any degree requirements at this or any other university.

COPYRIGHT DECLARATION

ACKNOWLEDGMENTS

Firstly, thanks are due to Prof. Joshua Edel and Dr. Alex Ivanov for all of their support and guidance throughout this PhD. Their encouragement and motivation even when things were not going to plan were invaluable and it was a privilege to work with them both.

Thanks to Binoy and Paolo for everything, their help in and out of the lab was invaluable and working with them was a joy; Binoy, keep Philosotea going! Giulia thank you for always being around to discuss new ideas, Debjani thank you for always being willing to double check logic and math, Annie thank you for helping with all of the fluorescent experiments, Shenglin thank you for all of the help, particularly when at the beginning of the process, and Ren thank you for answering many questions about the amplifiers and helping with fixing any issues.

Additional thanks are due to the wonderful people who agree to proofread this thesis, Annie, Debjani, and Caroline. Thank you again and sorry some of the chapters were a bit long. Their help was invaluable and it was wonderful spending time in the lab with them.

To all the members of the Edel group, past and present, thank you for making the lab a wonderful place to work. Debjani, Caroline, Annie, Tony, Zoe, Aleksandra, Alfie, Xiaoyi, Oliver, Yilin, Yaxian good luck with the rest of the PhD experience. Micol thank you so much for being there from the beginning, this PhD experience would not have been the same otherwise.

It would be remiss to not thank my wonderful brother Kieran. Thank you for proofreading drafts this year and always being encouraging.

Finally, thank you to my parents. Without all the support and love, nothing would have been achieved. Thank you for everything Mum and Dad!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

$\boldsymbol{\alpha}$-HL	Alpha-Hemolysin
$\boldsymbol{\delta}$	Debye Length
ε_ρ	Permittivity of the Medium
$\eta_{KCl}(c)$	Number Density
μ_{Cl^-}	Mobility of Cl^- Ions
μ_{K^+}	Mobility of K^+ Ions
$\boldsymbol{\sigma}$	Solution Conductivity
σ_p	Surface Charge Density of the Nanopore
a1	15 Base Thrombin Aptamer
a2	29 Base Thrombin Aptamer
AeL	Aerolysin
AFM	Atomic Force Microscopy
Ag	Silver
AgCl	Silver Chloride
Ar	Argon
BO	Biotin Oligonucleotide
BO2	Two Biotin Oligonucleotide
Boph	Biotin Oligonucleotide with Phosphorylation
bp	Base Pairs
BSA	Bovine Serum Albumin
c_0	Bulk Ion Concentration
^oC	Degrees Celsius
Cl	Chlorine
Cl^-	Chlorine Ions
CO_2	Carbon Dioxide
D	Diameter of the Pore
d_b	Diameter of the Pipette before Convergence
d_i	Capillary Diameter at the Tip
D_i	Inner Diameter at the Base
dI water	Deionised Water

DNA	Deoxyribonucleic Acid
d_p	Diameter of the Pore
dsDNA	Double-stranded DNA
e	Elementary Charge
EDL	Electrical Double Layer
EDTA	Ethylenediaminetetraacetic Acid
ELISA	Enzyme-linked Immunosorbent Assays
F	Faraday's Constant
G	Conductance
g	Conductance of the Solvent
G1	First Generation Dendrimer
G2	Second Generation Dendrimer
Gdm-HCl	Guanidium-chloride
H_2O	Water
H_3O^+	Hydronium
HCl	Hydrochloric Acid
I	Current
I-V	Current Voltage Curve
ICR	Ion Current Rectification
IDT	Integrated DNA Technologies
IgG	Immunoglobulin G
I_p	Peak Amplitude
JAK2	Janus Kinase 2
K	Potassium
K^+	Potassium Ions
Kbp	Kilo Base Pairs
KCl	Potassium Chloride
KHz	KiloHertzs
KOH	Sodium Hydroxide
l	Capillary Length
l	Length of the Conical Area of the Nanopipette

L_p	Length of the Nanopore
$MgCl_2$	Magnesium Chloride
μL	Microlitre
mM	Millimolar
MspA	Myobacterium Smegmatis Porin A
mV	Millivolts
Na	Sodium
NaCl	Sodium Chloride
NEB	New England Biolabs
nM	Nanomolar
nm	Nanometre
OmpG	Outer Membrane Protein G
pA	PicoAmperes
PBS	Phosphate Buffered Saline
PC	Plasmid Control
PC2ELigBO	Nicked, Ligated, and Cut Plasmid with Biotin Oligo
PC2ELigBO2	Nicked, Ligated, and Cut Plasmid with Two Biotin Oligo
PC2ELigBO2fil	Nicked, Ligated, Cut, and Filtered Plasmid with Two Biotin Oligo
PC2ELigBOfil	Nicked, Ligated, Cut, and Filtered Plasmid with Biotin Oligo
PC2ELigBOph	Nicked, Ligated, and Cut Plasmid with Biotin Oligo with Phosphorylation
PC2ELigBOphfil	Nicked, Ligated, Cut, and Filtered Plasmid with Biotin Oligo with Phosphorylation
PCNbBV	Nicked Plasmid
PCNbBVBO	Nicked Plasmid with Biotin Oligo
PCNbBVBO2	Nicked Plasmid with Two Biotin Oligo
PCNbBVBOph	Nicked Plasmid with Biotin Oligo with Phosphorylation
PCNbBVLigBO	Nicked and Ligated Plasmid with Biotin Oligo
PCNbBVLigBO2	Nicked and Ligated Plasmid with Two Biotin Oligo
PCNbBVLigBOph	Nicked and Ligated Plasmid with Biotin Oligo with Phosphorylation
PCSAL1	Cut Plasmid
PCSAL1NbBVfil	Cut, Nicked, and Filtered plasmid
PEG	Polyethylene Glycol

pI	Isoelectric Point
pM	Picomolar
Q_{area}	Integrated Area of the Event (Charge)
R	Resistance
SEM	Scanning Electron Microscope
smFRET	Single-molecule Fluorescence Resonance Energy Transfer
SiM-Pull	Single-molecule Pull-down
SiN_x	Silicon Nitride
SMT	Single-molecule Tracking
SNP	Single Nucleotide Polymorphism
SNR	Signal to Noise Ratio
ssDNA	Single-stranded DNA
S_{surf}	Silicon Surface
TBE	Tris-borate-EDTA (100 mM Tris, 90 mM Boric Acid, and 1 mM EDTA)
t_d	Dwell Time
TE	Tris Ethylenediaminetetraacetic Acid
TEM	Transmission Electron Microscopy
Tm	Melting Temperature
V	Volts
V_{bias}	Voltage Bias

Chapter 1 Introduction

Table of Contents

Synopsis: This chapter discusses the need for biosensing and single-molecule sensing specifically. It includes an exploration of the history of nanopores, their working principles, and the materials from which they are fabricated. The applications of nanopores and their challenges are also discussed. Some potential solutions for these challenges, in the form of nanopore modification and DNA carriers are examined. Lastly, this chapter contains a summary of the overall aims for the project and an outline for this thesis.

BIOSENSING AND SINGLE-MOLECULE SENSING

Over the last few decades there has been an explosion in the understanding of the chemical, biological, and nanotechnology worlds. This has allowed for an increased interest in biosensors, that is sensors that have a key biological sensing element. Biosensors have many uses and applications, which include disease monitoring in vivo, environmental, defence, and marine applications, food safety and processing, and drug discovery. Biosensors can process signals from a biological response and are able to target molecules or macromolecules. Key sensing techniques such as enzyme-linked immunosorbent assays (ELISA) and polymerase chain reactions (PCR) are extremely useful and widely available. However, these techniques do not allow for sensing at extremely low concentrations and only allow for bulk characterisations. Additionally, they require significant sample preparation and high sample volumes. [1–5] Thus, techniques that can perform single molecule sensing can offer many new insights that could be obscured when individual molecules cannot be analysed.

Recently research into single-molecule techniques has increased greatly. Single-molecule technologies offer the unique ability to image, label, and manipulate biomolecules individually. This offers previously unthought of access into biological systems. Some of these techniques include single-molecule fluorescence resonance energy transfer (smFRET), a single molecule imaging technique that can investigate conformational changes and molecular interactions, single-molecule pull-down (SiM-Pull), which can analyse the structure and function of protein complexes, and single-molecule tracking (SMT), which can reveal the cellular structures and dynamics of live cells.[6] Other techniques include the use of optical tweezers and atomic force microscopy (AFM). Optical tweezers are a tool which can trap particles in a laser beam. In single-molecule experiments molecules can be attached and immobilised to a bead which is trapped, allowing for the measurement of physical properties. AFM uses a cantilever to scan a surface or it can be attached to a molecule in

pulling experiments. The AFM detects motion through the observation of laser deflection off the cantilever. Despite the development of these techniques, there are still limitations. Many of these single molecule techniques require significant sample processing. Some of these can include modification to bind with dyes, immobilisation on a flat surface, or engineering the laser for a specific molecule.[7]

Nanopore sensing can offer the sensitivity of other single-molecule techniques without some of the limitations inherent to them. A nanopore is a small hole with a diameter at the nanometre scale. Nanopores can analyse single molecules at extremely low concentrations by detecting current changes as individual molecules pass through the pore. This technique is also label free and does not require any kind of immobilisation. Nanopore techniques have been used to study both proteins and DNA successfully. They can also identify and quantify concentrations of various biomolecules without much modification.[8–10]

NANOPORE SENSING

1.1 BACKGROUND

Nanopore sensing builds upon the principles behind the Coulter Counter. The Coulter counter, developed in the 1940s, uses pores to count and identify micro-sized objects, such as blood cells or bacteria.[11] Smaller pores, sub-micro, were used in the 1970s to detect viruses and other nanoscale particles.[12] However, the principles in the Coulter Counter were first used in nanopore sensing to study DNA and RNA in 1996.[13] A pore made out of a bacterial protein, α-hemolysin, was used after establishing the viability of such a pore; namely that spontaneous gating would not occur, the effect of polyethylene glycol on pore conductance, and that the pore diameter was ~1.1 nm.[14] The pore diameter was so small that only single-stranded DNA, and not double-stranded DNA was able to move through the pore.[13] Further experiments were performed to show that nanopore sensing could be used to sequence nucleic acids. These included establishing that information about the oligomeric structure, such as purine versus pyrimidine, could be determined from nanopore sensing.[15] This showed that two strands of identical length but different sequences had different blockade signals,[16] and demonstrated that there are differences in current traces for sequences that enter 5’-3’ and 3’-5’.[17–19] This allowed a single base substitution to be identified,[20] and finally confirmed

that all four DNA bases have distinguishable signals.[21] However, all of these experiments depended on immobilising the DNA strand in the pore, rather than allowing it to move through freely, as this allowed for a great increase in the signal-to-noise ratio.[14] This issue was eventually overcome by using DNA polymerases and blocking oligomers, allowing for good control of the translocation rate through the pore.[22–26] In addition to the initial α-hemolysin pore, solid-state nanopores were produced and used to detect DNA from 2001.[14] Solid-state nanopores allowed for different kinds of pore modification and size ranges. That, these resistive pulse sensing principles can be applied to a range of sizes and molecules makes them an extremely powerful sensing technique.[27]

1.2 WORKING PRINCIPLES

Nanopore sensing is a technique where molecules are transported across a membrane or pipette tip by an applied voltage. The pores are either biological or solid-state, of which nanopipettes are a subcategory. There are two compartments, cis and trans, containing electrolyte solution separated by a membrane that has one or several holes at the nanoscale, nanopores.[8] Two non-polarised electrodes are placed on either side of the membrane. These provide the driving force, and when chloride ions are used as the anion, these electrodes are typically Ag/AgCl.[27] The current in these pores is monitored continuously, as the passage of particles modulates the pore conductance. It is the flux of ions that determines the electric current. DNA and other biological molecules can be driven across the pore by the constant applied voltage and diffusion. As they move through the pore, the current changes. This temporary current modulation is called a translocation event, and these events can either block or enhance the current. A sample nanopore set up using potassium chloride as the salt solution is shown in **Figure 1.1**. The ions are driven across the pore by the applied voltage.[28]

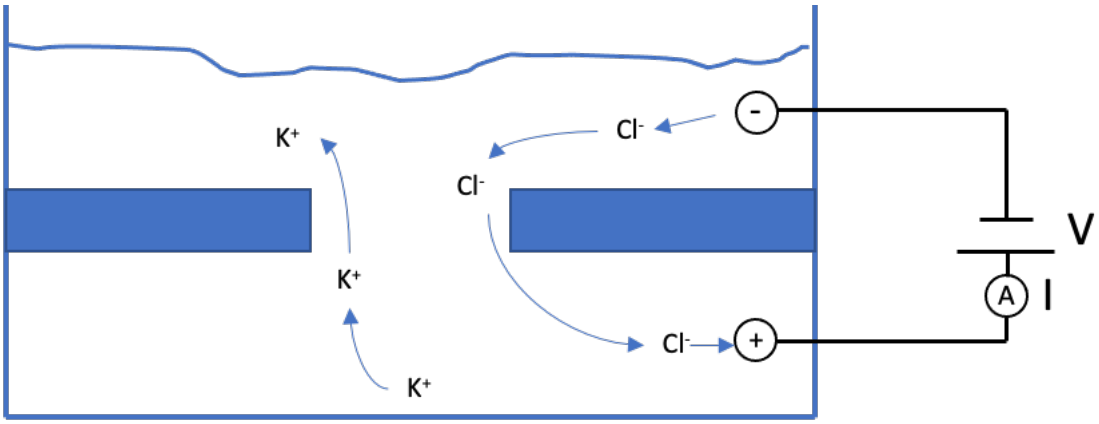


***Figure 1.1**: Sample nanopore setup using KCl as the salt solution. The applied voltage triggers ion movement across the pore. It is this electrophoretic force that is able to move both the ions and molecules of interest across the pore. The current is monitored continuously, such that modulations from molecules of interest are observed. Adapted from Reference 27.*[27]

Ag/AgCl wire is typically used as the electrode. As the free ions in solution that provide the open pore current are typically K^+ and Cl^-, and these are the ions used in this research, silver chloride is the ideal electrode. Silver chloride electrodes are non-polarisable and have fast charge transfer from electrode to solution. K^+ and Cl^- are typically used as they have similar ion mobilities and therefore move at similar speeds across the pore creating a stable current. It is simple to construct, easy and inexpensive to manufacture, has a stable potential, and has non-toxic components. The current is formed by the redox reactions with these electrodes, which are shown in **Equation 1.1** and **1.2**.[27]

$$AgCl_{(s)} + e^- \rightarrow Ag_{(s)} + Cl^- \quad \text{Equation 1.1}$$

$$Ag_{(s)} + Cl^- \rightarrow AgCl_{(s)} + e^- \quad \text{Equation 1.2}$$

***Equation 1.1 and 1.2**: The redox reactions that occur in the electrodes used in the nanopore experiments. Silver wire is used for the electrodes. Equation 1.1 occurs at the cathode(-) and Equation 1.2 occurs at the anode(+).*[27]

The reduction in **Equation 1.1** occurs at the cathode and the oxidation in **Equation 1.2** occurs at the anode. The electrons move along the circuit from anode to cathode and the ions move across the membrane. The Cl^- ions move from the cathode to anode through the electrolyte solution. Typically, the Cl^- ions move towards the reference electrode and the K^+ ions to the working electrode. The working electrode is defined as the electrode where the reaction of interest is occurring, while the reference electrode is a stable and well-

known electrode potential. It is the capture of Cl^- from solution that causes the electron migration through the wire, forming the current.[27,29]

The ions and the molecules of interest are driven across the nanopore using electrophoretic force via an applied voltage. The electric field around the nanopore is closely related to its conductance.[29] However, quartz nanopipettes are used in this research and these have a slight negative charge on their surface.[30] Therefore, surface charges can have an important effect, so it is important to consider this when discussing pore conductance. Thus, the open pore current is the combination of the bulk ions and the ions shielding the charged walls of the nanopore.[31] This is shown in **Equation 1.3**, which shows how the conductance of a cylindrical pore can be calculated. From this equation, it is possible to see that when the diameter is larger, the bulk dominates the conductance, but when the diameter is smaller the surface charges have a larger effect.[32] The cylindrical shape is a good model for nanopores fabricated from glass capillaries.[8]

$$G_0 = \frac{\pi {d_p}^2}{4L_p}(\mu_{K^+} + \mu_{Cl^-})Fc_0 + \frac{\pi d_p}{L_p}(\mu_{K^+})\sigma_p$$

***Equation 1.3**: μ_{K+} and μ_{Cl-} are the mobility of the cations and anions, c_0 is the bulk ion concentration in the nanopore, d_p is the diameter of the pore, L_p is the length of the nanopore, σ_p is the surface charge density of the nanopore, and F is Faraday's constant.*[32]

It is also necessary to consider the resistance in the nanopore setup. The total resistance is made up of the resistance of the pore and the resistance of the solution. However, for cylindrical pores, the resistance of the pore dominates the total resistance, as shown in **Equation 1.4**. As the pore resistance can be used as the total resistance, it is possible to consider that the applied bias, V, is equal to the potential drop across the pore.[8,33–36]

$$R_{solncis} + R_{pore} + R_{solntrans} \approx R_{pore}$$

***Equation 1.4**: Total resistance in the nanopore set up. For cylindrical nanopores, the resistance of the pore dominates the resistance of the solution.*

The movement of molecules of interest, rather than just the ions, has a different effect on the ions. As they move through, the current is temporarily modified with a translocation event. These translocation events can either be current enhancing or blockading.[28] The translocation event has several important characteristics; the dwell time, the length of the event, and the peak amplitude, the height of the event are the characteristics that are

discussed the most in this research.[37,38] Other typical attributes that can be calculated from these events are the capture rate, or the number of events in a certain amount of time, and the charge, which can be calculated using the area of the event.[39] The capture rate depends on factors such as the polyelectrolyte strength and concentration, nanopore geometry and surface chemistry, applied potential, and pH.[8] These events are characteristic of the molecules as they depend on attributes such as volume, charge, and conformation of the molecule.[35,40,41] Unfortunately, as solutions become more complex, that is, more similar to real biological samples, information on these attributes cannot be inferred from translocation events alone.[40] A schematic of a molecule moving through the pore and its corresponding translocation event is shown in **Figure 1.2**.

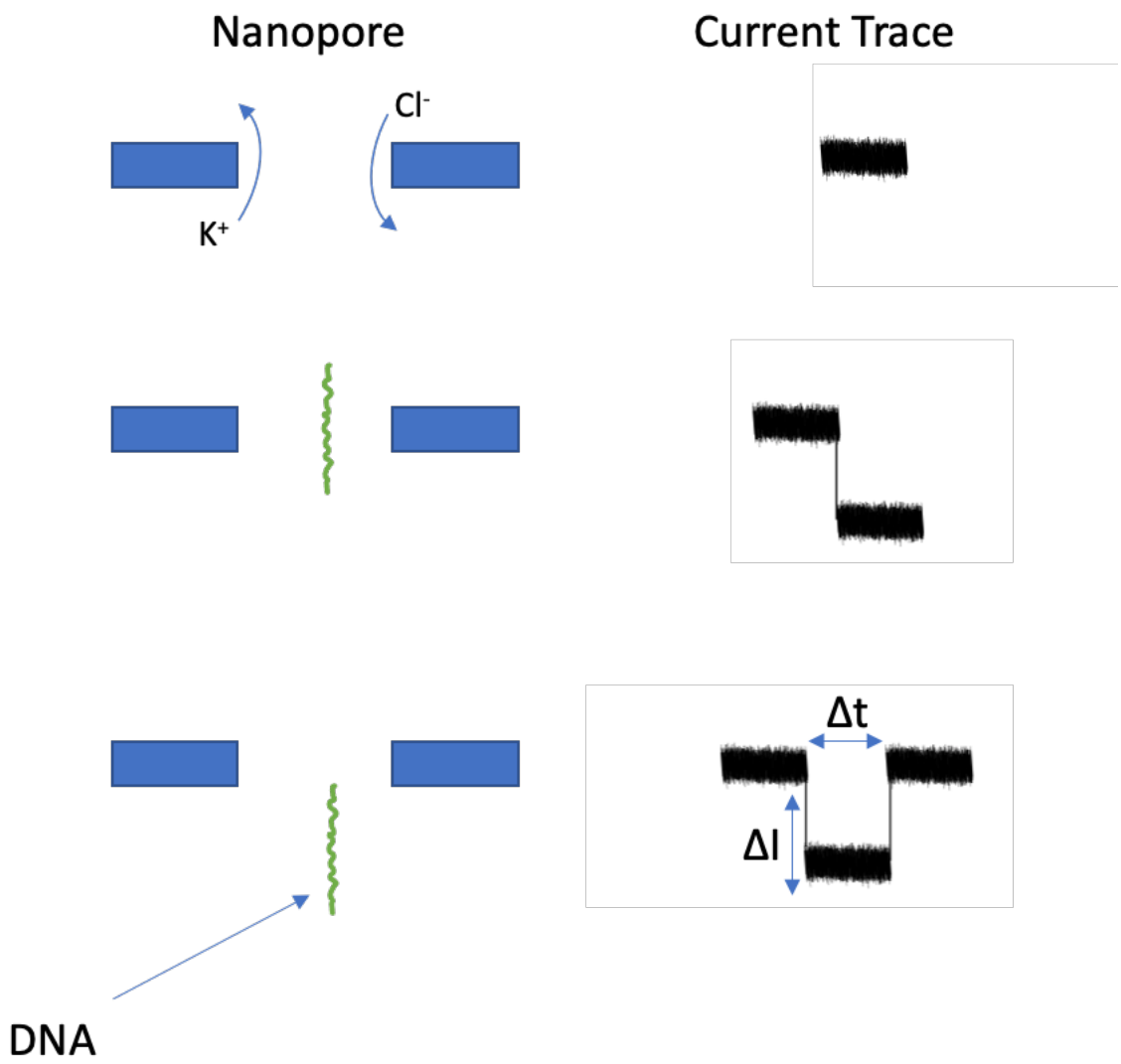


***Figure 1.2**: Schematic showing how the current trace changes as molecules move through the pore. The initial entry of the molecule causes the current to change, and the magnitude of this change is called the peak amplitude, ⊿I. The duration of this change is the dwell time, ⊿t. After the molecule moves through the pore, the current returns to the initial open pore current value.*

Some other factors also have effects on the movement of molecules through the pore. The first of these is Brownian motion. Brownian motion is always present and comes from the random motion of particles. Particles continually collide with each other randomly in the fluid, especially as some are moving at vastly different speeds. These collisions can cause some errors, as they can force a particle to move backwards through the pore.[8,42]

Another important aspect is the effect of surface charges. The surface charges stem from the material used to make the nanopore. In this research, quartz is used to form the nanopipettes, and this, along with the use of a polar solution, KCl, leads to the surface having

a slight negative charge. The negative charge derives from the dissociation of the terminal silanol groups on the surface of the quartz in the presence of aqueous solutions, which is shown by **Equation 1.5**. **Equation 1.5** only shows the dissociation reaction, but there is a corresponding protonation reaction as well. S_{surf} are the silicon surface atoms.[43]

$$S_{surf} - OH + H_2O \rightarrow S_{surf} - O^- + H_3O^+$$

Equation 1.5*: Dissociation equation of the silicon surface atoms in the presence of polar solutions.*[43]

The surface charge effect depends on several factors, including the surface itself, the solvent pH, the solvent ionic strength, and the temperature. This surface charge can be described using the double-layer model, the Gouy-Chapman-Stern model. A schematic of this model is shown in **Figure 1.3**. This model describes two different layers, an inner layer, called the Helmholtz/Stern layer, and the outer double-diffuse layer. In the Helmholtz layer, ions are immobilised, while in the outer diffuse layer they are able to move. Each ion described in this model is assumed to be a hard sphere with the same radius and the charge is located at the centre of each spherical ion.[44–46]

As there is a charged surface and subsequent migration of ions to the surface, a concentration gradient is created. While on a larger scale this effect is normally minimal, for nanopores these charged surfaces can have a significant effect, especially as the distribution of ions inside the pore can vary greatly from the bulk solution.[8] The effect of the surface charge is even more pronounced for very small pores, where the double layer can be affected by both the pore size and the curvature of the pore. Thus, it is possible that this effect can alter translocation events. This is especially true for lower salt conditions where the double layers are thicker.[30] With a lower concentration of ions, the bulk concentration is more influenced by the loss of ions that migrate towards the immobile Helmholtz layer.[47]

The length of this double layer can be measured using the Debye length, which is shown in **Equation 1.6**, where ε_r is the permittivity of the medium. The other symbols are standard.[45]

$$\delta = \sqrt{\frac{RT\varepsilon_r}{2F^2c}}$$

Equation 1.6*: The Debye length, which corresponds to the length of the double layer.* ε_r *is the permittivity of the medium, R is the gas constant, T is the temperature, F is Faraday's constant, and c is the concentration of the monovalent electrolyte.*[45]

The important takeaway from this equation is that the Debye length does not depend on the interface, but rather only on the properties of the liquid, particularly the permittivity and its concentration, as well as the temperature at which the experiments are conducted.[45] Also for pores where $d_p > \delta$, the double layer does not have a great effect, but for smaller pores, the effect can be meaningful. The more dilute an electrolyte solution is the longer the Debye length will be. At the conditions used in the experiments discussed in this thesis, 1 M KCl and 100 mM KCl the Debye length is approximately 0.3 nm and 1 nm respectively.[8,47]

While this model is useful to describe the ion behaviour at a charged surface within the nanopore, there are some limitations. The assumptions that the model makes about the ions' shape and charge, uniform charge along the surface, and the constant solvent characteristics, lead to errors.[46]

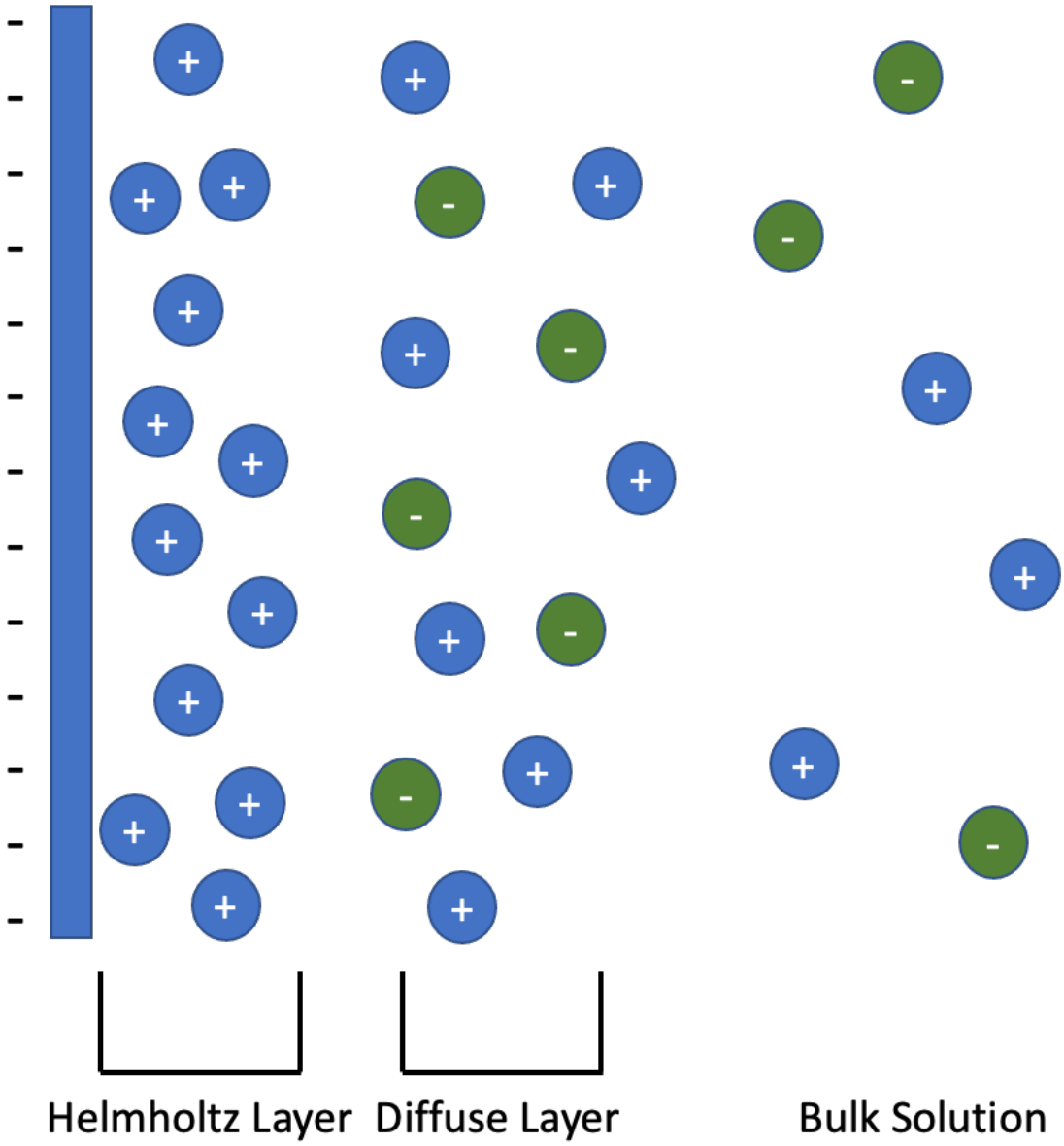


Figure 1.3: *Schematic of the double layer model. The initial inner layer, the Helmholtz layer contains immobilised ions. The out diffuse layer allows movement of ions, but does not allow the same free flow of ions as the bulk. Adapted from Reference 44.*[44]

In addition to the double layer, the charged surface of the pore can also lead to electroosmosis. Electroosmosis is fluid motion that occurs near a charged surface due to an external electric field.[30,48] In the experiments discussed in this research, this surface charge is negative[48] so is balanced by excess cations. The cations and anions then pull their respective solvation shells along once an electrical charge is applied, as they move along their potential gradients.[8] As there are unequal amounts of free ions, there is a net flow of liquid along the direction of the electric field. The cations move to counterbalance to the surface charge land

this leads to them being in excess relative to the bulk concentration.[49] This force is important in nanopore sensing, as electroosmosis can act in the same or in the opposite direction of the applied electric force.[50,51] The magnitude of the electroosmotic force depends on the pore material, analyte, and the solvent, and it is more prominent in long nanochannels. Typically, it will be less important for analytes that have a large fixed charge, but it does have more of an effect at lower salt concentrations.[8,30,48–52]

Another force to consider is the non-linear dielectrophoretic force (DEP) as there is a gradient within the pipette. This force is typically concentrated at the tip due to the shape of the pipette.[53] DEP can trap molecules at the tip of the pipette and this can be used to increase the detection level in nanopores.[54] This force is created from the change in electric field along the pipette's taper, and it is this variation in the electric field that leads to a gradient within the pipette. The DEP, depending on the direction of the applied voltage, can pull molecules towards the tip or it can oppose the electrophoretic force and trap molecules at the tip. This means that the forces affecting molecules can vary as they travel through the pipette.[55]

These surface charges are also thought to have another effect, ion current rectification (ICR). From Ohm's law, it would be expected that the current voltage curves (I-V curves) would be linear. However, in the presence of surface charges, the I-V curves become asymmetrical. Sample I-V curves illustrating this are shown in **Figure 1.4**. Typically the effect is much more pronounced at lower salt concentrations, and this is shown in **Figure 1.4**.[56]

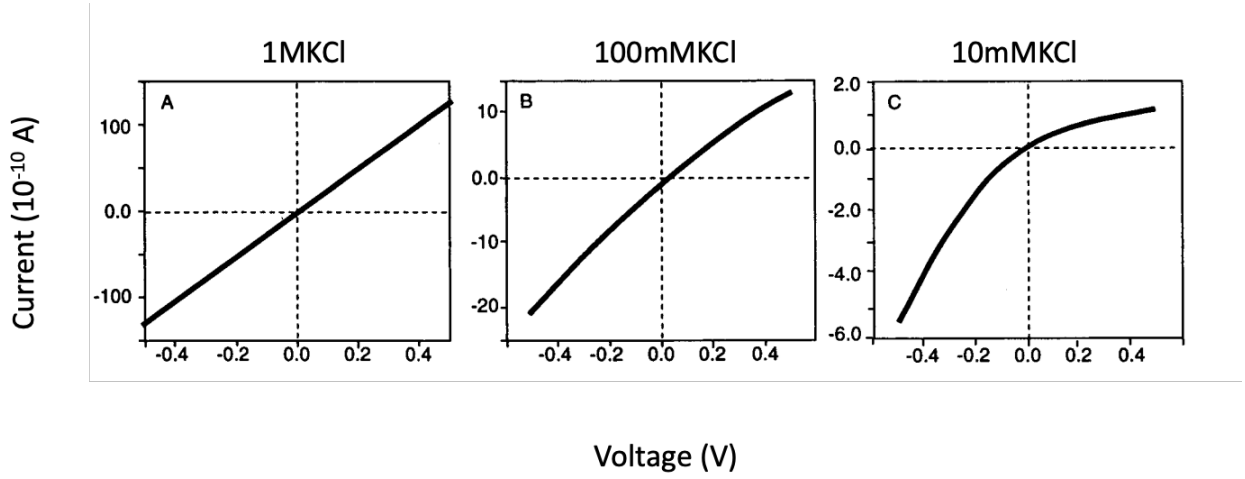


***Figure 1.4**: Sample I-V curves showing rectification. These are performed with a nanopipette at various different salt concentrations. The higher the salt concentration the less the rectification, as shown. Adapted from Reference 56.* [56]

This phenomenon seems to occur as on one side of the pore the current is repressed, while on the other it is enhanced. Typically, this occurs if the double layer is similar to the diameter of the nanopore, so the effect will be more pronounced in smaller pores.[30] For an uncharged pore, ICR would not occur, and thus the I-V curve would be symmetric. ICR is thought to be a result of uneven translocations of cations and anions, making the current greater on one side of the pore than the other.[57] For example, in low salt concentrations and an applied negative bias, where the diameter of the pore/nanopipette is similar to the Debye length, anions migrate from the tip to the middle of the pore, as they are rejected at the tip due to the negative charges on the surface. The cations can move freely through the pore, resulting in an increased ion concentration at the tip so that the conductivity of the solution at the tip is also increased. In the case of applying a positive bias, the opposite transpires and a subsequence decrease in conductivity occurs.[48,58] Surface charges are not the only property that can lead to ICR; asymmetric geometry and a blockage of the pore can also lead to rectification. The amount of ICR is dependent on properties like the surface charge density and the polarity of the applied potential, and as such, can be affected by the electrolyte concentration, the solvent pH, and surface modification.[57–60]

There are some other parameters, in addition to the electrophoretic force, that are important to consider during nanopore experiments. These in particular are the current amplification, the signal bandwidth and filter choice, and the noise considerations. For the first criteria, it is necessary to have a current amplifier for these experiments, as they are typically conducted in the picoAmpere range, and without amplification it would be impossible to get any meaningful data.[27] The exact range of amplifier can be tailored with the gain. For signal bandwidth and filter choices, it is important to consider the Nyquist theorem, so that no information is lost, as well as the signal to noise ratio (SNR). For the Nyquist theorem, it is important to sample at least twice as fast as the highest frequency that needs to be measured.[37] The bandwidth is the difference between the minimum and maximum frequencies in a continuous band of frequencies. Low-pass Bessel filters are used and the cut-off frequency is determined by balancing the expected time intervals for the molecule of interest and acceptable noise levels. Higher filter frequencies often result in more noise, but the greater the filter frequency the lower the time resolution. Applying a filter to attenuate the high frequency parts of the current signal can significantly improve the SNR, however they do lower the resolution and can distort the shape of events with small dwell times.[63] The

filtration frequencies typically used in this research are 10 KHz, for a resolution of up to 0.1 ms, and 30 KHz, for a resolution of up to 0.03 ms.[27] This resolution of the bandwidth of the recording is a significant constraint.[64] Keeping the signal to noise ratio (SNR) high is also important, as noise levels can obscure events. The noise in nanopore experiments is typically measured with power spectra. They show several types of noise, including the lower frequency noise, thermal noise, shot noise, and flicker noise, as well as the high frequency noise, the dielectric noise. Thermal noise is from thermal fluctuation of the charged molecules within the solution and is independent of the current, shot noise is the caused by random changes in ion motion in the pore, and flicker noise, often larger than thermal and shot noise, comes from the application of a bias across the pore.[8,27,65] For the high frequency noise, the nanopore material has a huge effect on the noise levels, as it is particularly dependent on the material's dielectric constant. This noise is associated with the capacitance and dielectric loss of the pore.[65,66] Sample power spectra are shown in **Figure 1.5**. Corresponding noise levels in the current trace are also illustrated in this figure. Both power spectra have similar noise levels in the high frequency noise, but significant differences in the low frequency noise.

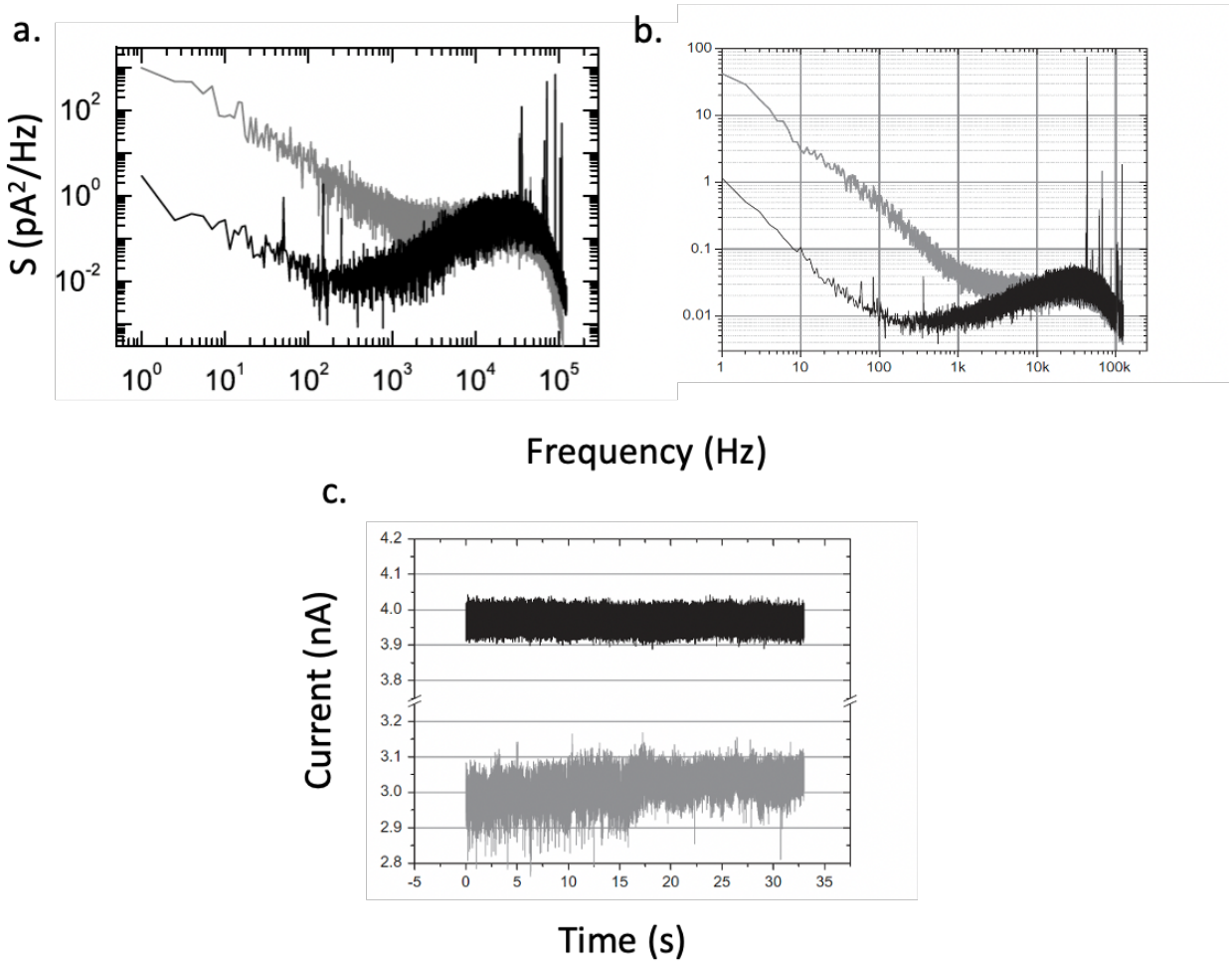


***Figure 1.5**: Sample power spectra are shown in a and b. The black is from a lower noise trace and the grey from a higher noise trace. Both have similar levels of noise at high frequencies, but the noise at lower frequencies is very different. Current traces for b are shown in c. In these current traces, it is easy to see that the grey has much higher noise levels. Adapted from References 8 and 67.*[8,67] *Copyright (1998) National Academy of Sciences.*

1.3 TYPES OF PORES

There are two types of nanopores, biological and solid-state. Nanopipettes are another commonly used pore, but they are a sub-category of solid-state pores. They can be made from organic or inorganic material.

1.3.1 BIOLOGICAL PORES

The use of biological nanopores started in the late 1980s with development in ion-channel physiology.[68] Bacterial proteins, such as α-hemolysin, are often used for biological

pores.[69] α-hemolysin is a toxin secreted by Staphyloccocus aureus as a 33.2 KDa water-soluble monomer. This toxin forms a mushroom shaped pore that has a height and diameter of 10 nm. α-hemolysin has a hydrophilic interior and hydrophobic exterior. There is a cap and stem, where the stem has a constriction of about 1.4 nm.[70] **Figure 1.6** shows a α-hemolysin pore.

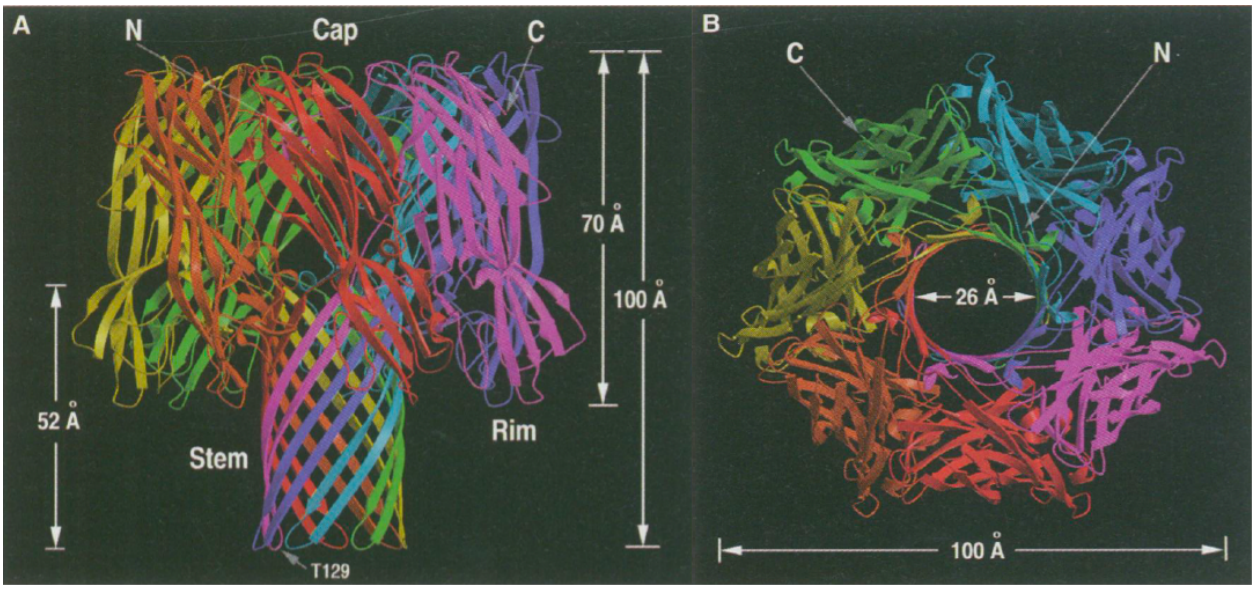


***Figure 1.6**: The mushroom shape of α-hemolysin. It is possible to clearly see the stem which forms the narrow opening that is used as the pore. Adapted from Reference 70.[70] From Song, L. et al. Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. Science 274, 1859–1866 (1996). Reprinted with permission from AAAS.*

α-hemolysin also is remarkably stable for a biological pore. It can remain functional in concentrations of up to 2 M guanidium-chloride (Gdm-HCl) and temperatures of 95°C.[71,72] It even remains stable under high concentrations of urea, such that α-hemolysin can be used to study denatured proteins.[73] α-hemolysin remains the most widely used biological pore, as it has repeatable self-assembly and an appropriately sized inner diameter, even if other biological pores are also used. Furthermore, compared to some other biological pores it remains stable over a wider variety of experimental conditions. [10,14,60,74–80].

Biological pores can also be short peptides, like gramicidin A, and these self-assemble to form channels, rather than large transmembrane proteins like α-hemolysin.[81] Other biological pores include myobacterium smegmatis porin A (MspA)[22,23,75], which is an octameric channel pore with a diameter of 1.2 nm, aerolysin (AeL)[80,81] , a heptameric pore with a bottleneck-shaped structure and a diameter of 1 nm, and outer membrane protein G

(OmpG)[68,80], which at its narrowest has a diameter of 1.3 nm and is a monomeric β-barrel protein. All of these protein nanopores have been used in a wide variety of applications.[66,69,73,75,78,79]

Biological pores are very customisable and can be tailored for specific molecules.[75] As such they are probably the single-molecule pore limit for both sensitivity and selectivity. However, this high level of selectivity and sensitivity comes with a lack of generalisability.[8] Biological pores have some other drawbacks. They are typically only usable for studying specific proteins or molecules as the pore size is not tuneable; this can be a limit to wider application. Furthermore, these biological pores have to be inserted into a lipid membrane, which can be tricky, and they are also susceptible to breaking down under the higher ion concentration conditions required by many nanopore experiments in order to have a reasonable SNR.[75,80]

1.3.2 SOLID-STATE PORES

Solid-state nanopores are another type of nanopore, that can be designed to resolve some of the issues with biological pores. As they are fabricated, their size and shape can be customised, and they typically are more robust. They can survive much harsher temperatures, pH conditions, and ionic concentrations, and they can have adjustable surface properties.[75,80] Although, they are intensive to make, they are reusable unlike biological pores. Current methods allow for the creation of reliable nanopores, although some variation can occur. Solid-state nanopores are typically made from silicon nitride, aluminium oxide, and graphene, and are commonly made through ion-beam sculpting, electron-beam drilling, and ion track etching.[82–90] Solid-state nanopores were first fabricated by Li *et al.* in 2001. These pores were made with ion-beam sculpting. Using an argon (Ar^+) ion beam, they were able to sculpt a cavity in the Si_3N_4 membrane, as shown in **Figure 1.7a**.[82] Transmission electron microscopy (TEM) was then used to determine that the diameter of this pore was 1.8 nm. They were able to have control over both the temperature and the ion beam exposure, allowing for the fine tuning of pore size. Other noble gas ion beams have also been shown to sculpt nanopores effectively. Electron beam drilling with the use of a TEM is another method for making solid-state nanopores from a silicon membrane. The high electron intensity from the TEM was used to modify the

dimensions of the silicon oxide nanopore, in a controlled manner, such that they achieved nanometre scale dimensions. This is shown in **Figure 1.7b**.[83] Ion track etching can also be used to make solid-state nanopores. Track etching uses accelerated heavy ions, like Xe and Au, to irradiate a polymeric foil. This is then followed by chemical etching of the latent ion tracks. The number of pores in each membrane corresponds to the number of ions that went through the foil, so it is possible to get a membrane with a single pore of nanometre dimension. This method is shown in **Figure 1.7c**.[88]

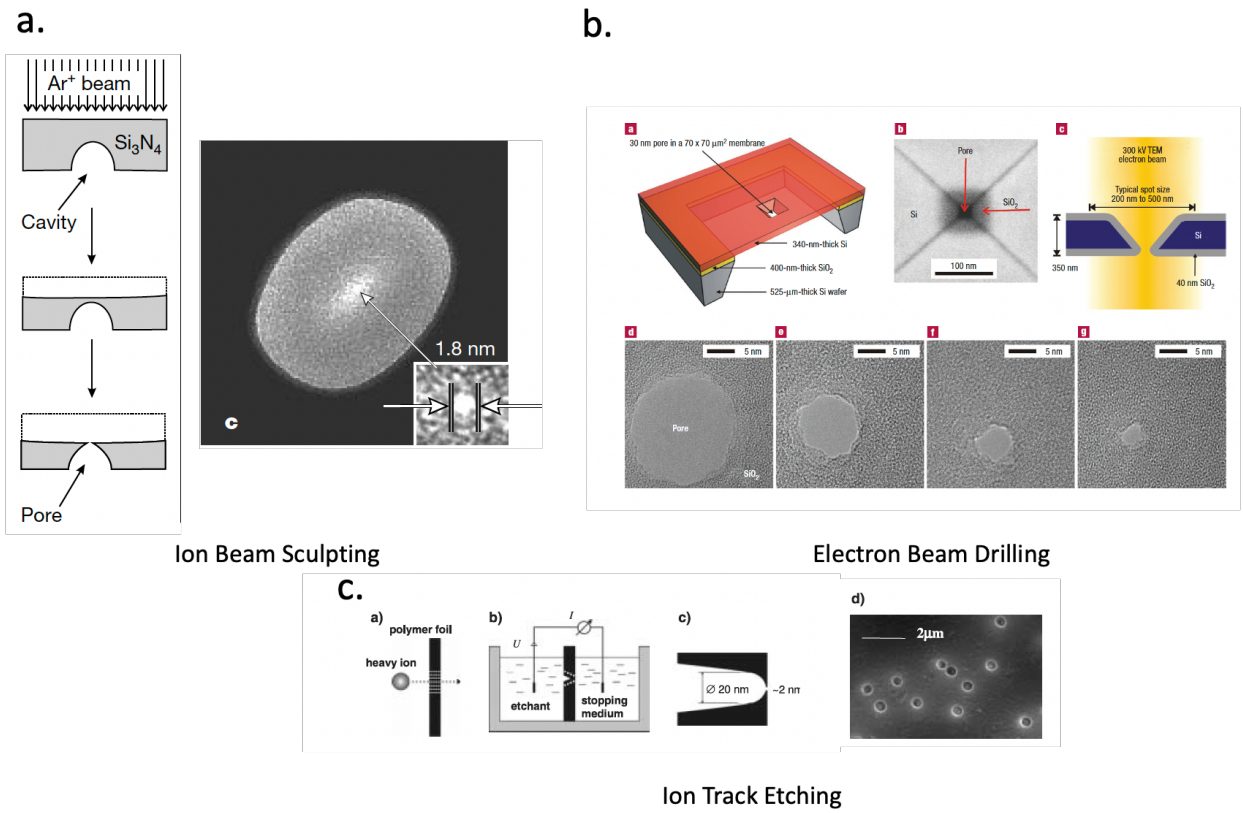


***Figure 1.7**: Three fabrication methods for solid-state nanopores. a. Ion beam sculpting uses an Ar^+ ion beam to form a pore through a silicon membrane. A TEM image of the pore is shown. b. Electron beam drilling uses TEM to shrink the size of a pore in a silicon membrane, and a sequence of images from the TEM shows this effect. c. In ion track etching, a heavy ion is used to form a track that is the base of forming the nanopore. An SEM image of the membrane is shown with the pores. Adapted from References 82, 83, and 90.*[82,83,90]

Solid-state nanopores have been used in a variety of different applications. These can vary from DNA sensing to protein sensing and DNA base detection. DNA sensing includes examinations of hairpins, unzipping of double stranded DNA, and investigations into folded states.[91–95] Protein sensing has been used to detect proteins such as bovine

serum albumin (BSA), DNA polymerase, and fibrinogen.[91] DNA base detection and sequencing has also been accomplished.[14]

While these solid-state nanopores can achieve good results, they are often time consuming to make, require optimisation, and can cost significantly more than biological pores. However, there is a class of solid-state pores that do not have these downsides: nanopipettes.

1.3.3 NANOPIPETTES

Nanopipettes, the solid-state nanopores used in this research, are very quick to make. They are glass capillaries made from either borosilicate or quartz using a laser puller. The glass capillary is clamped into position and then heat, via a carbon dioxide (CO_2) laser, is added to soften the centre of capillary. Directly after the application of this heat, both ends are pulled apart. This heat and pull allows a taper to form in the middle of the capillary. The taper is then heated and pulled again to form the two symmetric pipettes with a conical structure at the tip.[30,58,95,96] It is possible to tune the taper lengths and diameters by varying parameters in the puller, including heat and pulling velocity.[97] More information about the parameters of the puller can be found in Chapter 2, section 2.1. A schematic for the formation of these nanopipettes is shown in **Figure 1.8**. This whole process, including a plasma cleaning of the pipettes, takes under an hour; for each pipette the pulling time is under 10 seconds. Nanopipettes are also low cost, electrically and chemically stable, and have low noise.[97,98] They typically have pores ranging from 15 nm to several hundred nanometres.[97,99] For use in DNA sensing, the nanopipettes tend to vary from 14-45 nm.[58] These capillaries form into conical nanopores that have well-defined geometries. They are able to identify differences on an extremely small scale, such as determining coated and uncoated particles. The main downside of the use of these capillaries as nanopores is that they are not able to achieve diameters sub 10 nm.[58] Some experiments have been performed to further reduce the size of the nanopores, but this requires more time and equipment, and does not allow for the pipettes to be produced in batches, the main benefits of nanopipettes.[100,101]

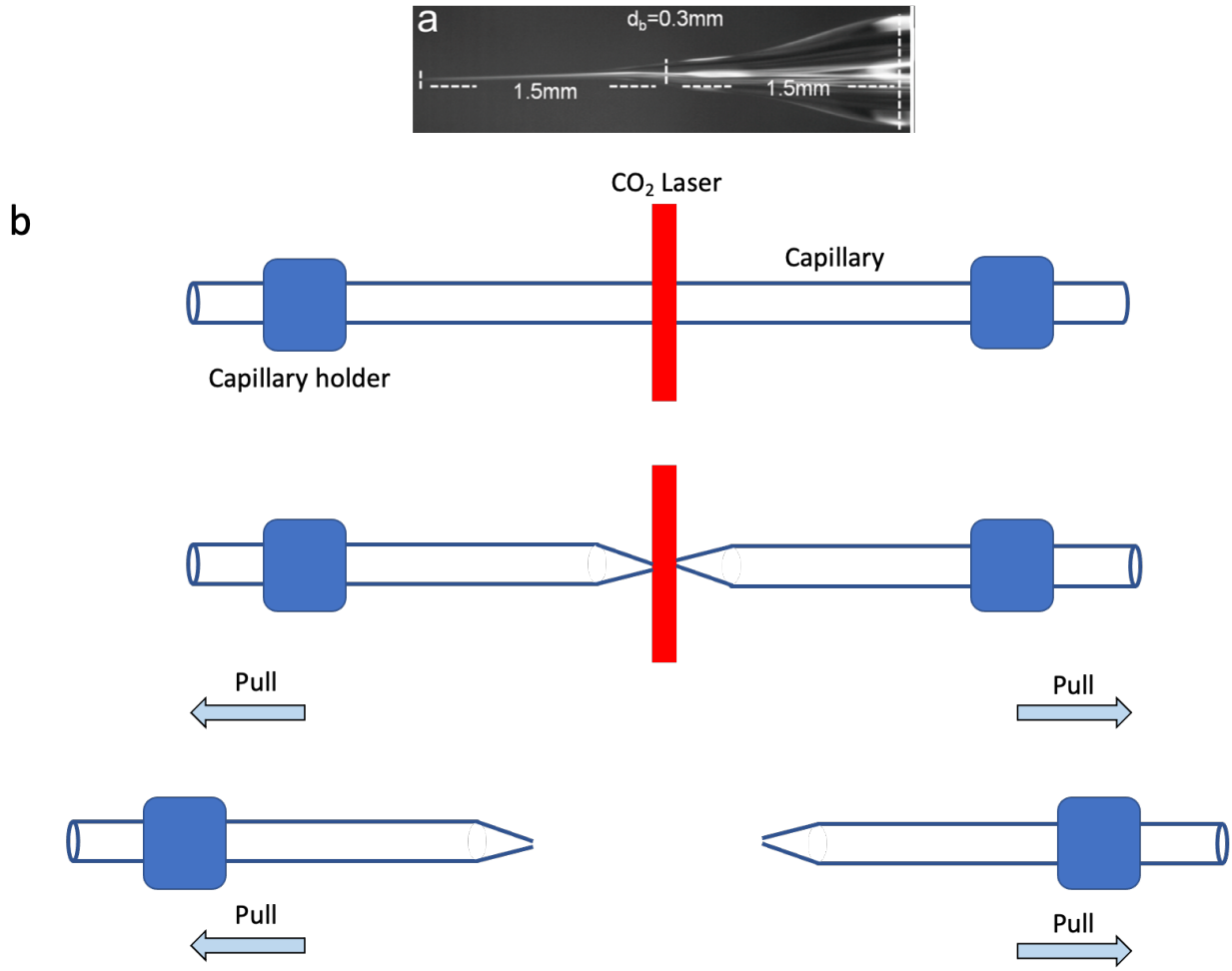


***Figure 1.8**: a. Micrograph of a nanopipette with several dimensions, including the diameter and tip length, labelled. b. Schematic of the laser pulling process for the formation of nanopipettes. The capillary is inserted into the holder and heated by the laser. It is then pulled apart and the tapered length in the centre is heated again. The capillary is then pulled apart forming symmetric pipettes. Adapted from References 97 and 98.*[97,98]

1.4 DNA TRANSLOCATIONS

1.4.1 WORKING PRINCIPLES

DNA is a molecule that has been well studied using nanopores. It is generally quite simple to identify a translocation event by monitoring the current. Furthermore, it is possible to do rapid sampling through the nanopore.[13,102] Typical DNA translocation can be described in three steps. At first, through diffusion and drift due to the applied voltage, the DNA moves towards the pore entrance. Then, it enters the pore. Lastly, the DNA moves through the pore, and while this occurs, it changes the electrical current.[41] For the DNA to enter to the nanopore a conformational change is required. It must unwind from a coil in

thermal equilibrium into an elongated state, which requires overcoming a free energy barrier. Thus, two free energy barriers have to be overcome for every translocation event: the unwinding of the chain and the entropic pressure that occurs at the pore entrance.[8,27] DNA translocation events occur in three types: nonfolded, partially folded, or fully-folded. Sample translocation events of these are shown in **Figure 1.10**. In nonfolded events, the DNA remains linear throughout the whole translocation event. In partially folded, the DNA enters partially pinched over, and remains like this, throughout the translocation. Folded is when the DNA enters and remains pinched completely in half.[88,103] The dwell time is dependent on both the length of the DNA and its configuration, so matching a specific translocation event to a specific DNA length is not possible.[41] However, analytes such as DNA, which are large relative to the size of the pore, have an event duration that can be related to the dimensions of the analyte.[102,104,105]

DNA is a particularly well-studied molecule in nanopores as it is has a large negative charge. This large negative charge allows it to be driven across the pore effectively with electrophoretic migration. Due to its negative charge, there is an increase in the cation concentration near the surface of the DNA, but the bulk solution remains neutral overall. However, with a decreased ion concentration in solution, behaviour deviates more as both the bulk and the surface charge effects impact the conductance. Furthermore, if the surface charge of the nanopore is larger than the DNA charge, the DNA cannot translocate at all. This can lead to the DNA molecule becoming stuck and translocating very slowly, and in this condition, it is not possible to distinguish between the electroosmotic and electrophoretic effects.[31,58,106] There is also a linear dependency between salt concentration and conductance change. However, high salt conditions affect the surface charge, where the salt concentration does not have a linear effect on ion mobility.[31] Additionally, for larger pores, there is a much larger distribution in dwell times than for smaller pores.[107] This effect likely stems from the reduced number of nanopore-DNA interactions for the larger pores. The capture rate for DNA tends to increase exponentially at first, until it becomes linear as the applied potential is increased. However, it can become lower at high voltages due to steric hinderance and DNA molecules crowding the pore entrance.[89,106–108]

Another important element to consider with DNA translocations is the direction of analyte movement. Depending on the salt concentration the DNA can be added either to the pipette, cis-to-trans direction, or to the bath, trans-to-cis direction. The choice of where

DNA is added is dependent on the salt concentration due to the electroosmotic forces. If DNA is added to the bath with a 100 mM KCl salt solution, the electroosmotic force is so strong that it will not translocate. For higher salt solutions, it becomes possible to add the DNA to the bath, as the electroosmotic force decreases significantly at higher salt concentrations. For cis-to-trans, the DNA is forced through the pore, by applying a negative voltage, and vice versa for trans-to-cis.[31] Typically the translocation event is shorter going from trans-to-cis than from cis-to-trans, and the velocity of the DNA remains constant in this direction. As the DNA moves out of the bath and into the pipette, the DNA buckles as it is under a compressive load, and this results in a smaller resistive force. For cis-to-trans, the DNA translocates as a straight rod and is subjected to a tension force rather than a compressive force.[111] A schematic of this is shown in **Figure 1.9**.

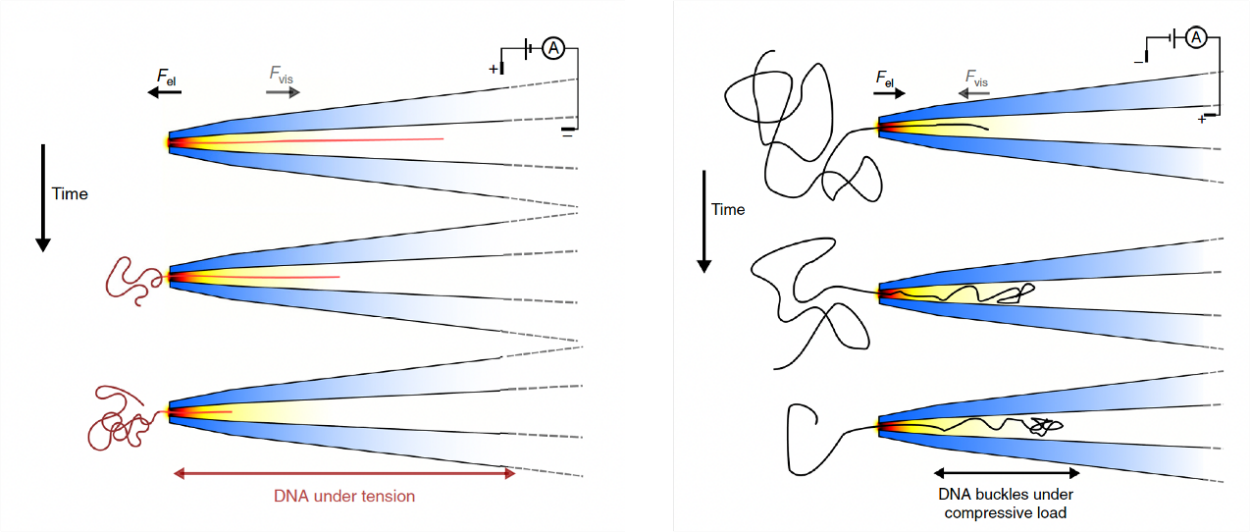


***Figure 1.9**: Schematic showing the difference for DNA translocation for cis-to-trans and trans-to-cis. In cis-to-trans the DNA is subjected to a tension force inside the pipette and a significant portion of it is affected by the viscous drag force. For trans-to-cis, most of the DNA is compressed and buckles. Only a short section of DNA, closest to the pore entrance, is affected by viscous drag. This could account for why the dwell times are normally shorter for trans-to-cis events. Adapted from Reference 111.*[111]

As the magnitude of force required for either direction varies, there is a much higher capture rate in the cis-to-trans direction than the trans-to-cis direction. For DNA, asymmetric pore geometry can lead to a large variety of distinct translocation signals, due to alterations of the salt concentration or the translocation direction. This is especially clear in 1 M KCl salt solutions, as the translocations look very different in the cis-to-trans direction compared to the trans-to-cis direction. The magnitude of the current change can also be

different even if the same magnitude of voltage was used. Typically, the current change is much greater for trans-to-cis. Another effect of asymmetric pores is that concentration polarisation can occur. This is when one ion transfers more readily, and as such, there are local depletions or enrichments of certain ions; coions are depleted and counterions are enriched. Concentration polarisation can have a significant effect on the transfer of DNA.[109–111]

As DNA has a negative charge, it can either enhance or block the current, depending on the salt condition in which the experiment is carried out. For low salt concentrations, such as 100 mM KCl, the DNA is negative enough that it can offset the more the negative ions it displaces as it travels through the pore, resulting in enhancement. However, when high salt concentrations are used, the DNA displaces too many negative ions. Thus its inherent negativity is not enough to offset the negative ions that are blocked as it moves through the pore. This leads to current blockades.[28]

1.4.2 APPLICATIONS

Nanopores have been used to study various different aspects of DNA. The detection of DNA folding has been performed, as well as investigations into DNA knots. For these two cases, the translocation events vary enough that it is possible to get DNA confirmation information from the nanopore. The folded and/or knotted DNA has an additional peak that corresponds to where the knot or fold is on the DNA. Additionally, as shown in **Figure 1.10**, circular DNA has a higher peak current, but also a shorter dwell time.[95,114] Other DNA structural information has also been studied, including comparisons of linear, circular relaxed, and supercoiled. It was again possible to separate these structures by their differences in dwell times and peak amplitudes.[115]

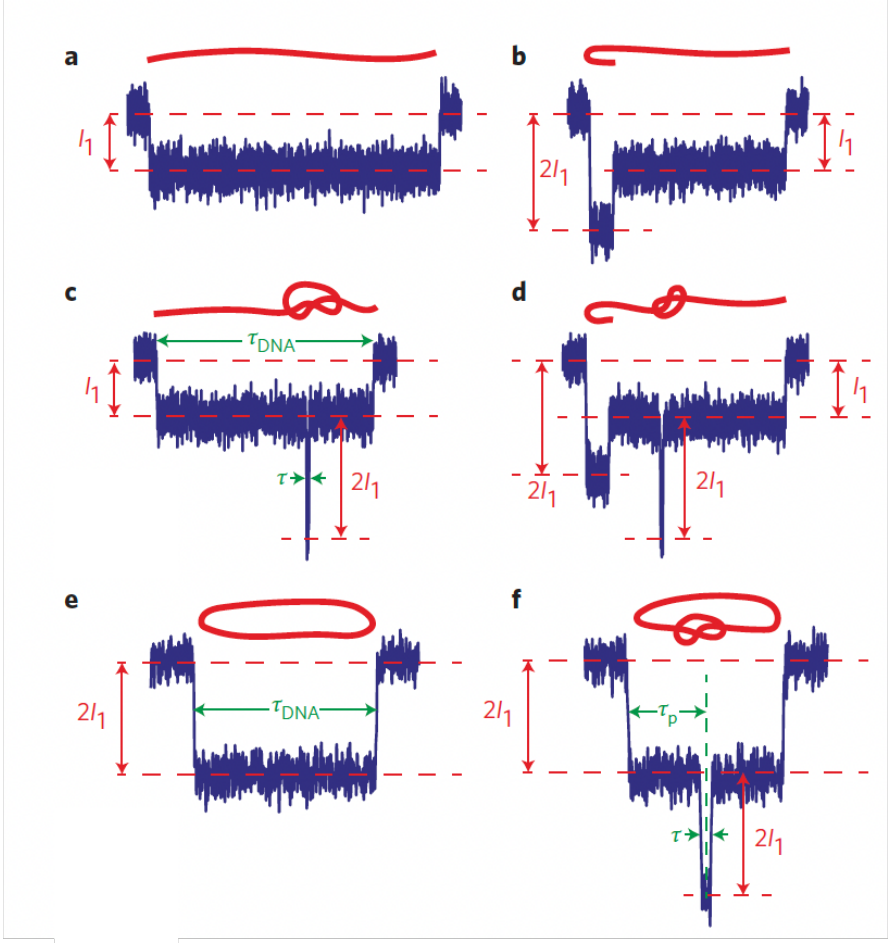


***Figure 1.10**: Sample DNA translocation events with knots and folds. The shape of these events differs enough that it is possible to get structural information from them. Adapted from Reference 114.*[114]

Nanopores have also been used to perform DNA sequencing. As previously described in this chapter, rapid, reliable, and continuous DNA sequencing was a long-term goal. In 2014, Oxford Nanopore Technologies released their MinION device and this was realised. The MinION is able to regularly sequence fragments of up to 48,000 bases.[14]

In addition to using nanopores to study DNA, DNA has been harnessed as a tool to use with nanopores. For example, DNA probes were designed to identify mercury ions. The binding of the mercury to the DNA probe changed the DNA structure such that the translocation events from the pore were significantly different as to identify probes with and without mercury. DNA was also able to be used as a probe to identify circulating microRNAs. There were similar changes to translocation events reported for this probe, so that DNA with and without microRNA could be identified. The proportion of microRNA reported from the translocation events matched the microRNA concentration.[116,117]

1.4.3 CHALLENGES AND SOLUTIONS

The main challenge with DNA is that it can move through the pore too quickly to either sequence it or to find out some information about its structure, so fast translocation speeds are an issue. The fast translocation rates lead to poor SNR and resolution.[118] To counter this problem many solutions have been posited. These include, modifying the molecule, the viscosity of the solution, the salt concentration of the electrolyte solution, and the temperature, in addition to adding a salt gradient, lowering the bias voltage, or modifying the nanopore itself. Modifying the molecule can work very effectively, but can also introduce other challenges, particularly more processing steps. The translocation times are directly proportional to the viscosity of the molecule, so increasing it can increase the dwell times. If the salt concentration is increased, DNA molecules can get saturated with counterions without any effect in the bulk solution. This leads to a slight increase in the dwell times. Changing the electrolyte solution, for example from KCl to LiCl, can greatly improve DNA translocation event resolution. This due to the smaller ion radii for Li^+ ions compared to K^+ ions. This allows for more bonding and thus the negative charge of the DNA is more neutralised. The lower negative charge reduces the effect of the electrophoretic force and therefore the velocity of the DNA. The temperature can affect several parameters including the conductivity, viscosity, and ion mobilities, such that the overall effect of decreasing temperature is a decrease in dwell time. For the applied voltage, translocation times are inversely proportional, so decreasing it will increase translocation times. There are several ways to modify the pore to slow down DNA translocation. These include adding a surface coating, having the DNA interact with a nanofiber mesh, shrinking the pore, or functionalising the pore to interact with the DNA more. [58,87,118–126]

1.5 PROTEIN TRANSLOCATIONS

1.5.1 WORKING PRINCIPLES

Proteins have also been studied in nanopores, but they are a more recent application of nanopore sensing. However, they have several different characteristics than DNA that affect nanopore experiments. Proteins do not have a universal charge, unlike DNA, and instead have a variety of charges. Their shape is also quite different, and they can have

several transitional states and substates.[40,125,127] Additionally, they tend to have dimensions that are similar to the pore, particularly if they are globular, rather than one dimension that is much longer than the pore. This means that their translocation is often very fast, particularly compared to DNA; the average dwell time is 100 μs. Thus, there are often many lost events in protein translocation, and the amount of events lost depends on the protein's size, net charge, and the minimum time resolution. Furthermore, for small proteins, this fast translocation speed can often result in poor SNR.[118] For samples with multiple proteins, in particular, there can be a lack of specificity. As the charges vary, it is also possible to have protein-nanopore wall interactions, and this can lead to difficulties in identifying different protein species. Particularly as nanopores cannot chemically differentiate, without modification, between analytes of similar sizes.[128] As proteins enter the nanopore, there are several forces that govern this event. These include the applied electrophoretic force, electroosmosis, and drift, stemming from free energy gradients. The residence time increases with peptide length.[61,116,125–128]

1.5.2 APPLICATIONS

Nanopores can be used to study the folding and misfolding of proteins. They also have been used to study different conformations of a protein within one sample. They have been used with alpha-synuclein, beta-amyloid, and prions, all of which are proteins that are of great interest medically.[132,133] Studies have analysed both mutant and wild-type proteins, and they were able to confirm percentages of folded and unfolded protein. The unfolded events had a significantly larger peak amplitude. Although these did have to be performed at low concentrations to prevent aggregation. It is very promising for studying protein folding at the single-molecule level.[134]

The study of proteins in the nanopore has many important applications, particularly when it comes studying protein confirmations at a single-molecule level. Sutherland *et al*. (2014) investigated proteins using nanopores and could see intermediated confirmations of collagen-like peptides.[129] Nanopores have also been used to study protein confirmation changes[135,136] and pH responsiveness.[131] Additionally, there is some work, although problems persist, in using nanopore sensing for protein sequencing.[137–140] Quantification of

proteins has also been performed using aptamer functionalised nanopores.[110] Nanopore sensing has also been used to examine post-translocation modification in proteins.[61,130–142]

1.5.3 CHALLENGES AND SOLUTIONS

The main challenge with protein translocations is the speed the protein moves through the pore. It is possible to slow it down by modifying the nanopore, for example coating the nanopore with polymer, and controlling the pore and protein surface charge. Other challenges are the non-linear shape, particularly globular, and their non-uniform charge. They can also absorb to pore walls or aggregate. These challenges all lead to extremely fast dwell times.[118,127] There are two options to solve this issue: either slowing down the molecules or improving the bandwidth and resolution. The bandwidth has been improved through the use of new amplifiers. In these experiments, thin solid-state nanopores were used to the study small proteins, <30 KDa, RNase and proteinase K, and dwell times as low as 2-2.5 μs were able to be resolved.[121] It is also possible to unfold the protein, using a polyethylene glycol (PEG) chain, although this means structural information about the protein is lost. Optical and AFM tweezers have also been used to improve protein translocation results.[127]

There are several ways to control the protein movement through the pore. This includes affecting the proteins charge by varying pH, controlling the electroosmotic flow, the modification of the surface charges on both the protein and the nanopore, varying the electrophoretic force, chemically modifying the pore, and coating the pore with fluids or polymers.[63,131,147–149] Another important technique is by binding the proteins to a DNA carrier, and this can particularly negate difficulties in getting the protein to move through the pore.

1.6 MODIFICATION OF THE NANOPORE

To negate some of the issues discussed in both the DNA and protein translocation sections, it is possible to modify the nanopore. Wei *et al*. (2012) demonstrated that it is possible to modify the aperture size of a solid-state SiN nanopore using DNA origami gatekeepers. This allowed for the production of solid state nanopores that have extremely small diameters, about 9 nm, on par with biological ones. **Figure 1.11a** shows a schematic of

what these pores look like. They demonstrated that they were able to get good results for certain proteins such as streptavidin and immunoglobulin G (IgG), with far better SNR.[150] Other groups have functionalised the nanopore with molecular-recognition elements such as aptamers or antibodies. Das *et al*. (2018) modified nanopipettes so that thrombin aptamers coated the surface. **Figure 1.11b** shows a schematic of these functionalised nanopores. With these functionalised aptamers, they were able to quantify thrombin down to 50 pM in real time.[110] However, all these methods require significant processing and are time consuming. There are also stability issues that stem from uneven coating. This is particularly important in nanopore sensing as this could affect the baseline and make it harder to identify binding events. Additionally, functionalising the pore leads to pores only being able to select for one protein. As such, there are significant downsides to modifying the pore directly.[37,98,149–152]

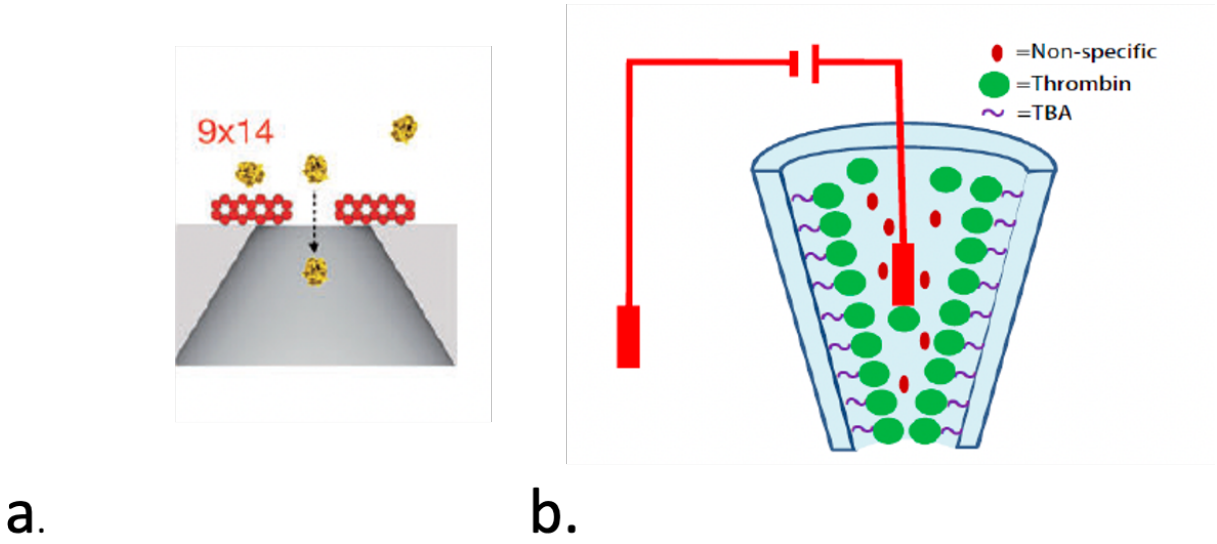


Figure 1.11*: Examples of functionalised nanopores. a. shows a solid state nanopore with a DNA origami gatekeeps. This allows for a reduction in pore aperture. b. shows a nanopipette functionalised with thrombin binding aptamers. The pore is able to select directly for thrombin and even quantify it in the picomolar range. Adapted from References 110 and 150.*[110,150]

1.7 DNA CARRIERS

1.7.1 WHY DNA CARRIERS

DNA strands that carry a ligand of interest through the nanopore are called DNA carriers. DNA is a suitable molecule for this as it is well studied in nanopore sensing and can

be easily detected. It is possible to design them to bind to any ligand, thus making it possible to select for specific proteins. They are particularly useful for globular proteins that would need to unfold to translocate or proteins that do not have enough negative charge to be driven across the pore.[148] Another benefit is that DNA carriers can significantly slow down the translocation event. This prevents the loss of protein translocation events, particularly for small proteins.[128] When these carriers translocate through the pore, the translocation events look different with and without bound protein. Sample events for a carrier with and without a subpeak are shown in **Figure 1.12**. These translocation events have both a primary peak and a secondary subpeak. The subpeak is much shorter and corresponds to the protein only. They allow for lower detection limits for the protein and ensure that the complex has an overall negative charge. The peaks make it possible to gain significant information about the protein, in addition to confirming it is present, therefore giving both identification and quantification. If the K_d is known, it is possible to quantify the protein.[128] Additionally, it is possible to bind more than one type of protein or more than one protein to a carrier, depending on how it is designed. If more than one protein is bound, the amplitude of the subpeak increases.[147] However, one downside of using DNA carriers is that the subpeaks can look similar to folded events. Therefore, using a shorter DNA strand can be advantageous as it prevents large numbers of folded events. Another disadvantage is that it can involve significant DNA modification or processing to make the sample. However, it generally does not require the same amount of time and equipment that modifying the nanopore itself does, and it is often possible to select for the modified DNA. Other problems that previous DNA carriers have had are the need for high salt concentrations, lots of excess protein, and extensive engineering. [40,128,147,148,155] This research will try to address these problems by creating two DNA carriers that could offer some solutions.

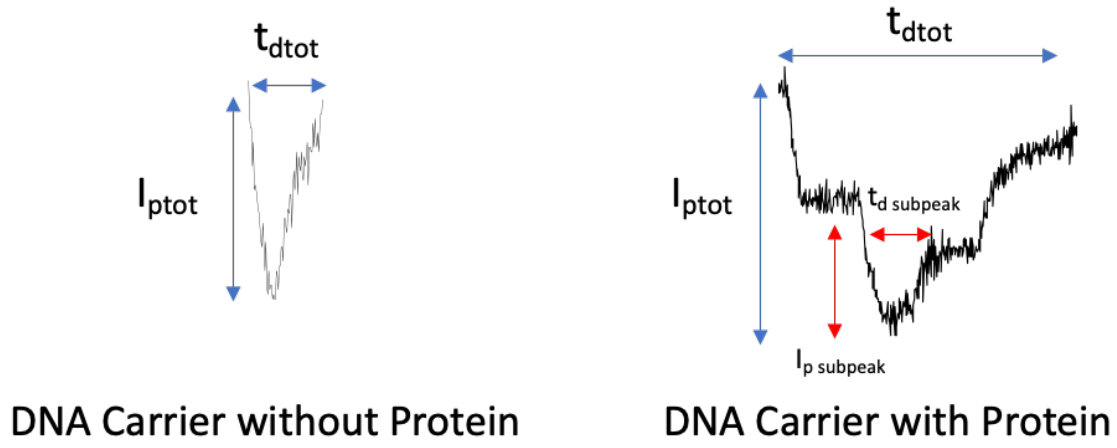


***Figure 1.12**: Sample events for a DNA carrier with and without protein. t_{dtot} is the total dwell time and I_{ptot} is the total peak amplitude. $t_{dsubpeak}$ is the dwell time of the subpeak and $I_{psubpeak}$ is the peak height of the subpeak. For the event with protein, it is possible to clearly see that there is a secondary subpeak that corresponds to the protein.*

1.7.2 EXAMPLES OF DNA CARRIERS

There are several examples of DNA carriers being used in nanopore sensing. Lambda DNA has been shown to be a successful carrier. It was first accomplished with the binding of lambda DNA and anti-DNA antibodies. In 1 M KCl pH 8.0 conditions, these antibodies are positively charged and therefore able to bind strongly to the negatively charged DNA. With the antibodies a clear subpeak emerged and the more antibodies that were bound, the longer the total event was. Additionally, the spikes were found randomly all over the events, which corresponds to the random binding based on charge.[155] Lambda DNA also has been modified directly to form a carrier. Three probes were able to be attached to lambda DNA, such that it was able to be used to perform multiplexed sensing and sense specific proteins in serum.[40] **Figure 1.13a** shows sample events from this carrier. A 7.2 Kbp DNA molecule with chemical motifs has also been successfully used as a carrier. This carrier was formed by hybridising a single-stranded DNA scaffold to probes. These chemical motifs appear at regular intervals so that it is possible to bind to more than one protein at a time. As more proteins were bound, the larger the secondary peak was. Additionally, with this carrier specific proteins were able to be detected from a mixture.[147] **Figure 1.13b** shows what this carrier and its corresponding translocation events look like. This carrier has also been used to successfully quantify protein concentrations[128] and perform single nucleotide polymorphism (SNP) detection. For the SNP detection, biotin is hybridised to the carrier and

then incubated with streptavidin.[128] The presence of the protein of interest, a mutant Janus kinase 2 (JAK2), was confirmed by its displacement of the streptavidin, and the subsequent lack of secondary peaks.[156] A probe made from a breast cancer cell-specific aptamer and complimentary single-stranded DNA has also been used as a DNA carrier. When the aptamer binds, the strands are amplified to form three-way junction structure DNA. This DNA and protein can then be identified using a nanopore.[157]

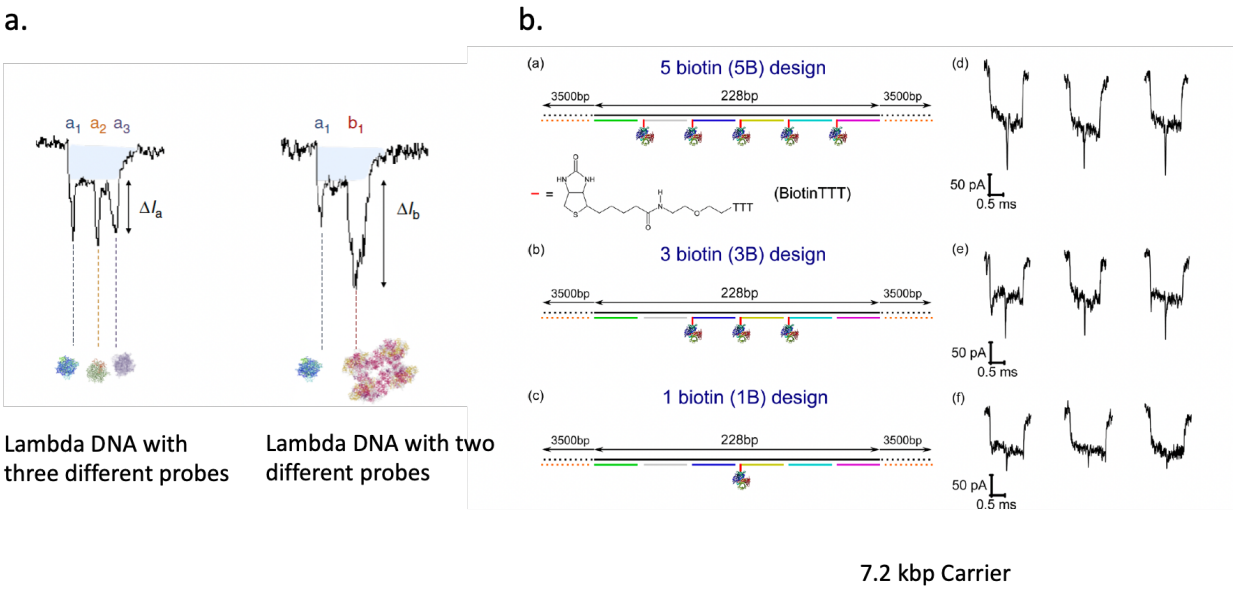


Figure 1.13: *Sample events for lambda DNA with multiple probes. When all of the probes are different, three different subpeaks occur, but only two when there are two probes. Additionally in contrast to the 7.2 Kbp where all of the biotins are very close together, there is enough space between probes for separate subpeaks. In the 7.2 Kbp carrier, sample events are also shown. However, as the total distance between probes is very short, a max of 228 bases, instead of separate peaks occurring when more streptavidin are bound, the subpeak's amplitude increases. Adapted from References 40 and 147.*[40,147]

1.7.3 APTAMERS

Molecular recognition elements are a key part to any DNA carrier. Given the desire for small size and easy incorporation, aptamers are used in this research. Antibodies are another common tool for molecular recognition elements, but they are often too large, not easily incorporated, vary from batch to batch, and do not have the require sensing capabilities.[158]

Aptamers are made through the systematic evolution of ligands by exponential enrichment or SELEX process.[159] The process works by screening large combinatorial libraries in an iterative manner. The library is made up of linear oligonucleotides of a unique

sequence. The first step is incubating the target at a specific temperature with the library, and this cycle is repeated, often for 8-15 cycles until a target is found. Each cycle selects for and enriches specific nucleic acid sequences. The sequences are amplified using PCR.[158] Aptamers are a valuable tool as the process of making them is in vitro and does not rely upon animals or cells, and as such the desired properties are easily tailored. The purity and reproducibility levels are also very high. They are suitable for long term storage and are easier and cheaper to work with than antibodies. Additionally, as they can be tailored so precisely, they are able to identify molecules that differ only by a methyl group, making them more specific than antibodies. They also can have dissociation constants in the low picomolar range, which is a requirement for nanopore sensing. As they are oligonucleotides, they provide a way to avoid extensive additional pre-treatment, particularly if the DNA carrier is already being modified. Their main limitation is nuclease degradation.[158,160,161] Given all of these characteristics aptamers make the ideal molecular recognition element to use for DNA carriers, and some aptamers, specifically for thrombin, will be used in this research.[162–164]

OVERALL AIMS OF THE PROJECT AND OUTLINE

The overall aims of this project are to investigate the use of novel DNA dendrimers and plasmids as DNA carriers. Establishing that both these carriers are able to form distinctive nanopore signals with and without protein is a key aspect. These experiments will use well-established proteins for the proof of concept, thrombin for the dendrimers and the biotin streptavidin bond for the plasmids. Lastly, determining if these carriers are able to perform multiplexed sensing is an important goal. This thesis will describe the results of these DNA carrier experiments and will aim to address some of the common challenges in the nanopore field and DNA carriers specifically.

Chapter 1 discusses the necessary background information for this research, starting with the origin of nanopores and their working principles. A discussion of the various types of pores is also included. DNA and protein translocations working principles, applications, and challenges are examined. Then, nanopore modifications and DNA carriers are discussed, including an in-depth exploration of DNA carriers and the aptamers as a molecular

recognition element. The aims of this project are briefly summarised, followed by an overview of the chapters of this thesis.

Chapter 2 discusses the materials and methods used in this thesis. It includes the fabrication of the nanopipettes using a laser pipette puller and the general techniques used for DNA modification to modify the carriers. Techniques that were used to confirm successful modification of the carriers are also discussed.

Chapter 3 contains an in-depth exploration of DNA dendrimers and why they were selected as a potential carrier structure. The formation of the dendrimer is discussed thoroughly. Lastly, the nanopore results for these DNA carriers are discussed, including results for the carrier at several different sizes and with and without protein. The final section examines why these DNA dendrimers are not effective carriers.

Chapter 4 demonstrates that plasmids offer much potential as a DNA carrier. The modification of a specific 10 Kbp plasmid is shown, as well as its behaviour in the nanopore. Voltage studies and the effect of pH are also discussed. Finally, successful modification is confirmed using quadrivalent streptavidin binding.

Chapter 5 explores the use of the plasmids as a DNA carrier. The necessary data analysis to automate the event results was discussed. Cis-to-trans binding curves for both one and two biotins are shown, as well as trans-to-cis binding curves for one biotin. For cis-to-trans, there is also a comparison of monovalent and quadrivalent streptavidin, and the use of the carrier in sandwich-like assay.

Chapter 6 summarises the conclusions of this thesis and gives a brief outlook on potential future work.

REFERENCES

1. Wu, L. *et al.* Application of nano-ELISA in food analysis: Recent advances and challenges. *TrAC - Trends Anal. Chem.* **113**, 140–156 (2019).
2. Zhu, H. *et al.* PCR past, present and future. *Biotechniques* **69**, 317–325 (2020).
3. Peng, P. *et al.* Emerging Elisa Derived Technologies for in Vitro Diagnostics. *SSRN Electron. J.* (2022). doi:10.2139/ssrn.3988122
4. Mehrotra, P. Biosensors and their applications - A review. *J. Oral Biol. Craniofacial Res.* **6**, 153–159 (2016).
5. Perumal, V. & Hashim, U. Advances in biosensors: Principle, architecture and applications. *J. Appl. Biomed.* **12**, 1–15 (2014).
6. Ha, T., Kaiser, C., Myong, S., Wu, B. & Xiao, J. Next generation single-molecule techniques: Imaging, labeling, and manipulation in vitro and in cellulo. *Mol. Cell* **82**, 304–314 (2022).
7. Cornish, Peter V., Ha, T. R eview. *ACS Chem. Biol.* **2**, 53–61 (2007).
8. Edel, J. B. & Albrecht, T. *Engineered Nanopores for Bioanalytical Applications*. (Elsevier, 2013).
9. Keyser, U. F. Enhancing nanopore sensing with DNA nanotechnology. *Nat. Nanotechnol.* **11**, 106–108 (2016).
10. Branton, D. *et al.* The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **26**, 1146–1153 (2008).
11. Graham, M. D. The Coulter Principle: Foundation of an industry. *JALA - J. Assoc. Lab. Autom.* **8**, 72–81 (2003).
12. Graham, M. D., DeBlois, R. W., Bean, C. P. & Wesley, R. K. A. Electrokinetic measurements with submicron particles and pores by the resistive pulse technique. *JALA - J. Assoc. Lab. Autom.* **61**, 323–335 (2003).
13. Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci.* **93**, 13770–13773 (1996).
14. Deamer, D., Akeson, M. & Branton, D. Three decades of nanopore sequencing. *Nat. Biotechnol.* **34**, 518–524 (2016).
15. Akeson, M., Branton, D., Kasianowicz, J. J., Brandin, E. & Deamer, D. W. Microsecond

time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* **77**, 3227–3233 (1999).

16. Meller, A., Nivon, L., Brandin, E., Golovchenko, J. & Branton, D. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1079–1084 (2000).
17. Wang, H., Dunning, J. E., Huang, A. P. H., Nyamwanda, J. A. & Branton, D. DNA heterogeneity and phosphorylation unveiled by single-molecule electrophoresis. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13472–13477 (2004).
18. Mathé, J., Aksimentiev, A., Nelson, D. R., Schulten, K. & Meller, A. Orientation discrimination of single-stranded DNA inside the α-hemolysin membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12377–12382 (2005).
19. Butler, T. Z., Gundlach, J. H. & Troll, M. A. Determination of RNA orientation during translocation through a biological nanopore. *Biophys. J.* **90**, 190–199 (2006).
20. Ashkenasy, N., Sánchez-Quesada, J., Bayley, H. & Ghadiri, M. R. Recognizing a Single Base in an Individual DNA Strand: A Step Toward DNA Sequencing in Nanopores. *Angew. Chemie* **117**, 1425–1428 (2005).
21. Stoddart, D., Heron, A. J., Mikhailova, E., Maglia, G. & Bayley, H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7702–7707 (2009).
22. Manrao, E. A. *et al.* Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* **30**, 349–353 (2012).
23. Laszlo, A. H. *et al.* Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18904–18909 (2013).
24. Schreiber, J. *et al.* Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18910–18915 (2013).
25. Laszlo, A. H. *et al.* Decoding long nanopore sequencing reads of natural DNA. *Nat. Biotechnol.* **32**, 829–833 (2014).
26. Wescoe, Z. L., Schreiber, J. & Akeson, M. Nanopores discriminate among five C5-cytosine variants in DNA. *J. Am. Chem. Soc.* **136**, 16582–16587 (2014).

27. Wanunu, M. Nanopores: A journey towards DNA sequencing. *Phys. Life Rev.* **9**, 125–158 (2012).
28. Kesselheim, S., Müller, W. & Holm, C. Origin of current blockades in nanopore translocation experiments. *Phys. Rev. Lett.* **112**, 1–5 (2014).
29. Ren, R. Novel tools and strategies of nanopore single-molecule biosensing. (Imperial College London, 2019).
30. Haywood, D. G., Saha-Shah, A., Baker, L. A. & Jacobson, S. C. Fundamental studies of nanofluidics: Nanopores, nanochannels, and nanopipets. *Anal. Chem.* **87**, 172–187 (2015).
31. Smeets, R. M. M. *et al.* Salt dependence of ion transport and DMA translocation through solid-state nanopores. *Nano Lett.* **6**, 89–95 (2006).
32. Zanjani, M. B., Engelke, R. E., Lukes, J. R., Meunier, V. & Drndić, M. Up and down translocation events and electric double-layer formation inside solid-state nanopores. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **92**, (2015).
33. Ayub, M. *et al.* Nanopore/electrode structures for single-molecule biosensing. *Electrochim. Acta* **55**, 8237–8243 (2010).
34. Siwy, Z. & Fuliński, A. A nanodevice for rectification and pumping ions. *Am. J. Phys.* **72**, 567–574 (2004).
35. Makra, I. & Gyurcsányi, R. E. Electrochemical sensing with nanopores: A mini review. *Electrochemistry Communications* **43**, 55–59 (2014).
36. Ying, Y. & Sze, J. Selective Single Molecule Sensing in Nanopore. (2017).
37. Nehra, A., Ahlawat, S. & Singh, K. P. A biosensing expedition of nanopore: A review. *Sensors Actuators, B Chem.* **284**, 595–622 (2019).
38. Gu, Z. *et al.* Exponentially modified Gaussian relevance to the distributions of translocation events in Nanopore-based single molecule detection. *Chinese Chem. Lett.* **25**, 1029–1032 (2014).
39. Varongchayakul, N., Song, J., Meller, A. & Grinstaff, M. W. Single-molecule protein sensing in a nanopore: a tutorial. *Chem. Soc. Rev.* (2018). doi:10.1039/C8CS00106E
40. Sze, J. Y. Y., Ivanov, A. P., Cass, A. E. G. & Edel, J. B. Single molecule multiplexed nanopore protein screening in human serum using aptamer modified DNA carriers. *Nat. Commun.* (2017). doi:10.1038/s41467-017-01584-3
41. Bell, N. A. W., Muthukumar, M. & Keyser, U. F. Translocation frequency of double-

stranded DNA through a solid-state nanopore. *Phys. Rev. E* **93**, 1–10 (2016).

42. Lu, B., Albertorio, F., Hoogerheide, D. P. & Golovchenko, J. A. Origins and consequences of velocity fluctuations during DNA passage through a nanopore. *Biophys. J.* **101**, 70–79 (2011).
43. Lowe, B. M., Skylaris, C. K. & Green, N. G. Acid-base dissociation mechanisms and energetics at the silica-water interface: An activationless process. *J. Colloid Interface Sci.* **451**, 231–244 (2015).
44. Allagui, A., Benaoum, H. & Olendski, O. On the Gouy–Chapman–Stern model of the electrical double-layer structure with a generalized Boltzmann factor. *Phys. A Stat. Mech. its Appl.* **582**, 126252 (2021).
45. Oldham, K. B. A Gouy-Chapman-Stern model of the double layer at a (metal)/(ionic liquid) interface. *J. Electroanal. Chem.* **613**, 131–138 (2008).
46. Torrie, G. M. & Valleau, J. P. Electrical double layers. 4. Limitations of the Gouy-Chapman theory. *J. Phys. Chem.* **86**, 3251–3257 (1982).
47. Bearden, S., Simpanen, E. & Zhang, G. Active current gating in electrically biased conical nanopores. *Nanotechnology* **26**, 1–11 (2015).
48. Experton, J., Wu, X. & Martin, C. R. From ion current to electroosmotic flow rectification in asymmetric nanopore membranes. *Nanomaterials* **7**, (2017).
49. Schoch, R. B., Han, J. & Renaud, P. Transport phenomena in nanofluidics. *Rev. Mod. Phys.* **80**, 839–883 (2008).
50. Firnkes, M., Pedone, D., Knezevic, J., Döblinger, M. & Rant, U. Electrically facilitated translocations of proteins through silicon nitride nanopores: Conjoint and competitive action of diffusion, electrophoresis, and electroosmosis. *Nano Lett.* **10**, 2162–2167 (2010).
51. Hugh, J. M., Andresen, K. & Keyser, U. F. Cation dependent electroosmotic flow in glass nanopores. *arXiv* **1**, 1–5 (2019).
52. Schmid, G. Electrochemistry of capillary systems with narrow pores. VI. Convection conductivity (theoretical considerations). *J. Memb. Sci.* **150**, 211–225 (1998).
53. Ying, L. Applications of nanopipettes in bionanotechnology. *Biochem. Soc. Trans.* **37**, 702–706 (2009).
54. Freedman, K. J. *et al.* Nanopore sensing at ultra-low concentrations using single-molecule dielectrophoretic trapping. *Nat. Commun.* **7**, 1–9 (2016).

55. Ying, L. *et al.* The scanned nanopipette: A new tool for high resolution bioimaging and controlled deposition of biomolecules. *Phys. Chem. Chem. Phys.* **7**, 2859–2866 (2005).
56. Weit, C., Bard, A. J. & Feldberg, S. W. Current Rectification at Quartz Nanopipet Electrodes. *Anal. Chem.* **69**, 4627–4633 (1997).
57. Siwy, Z. S. Ion-current rectification in nanopores and nanotubes with broken symmetry. *Adv. Funct. Mater.* **16**, 735–746 (2006).
58. Wang, Z. *et al.* Nanopipettes: A potential tool for DNA detection. *Analyst* **144**, 5037–5047 (2019).
59. Cai, X. H. *et al.* Reversing current rectification to improve DNA-sensing sensitivity in conical nanopores. *Electrophoresis* **40**, 2098–2103 (2019).
60. Dessaux, D., Mathé, J., Ramirez, R. & Basdevant, N. Current rectification and ionic selectivity of alpha-hemolysin: Coarse-Grained Molecular Dynamics simulations. **1**, (2022).
61. Liu, G. C. *et al.* Ion current rectification in combination with ion current saturation. *Anal. Chim. Acta* **1117**, 35–40 (2020).
62. Zhao, C. *et al.* Effect of Anion Species on Ion Current Rectification Properties of Positively Charged Nanochannels. *ACS Appl. Mater. Interfaces* **12**, 28915–28922 (2020).
63. Oukhaled, A., Bacri, L., Pastoriza-Gallego, M., Betton, J. M. & Pelta, J. Sensing proteins through nanopores: Fundamental to applications. *ACS Chem. Biol.* **7**, 1935–1949 (2012).
64. Bell, N. A. W. & Keyser, U. F. Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. *Nat. Nanotechnol.* **11**, 645–651 (2016).
65. Liang, S. *et al.* Noise in nanopore sensors: Sources, models, reduction, and benchmarking. *Nami Jishu yu Jingmi Gongcheng/Nanotechnology Precis. Eng.* **3**, 9–17 (2020).
66. Tabard-Cossa, V., Trivedi, D., Wiggin, M., Jetha, N. N. & Marziali, A. Noise analysis and reduction in solid-state nanopores. *Nanotechnology* **18**, (2007).
67. Smeets, R. M. M., Keyser, U. F., Dekker, N. H. & Dekker, C. Noise in solid-state nanopores. *Proc. Natl. Acad. Sci.* **105**, 417–421 (2008).
68. Crnković, A., Srnko, M. & Anderluh, G. Biological nanopores: Engineering on demand.

Life **11**, 1–30 (2021).

69. Krasniqi, B. Nanopore sensing of peptides and proteins. (University of Saskatchewan, Saskatoon, Saskatchewan S7N 0W0, 2013).
70. Song, L. *et al.* Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. *Science (80-.).* **274**, 1859–1866 (1996).
71. Braha, O. *et al.* Designed protein pores as components for biosensors. *Chem. Biol.* **4**, 497–505 (1997).
72. Oukhaled, G. *et al.* Unfolding of proteins and long transient conformations detected by single nanopore recording. *Phys. Rev. Lett.* **98**, 98–101 (2007).
73. Lathrop, D. K. *et al.* Monitoring the escape of DNA from a nanopore using an alternating current signal. *J. Am. Chem. Soc.* **132**, 1878–1885 (2010).
74. Ying, Y.-L., Cao, C. & Long, Y.-T. Single molecule analysis by biological nanopore sensors. *Analyst* **139**, 3826–3835 (2014).
75. Haque, F., Li, J., Wu, H. C., Liang, X. J. & Guo, P. Solid-state and biological nanopore for real-time sensing of single chemical and sequencing of DNA. *Nano Today* **8**, 56–74 (2013).
76. Soskine, M. *et al.* An engineered ClyA nanopore detects folded target proteins by selective external association and pore entry. *Nano Lett.* **12**, 4895–4900 (2012).
77. Stefureac, R., Long, Y. T., Kraatz, H. B., Howard, P. & Lee, J. S. Transport of α-helical peptides through α-hemolysin and aerolysin pores. *Biochemistry* **45**, 9172–9179 (2006).
78. Wendell, D. *et al.* Translocation of double-stranded DNA through membrane-adapted phi29 motor protein nanopores. *Nat. Nanotechnol.* **4**, 765–772 (2009).
79. tan, cherie, Fleming, A. M., Ren, H., Burrows, C. J. & White, H. S. γ-Hemolysin Nanopore is Sensitive to Guanine-to-Inosine Substitutions in Double-Stranded DNA at the Single-Molecule Level. *J. Am. Chem. Soc.* jacs.8b08153 (2018). doi:10.1021/jacs.8b08153
80. Shi, W., Friedman, A. K. & Baker, L. A. Nanopore Sensing. *Anal. Chem.* **89**, 157–188 (2017).
81. Majd, S. *et al.* Applications of biological pores in nanomedicine, sensing, and nanoelectronics. *Curr. Opin. Biotechnol.* **21**, 439–476 (2010).
82. Li, J. *et al.* Ion-beam sculpting at nanometre length scales. *Nature* **412**, 166–169

(2001).

83. Storm, A. J., Chen, J. H., Ling, X. S., Zandbergen, H. W. & Dekker, C. Fabrication of solid-state nanopores with single-nanometre precision. *Nat. Mater.* **2**, 537–540 (2003).

84. Sugita, T., Hiramatsu, K., Ikeda, S. & Matsumura, M. Fabrication of pores in a silicon carbide wafer by electrochemical etching with a glassy-carbon needle electrode. *ACS Appl. Mater. Interfaces* **5**, 2580–2584 (2013).

85. Wu, M.-Y., Krapf, D., Zandbergen, M., Zandbergen, H. & Batson, P. E. Formation of nanopores in a SiN⁄SiO2 membrane with an electron beam. *Appl. Phys. Lett.* **87**, 113106 (2005).

86. Venkatesan, B. M., Shah, A. B., Zuo, J. M. & Bashir, R. DNA sensing using nanocrystalline surface-enhanced Al2O 3 nanopore sensors. *Adv. Funct. Mater.* **20**, 1266–1275 (2010).

87. Garaj, S. *et al.* Graphene as a subnanometre trans-electrode membrane. *Nature* **467**, 190–193 (2010).

88. Schneider, G. F. *et al.* DNA Translocation through Graphene Nanopores. doi:10.1021/nl102069z

89. Cai, Q., Ledden, B., Krueger, E., Golovchenko, J. A. & Li, J. Nanopore sculpting with noble gas ions. *J. Appl. Phys.* **100**, 24914 (2006).

90. Fuliski, A., Kosiska, I. D., Siwy -, Z., Yu Apel, P. & Blonskaya, I. V. Rectification and voltage gating of ion currents in a nanofabricated pore On the validity of continuous modelling of ion transport through nanochannels Effect of nanopore geometry on ion current rectification. (2002).

91. Howorka, S. & Siwy, Z. Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.* **38**, 2360 (2009).

92. Venkatesan, B. M. *et al.* Highly sensitive, mechanically stable nanopore sensors for DNA analysis. *Adv. Mater.* **21**, 2771–2776 (2009).

93. Mussi, V. *et al.* DNA-functionalized solid state nanopore for biosensing. *Nanotechnology* **21**, (2010).

94. Schneider, G. F. *et al.* DNA Translocation through Grapheme Nanopores. *Nano Lett.* **10**, 3163–3167 (2010).

95. Steinbock, L. J., Otto, O., Chimerel, C., Gornall, J. & Keyser, U. F. Detecting DNA

folding with nanocapillaries. *Nano Lett.* **10**, 2493–2497 (2010).

96. Morris, C. A., Friedman, A. K. & Baker, L. A. Applications of nanopipettes in the analytical sciences. *Analyst* **135**, 2190–2202 (2010).
97. Stanley, J. & Pourmand, N. Nanopipettes - The past and the present. *APL Mater.* **8**, (2020).
98. Gong, X. *et al.* Label-free in-flow detection of single DNA molecules using glass nanopipettes. *Anal. Chem.* **86**, 835–841 (2014).
99. Steinbock, L. J., Lucas, A., Otto, O. & Keyser, U. F. Voltage-driven transport of ions and DNA through nanocapillaries. *Electrophoresis* **33**, 3480–3487 (2012).
100. Actis, P., Mak, A. C. & Pourmand, N. Functionalized nanopipettes: Toward label-free, single cell biosensors. *Bioanal. Rev.* **1**, 177–185 (2010).
101. Steinbock, L. J. *et al.* Probing DNA with micro- and nanocapillaries and optical tweezers. *J. Phys. Condens. Matter* **22**, (2010).
102. Fyta, M. Threading DNA through nanopores for biosensing applications. *Journal of Physics Condensed Matter* (2015). doi:10.1088/0953-8984/27/27/273101
103. Storm, A. J., Chen, J. H., Zandbergen, H. W. & Dekker, C. Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **71**, 1–10 (2005).
104. Storm, A. J. *et al.* Fast DNA translocation through a solid-state nanopore. *Nano Lett.* **5**, 1193–1197 (2005).
105. Meller, A., Nivon, L. & Branton, D. Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.* **86**, 3435–3438 (2001).
106. Kubota, T. *et al.* Clog and Release, and Reverse Motions of DNA in a Nanopore. *Polymers (Basel).* **11**, 84 (2019).
107. Li, J. & Talaga, D. S. The distribution of DNA translocation times in solid-state nanopores. *J. Phys. Condens. Matter* **22**, (2010).
108. Wanunu, M., Sutin, J., McNally, B., Chow, A. & Meller, A. DNA translocation governed by interactions with solid-state nanopores. *Biophys. J.* **95**, 4716–4725 (2008).
109. Chen, P. *et al.* Probing single DNA molecule transport using fabricated nanopores. *Nano Lett.* **4**, 2293–2298 (2004).
110. Das, N., Ray, R., Ray, S. & Chaudhuri, C. R. Intelligent Quantification of Picomolar Protein Concentration in Serum by Functionalized Nanopores. *IEEE Sens. J.* **1748**, 1–1

(2018).

111. Bell, N. A. W., Chen, K., Ghosal, S., Ricci, M. & Keyser, U. F. Asymmetric dynamics of DNA entering and exiting a strongly confining nanopore. *Nat. Commun.* **8**, 1–8 (2017).
112. Chen, K., Bell, N. A. W., Kong, J., Tian, Y. & Keyser, U. F. Direction- and Salt-Dependent Ionic Current Signatures for DNA Sensing with Asymmetric Nanopores. *Biophys. J.* **112**, 674–682 (2017).
113. Kowalczyk, S. W. & Dekker, C. Measurement of the docking time of a DNA molecule onto a solid-state nanopore. *Nano Lett.* **12**, 4159–4163 (2012).
114. Plesa, C. *et al.* Direct observation of DNA knots using a solid-state nanopore. *Nat. Nanotechnol.* **11**, 1093–1097 (2016).
115. Fologea, D., Brandin, E., Uplinger, J., Branton, D. & Li, J. DNA conformation and base number simultaneously determined in a nanopore. *Electrophoresis* **28**, 3186–3192 (2007).
116. Wang, Y., Zheng, D., Tan, Q., Wang, M. X. & Gu, L. Q. Nanopore-based detection of circulating microRNAs in lung cancer patients. *Nat. Nanotechnol.* **6**, 668–674 (2011).
117. Liu, L. & Wu, H. C. DNA-Based Nanopore Sensing. *Angew. Chemie - Int. Ed.* **55**, 15216–15222 (2016).
118. Plesa, C. *et al.* Fast translocation of proteins through solid state nanopores. *Nano Lett.* **13**, 658–663 (2013).
119. Yuan, Zhishan; Liu, Youming; Dai, Min; Yi, Xin; Wang, C. Controlling DNA Translocation Through Solid-state Nanopores. *Nanoscale Res. Lett.* **15**, (2020).
120. Fologea, D., Uplinger, J., Thomas, B., McNabb, D. S. & Li, J. Slowing DNA translocation in a solid-state nanopore. *Nano Lett.* **5**, 1734–1737 (2005).
121. Larkin, J. *et al.* Slow DNA transport through nanopores in hafnium oxide membranes. *ACS Nano* **7**, 10121–10128 (2013).
122. Liu, S. *et al.* Boron nitride nanopores: Highly sensitive DNA single-molecule detectors. *Adv. Mater.* **25**, 4549–4554 (2013).
123. Squires, A. H., Hersey, J. S. & Meller, A. A Nanopore − Nano fi ber Mesh Biosensor To Control DNA Translocation. (2013). doi:10.1021/ja408685x
124. Steinbock, L. J., Steinbock, J. F. & Radenovic, A. Controllable shrinking and shaping of glass nanocapillaries under electron irradiation. *Nano Lett.* **13**, 1717–1723 (2013).
125. Cadinu, P. *et al.* Single Molecule Trapping and Sensing Using Dual Nanopores

Separated by a Zeptoliter Nanobridge. *Nano Lett.* **17**, 6376–6384 (2017).

126. Keyser, U. F. Controlling molecular transport through nanopores. *J. R. Soc. Interface* **8**, 1369–1378 (2011).
127. Movileanu, L. Interrogating single proteins through nanopores: challenges and opportunities. *Trends Biotechnol.* **27**, 333–341 (2009).
128. Kong, J., Bell, N. A. W. & Keyser, U. F. Quantifying Nanomolar Protein Concentrations Using Designed DNA Carriers and Solid-State Nanopores. *Nano Lett.* (2016). doi:10.1021/acs.nanolett.6b00627
129. Sutherland, T. C. *et al.* Structure of peptides investigated by nanopore analysis. *Nano Lett.* **4**, 1273–1277 (2004).
130. Movileanu, L. Interrogating single proteins through nanopores: challenges and opportunities. *Trends in Biotechnology* (2009). doi:10.1016/j.tibtech.2009.02.008
131. Luo, Y., Wu, L., Tu, J. & Lu, Z. Application of solid-state nanopore in protein detection. *Int. J. Mol. Sci.* **21**, (2020).
132. Jetha, N. N., Semenchenko, V., Wishart, D. S., Cashman, N. R. & Marziali, A. Nanopore Analysis of Wild-Type and Mutant Prion Protein (PrPC): Single Molecule Discrimination and PrPC Kinetics. *PLoS One* (2013). doi:10.1371/journal.pone.0054982
133. Houghtaling, J., List, J. & Mayer, M. Nanopore-Based, Rapid Characterization of Individual Amyloid Particles in Solution: Concepts, Challenges, and Prospects. *Small* **14**, (2018).
134. Madampage, C., Tavassoly, O., Christensen, C., Kumari, M. & Lee, J. S. Nanopore analysis: An emerging technique for studying the folding and misfolding of proteins. in *Prion* (2012). doi:10.4161/pri.18665
135. Wang, X. *et al.* Single-molecule nanopore sensing of actin dynamics and drug binding. *bioRxiv* (2019). doi:10.1101/829150
136. Si, W. & Aksimentiev, A. Nanopore Sensing of Protein Folding. *ACS Nano* **11**, 7091–7100 (2017).
137. Cressiot, B., Bacri, L. & Pelta, J. The Promise of Nanopore Technology: Advances in the Discrimination of Protein Sequences and Chemical Modifications. *Small Methods* **4**, 1–13 (2020).
138. Kennedy, E., Dong, Z., Tennant, C. & Timp, G. Reading the primary structure of a

protein with 0.07 nm 3 resolution using a subnanometre-diameter pore. *Nat. Nanotechnol.* **11**, 968–976 (2016).

139. Howorka, S. & Siwy, Z. S. Reading amino acids in a nanopore. *Nat. Biotechnol.* **38**, 159–160 (2020).
140. Ouldali, H. *et al.* Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore. *Nat. Biotechnol.* **38**, 176–181 (2020).
141. Saharia, J. *et al.* Molecular-Level Profiling of Human Serum Transferrin Protein through Assessment of Nanopore-Based Electrical and Chemical Responsiveness. *ACS Nano* **13**, 4246–4254 (2019).
142. Dong, Z., Kennedy, E., Hokmabadi, M. & Timp, G. Discriminating Residue Substitutions in a Single Protein Molecule Using a Sub-nanopore. *ACS Nano* **11**, 5440–5452 (2017).
143. Restrepo-Pérez, L., Wong, C. H., Maglia, G., Dekker, C. & Joo, C. Label-Free Detection of Post-translational Modifications with a Nanopore. *Nano Lett.* **19**, 7957–7964 (2019).
144. Houghtaling, J., List, J. & Mayer, M. Nanopore-Based, Rapid Characterization of Individual Amyloid Particles in Solution: Concepts, Challenges, and Prospects. *Small* **14**, (2018).
145. Rotem, D., Jayasinghe, L., Salichou, M. & Bayley, H. Protein detection by nanopores equipped with aptamers. *J. Am. Chem. Soc.* (2012). doi:10.1021/ja2105653
146. Harrington, L., Alexander, L. T., Knapp, S. & Bayley, H. Single-Molecule Protein Phosphorylation and Dephosphorylation by Nanopore Enzymology. *ACS Nano* acsnano.8b07697 (2018). doi:10.1021/acsnano.8b07697
147. Bell, N. A. W. & Keyser, U. F. Specific protein detection using designed DNA carriers and nanopores. *J. Am. Chem. Soc.* **137**, 2035–2041 (2015).
148. Acharya, S., Edwards, S. & Schmidt, J. Research highlights: nanopore protein detection and analysis. *Lab Chip* **15**, 3424–3427 (2015).
149. Li, Q., Ying, Y. L., Liu, S. C., Lin, Y. & Long, Y. T. Detection of Single Proteins with a General Nanopore Sensor. *ACS Sensors* **4**, 1185–1189 (2019).
150. Wei, R., Martin, T. G., Rant, U. & Dietz, H. DNA origami gatekeepers for solid-state nanopores. *Angew. Chemie - Int. Ed.* **51**, 4864–4867 (2012).
151. Schibel, A. E. P. & Ervin, E. N. Antigen detection via the rate of ion current rectification change of the antibody-modified glass nanopore membrane. *Langmuir*

30, 11248–11256 (2014).

152. Umehara, S., Karhanek, M., Davis, R. W. & Pourmand, N. Label-free biosensing with functionalized nanopipette probes. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4611–4616 (2009).
153. Ding, S., Gao, C. & Gu, L. Q. Capturing single molecules of immunoglobulin and ricin with an aptamer-encoded glass nanopore. *Anal. Chem.* **81**, 6649–6655 (2009).
154. Cairns-Gibson, D. F. & Cockroft, S. L. Functionalised nanopores: chemical and biological modifications. *Chem. Sci.* **13**, 1869–1882 (2022).
155. Plesa, C., Ruitenberg, J. W., Witteveen, M. J. & Dekker, C. Detection of individual proteins bound along DNA using solid-state nanopores. *Nano Lett.* **15**, 3153–3158 (2015).
156. Kong, J., Zhu, J. & Keyser, U. F. Single molecule based SNP detection using designed DNA carriers and solid-state nanopores. *Chem. Commun.* **53**, 436–439 (2017).
157. Li, X. *et al.* Detection of Circulating Tumor Cells in Breast Cancer Patients by Nanopore Sensing with Aptamer-Mediated Amplification. *ACS Sensors* **5**, 2359–2366 (2020).
158. Jayasena, S. D. Aptamers: An Emerging Class of Molecules That Rival Antibodies in Diagnostics.
159. Ying, Y. L., Wang, H. Y., Sutherland, T. C. & Long, Y. T. Monitoring of an ATP-binding aptamer and its conformational changes using an α-hemolysin nanopore. *Small* **7**, 87–94 (2011).
160. Tombelli, S., Minunni, M. & Mascini, M. Analytical applications of aptamers. *Biosens. Bioelectron.* **20**, 2424–2434 (2005).
161. Meirinho, S. G., Dias, L. G., Peres, A. M. & Rodrigues, L. R. Electrochemical aptasensor for human osteopontin detection using a DNA aptamer selected by SELEX. *Anal. Chim. Acta* **987**, 25–37 (2017).
162. Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A. & Feigon, J. Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci.* (1993). doi:10.1073/pnas.90.8.3745
163. Deng, B. *et al.* Aptamer binding assays for proteins: The thrombin example-A review. *Anal. Chim. Acta* **837**, 1–15 (2014).
164. Xu, Y., Yang, L., Ye, X., He, P. & Fang, Y. An aptamer-based protein biosensor by detecting the amplified impedance signal. *Electroanalysis* **18**, 1449–1456 (2006).9

Chapter 2 Materials and Methods

Table of Contents

INTRODUCTION

This chapter discusses the general lab procedures used throughout this work. It explains how the concept of nanopore sensing was applied for the experiments performed in this thesis. While the principles behind nanopore sensing are discussed in detail in Chapter 1, the experiments used in this thesis all require nanopipettes, made with a pipette puller, and the measurement of small changes in the current, picoAmperes to nanoAmperes, as the charged molecules move through the pore. Certain procedures are discussed in more detail in further chapters, but a general overview of the materials and methods is provided here. This chapter splits the protocols into three sections, (i) nanopore, (ii) DNA and protein sample preparation and modification, and (iii) data analysis. This chapter also has a complete list of all reagents used in this thesis.

NANOPIPETTE FABRICATION

2.1 LASER PULLER

All of the nanopipettes used in these experiments were fabricated using a P-2000 laser-based micropipette puller (Sutter Instrument Co.). The pipettes were made from quartz capillaries that have a filament, an internal ~150 μm diameter glass rod, an outer diameter of 1.0 mm, an inner diameter of 0.5 mm, and a length of 7.5 cm (World Precision Instruments). The filament allows for easy filling through capillary action. They were plasma cleaned (Harrick Plasma) for 20 minutes to remove dust and other organic contaminants. The capillary was then loaded into the laser puller, into the grooves between the two knobs that are tightened to hold the capillary in place, which heats the centre of the capillary with a CO_2 laser and pulls both ends apart, allowing for the formation of pores of very small diameters; **Figure 1.8** shows the working principle of these laser pullers. Various parameters can be controlled, and the settings used for these experiments are shown in **Table 2.1**.

	Program 1	Program 2
HEAT Line 1	775	790
FIL Line 1	4	4
VEL Line 1	30	30
DEL Line 1	170	160
PUL Line 1	80	90
HEAT Line 2	825	825
FIL Line 2	3	3
VEL Line 2	20	20
DEL Line 2	145	135
PUL Line 2	180	185

Table 2.1*: Pulling parameters for the pipette puller to make nanopipettes for these experiments. Two parameters were used as an adjustment of these parameters is often needed to maintain similar sized and shaped pipettes.*

These parameters allow for control of the pore diameter and geometry. HEAT is the amount of energy applied to the quartz with a range of 0 to 999 and the higher the heat the longer the taper length. FIL is the length of the quartz capillary that is heated with a range of 0 to 15. VEL is the velocity of the capillary before its hard final pull is completed with a range of 0 to 255. DEL is the delay between the laser being switched off and the hard final pull with a range of 0 to 255. PUL is the force of the final pull to split it in two with a range value of 0 to 255. An increase in the HEAT, VEL, or PUL parameters generally leads to an increase in taper length and a decrease in pore diameter.[1] Two different programs were used as the puller is susceptible to changes in humidity and temperature, and as such, the parameters sometimes require adjustment to continue forming reliably small pipettes with a similar shape. Additionally, while this research was conducted, the laser puller was moved from one lab building to another, which also contributed to the need for tweaking the parameters. Ideally, the shape and diameter of the pipettes was very similar for both of these programs, but some variation is to be expected. Pipettes that varied from the expected pulling time values, normally between 5.5-6.5 seconds were discarded. This led to the fabrication of pipettes that typically had diameters between 15 and 30 nm. Pipettes that

varied greatly from this diameter range, as measured using electrical characterisation, were discarded.

2.2 OPTICAL AND ELECTRICAL CHARACTERISATION

The pipette tips were characterised using both electrical and optical methods. The electrical method determines pipette size from individual I-V curves over a range of voltages (-400 mV to 400 mV). Ideally, linear behaviour would be expected for these curves due to Ohm's law, where the slope is the conductance of the pore. This is shown in **Equation 2.1**. The conductance of the pore is also inversely proportional to the total resistance, which as discussed in Chapter 1, can be assumed to be the resistance of the pore.

$$\mathrm{G_{pore}} = \frac{I}{V} = \frac{1}{\mathrm{R_{pore}}}$$

***Equation 2.1**: The conductance of the pore is related to the current flow through the pore, I, and the applied bias, V. It is also the reciprocal of the total resistance, as the R_{pore} dominates cis and trans resistances*.

However, the conductance can also be affected by the solvent concentration and surface charges and as such the conductance can be described with more complete equations. **Equation 2.2** describes the conductance for solutions with KCl > 200 mM, that is the bulk solution having a greater effect on the conductance. **Equation 2.3** also includes the effects of the surface charges that have an significant impact on the conductance at low KCl concentrations.[2] These two equations show agreement with the I-V curve results presented in **Figure 2.1** and **Figure 2.2**, where at 100 mM KCl, there is slight rectification, and at 1M KCl the I-V is completely linear.

$$\mathrm{G_{pore}} = \frac{\mathrm{d_i D_i}\pi}{4(\mathrm{l} + \pi/8(\mathrm{d_i} + \mathrm{D_i}))}(\mu_{\mathrm{K}} + \mu_{\mathrm{Cl}})\eta_{\mathrm{KCl}}(\mathrm{c})\mathrm{e}$$

***Equation 2.2**: Equation describing the conductance of the pore where the KCl concentration is above 100 mM KCl. The variables are as follows, d_i the capillary diameter at the tip, D_i is the inner diameter at the base, l is the capillary length , μ_K and μ_{Cl} are the ion mobilities of K^+ and Cl^- respectively, $\eta_{KCl}(c)$ is the number density of the potassium or chloride ions (it is a function of the ion concentration), and e is the elementary charge.*[2]

$$G_{pore} = \frac{d_i D_i \pi(\mu_K + \mu_{Cl})\eta_{KCl}(c)e}{4(l + \pi/8(d_i + D_i))} + \frac{d_i D_i^* \pi}{4(l^* + \pi/8(d_i + D_i^*))}\mu_K 4\sigma(c)/d_i$$

***Equation 2.3**: Equation describing the conductance of the pore where the KCl concentration is less than 100 mM KCl. The variables are as follows, d_i the capillary diameter at the tip, D_i is the inner diameter at the base, l is the capillary length , μ_K and μ_{Cl} are the ion mobilities of K^+ and Cl^- respectively, $\eta_{KCl}(c)$ is the number density of the potassium or chloride ions (it is a function of the ion concentration), e is the elementary charge, $\mu_K 4\sigma(c)/d_i$ is the counter ions' (K^+) contribution to the conductance, D_i^* is the inner diameter at the base at the far end of the capillary, and l^* is the length at the far end. The last two variables only apply at the end of the capillary, as this is where the pipette narrows enough such that surface charges can have an effect. Thus, this equation, unlike Equation 2.2, describes the effects of both the bulk and the surface charges.*[2]

An average of some I-V curves measured in 100 mM KCl TE buffer pH 8.0 is shown in **Figure 2.1**. The measured diameters of these pipettes range from 22-25 nm. This characterisation was done periodically while recording to ensure that the pipette remained unblocked in order to assess the stability of the pipette and calculate the pipette size; the I-V Curves used to determine pipette size were measured at the beginning of recording. This I-V shows minimal rectification, suggesting that the pore is not blocked and is functioning as would be expected. Some rectification due to the charged walls and low electrolyte solution concentration is to be expected. I-Vs were conducted regularly in order to ensure that the pipette was stable and had not become blocked.

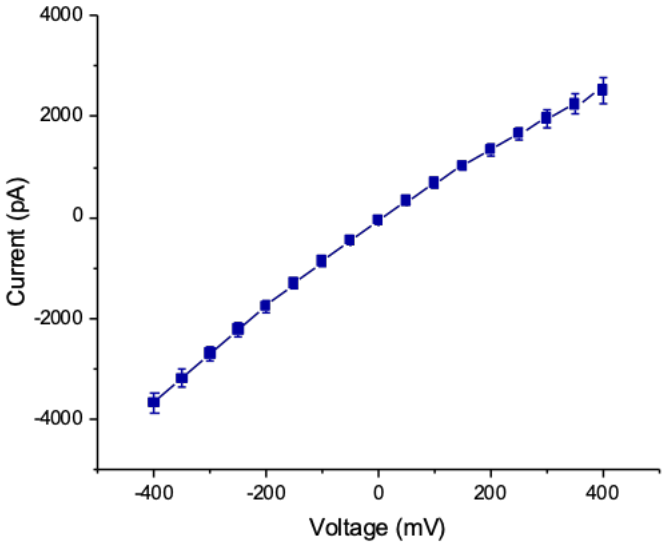


***Figure 2.1**: Average I-V for nanopipette in 100 mM KCl in TE pH 8.0 recorded in cis-to-trans conditions, where the cis side is inside the pipette. Voltage was varied from -400 mV to 400 mV. The experimental conditions were a 30 KHz filter, 100 KHz sampling rate, and 100 pM DNA concentration. The average calculated pore diameter is 23.2 ± 1.1 nm. The data presented is the average of I-Vs from various pipettes and the error is the standard deviation.*

A linear fitting is performed between -50 mV and 50 mV on the I-V graph to determine the diameter of the pores. The slope of this fitting is G, the conductance, and can be inserted into **Equation 2.4** to find the diameter of the nanopore.[3]

$$D = 4Gl/(\pi g d_b)$$

***Equation 2.4**: Equation for finding the diameter of the pore from the I-V curve. D is the diameter of the pore, G is conductance from linearly fitting the I-V curve in nS, l is the length of conical area of the nanopipette, which is assumed to be 1.5 mm, g is the conductance of solvent, either 1.285 S/m for 100 mM KCl or 11.13 S/m for 1 M KCl, and d_b is the diameter of the pipette before convergence, 500 µm.*[3]

These experiments were performed under two different salt solutions, 100 mM KCl and 1M KCl in TE buffer pH 8.0. For experiments conducted at the lower salt conditions, surface charges, in addition to the bulk conductivity, also play a role. This explains the rectification that occurs in the I-V curves in **Figure 2.1**. Sample I-V curves for pipettes used in 1M KCl in TE buffer pH 8.0 are shown in **Figure 2.2**. These pipettes typically were slightly smaller than the ones used for cis-to-trans, ranging from 13-18 nm, as larger pipettes would require a change in the gain so that a wider range of current could be measured, but this

resulted in a loss of SNR. This was especially true for pipettes used in voltage studies, as a slight increase in diameter results in a current above 20,000 pA at 400 mV, the typical upper limit used. It is also possible to see how little rectification occurs at these salt conditions, especially compared to the voltage studies conducted at 100 mM KCl. In both **Figure 2.1** and **2.2**, it is possible to see some of the variation that occurs due to the nanopipette fabrication process that results in differences between the pipettes.

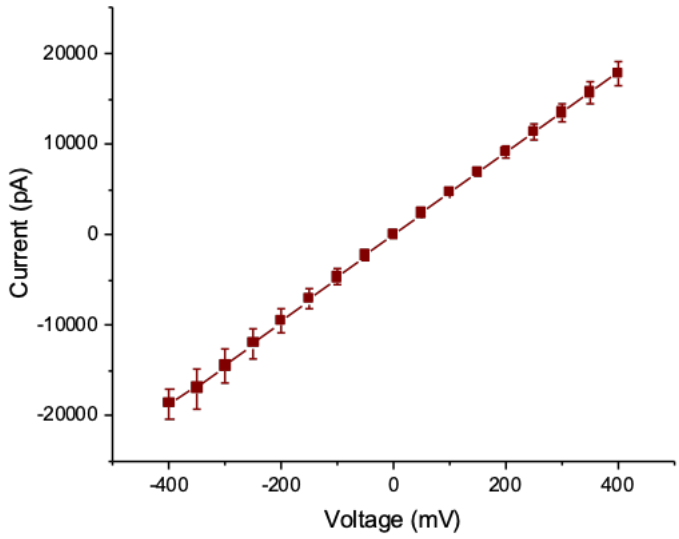


***Figure 2.2**: Average I-Vs for nanopipette in 1M KCl in TE pH 8.0 recorded in trans-to-cis conditions, where the cis side is inside the pipette. Voltage was varied from -400 mV to 400 mV. The experimental conditions were 30 KHz filter, 100 KHz sampling rate, and 100 pM DNA concentration. The average calculated pore diameter is 16.0 ± 1.7 nm. The data presented is the average of I-Vs from various pipettes and the error is the standard deviation.*

The main factors affecting the pore conductance, and thus the I-Vs, are the pore dimensions, including diameter, length, shape, and ionic strength. For the smaller pores, at the lower salt conditions, the surface charges would be a significant factor. However, at the sizes typically used in this thesis for 100 mM KCl, >20 nm, the diameter is greater than the Debye length, and therefore the surface charge will only have a minimal effect.[4]

The optical method for characterisation uses a scanning electron microscope (SEM). The pipette is imaged and the dimensions of the pipettes are measured directly. Pipettes used in these experiments were not imaged with the SEM for any of the plasmid data, due to restricted access to the SEM during Covid-19. **Figure 2.3** shows a sample SEM image for a

pipette used during the dendrimer experiments, and similar figures would be expected for the pipettes used during the plasmid experiments.

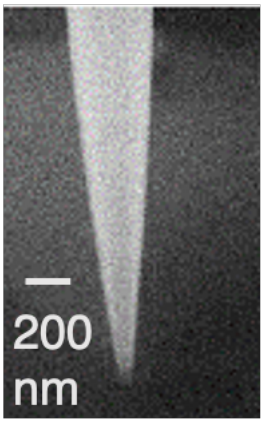


***Figure 2.3**: SEM image of a pipette used in the following experiments. SEM images are used to measure the pipette size as well as to ensure that the pipette shape is as expected. SEM image taken by Dr. Minkyung Kang.*

2.3 EXPERIMENTAL SETUP

The general setup for all nanopore experiments creates a small circuit. **Figure 2.4** shows the general setup schematically. The electrolyte solution is added to a small bath and inside the pipette. Two electrodes, made of silver wire, the reference added to the bath directly (the trans) and the working added through the pipette tip (the cis), are inserted into the bath, along with the pipette tip. The DNA or protein sample can be added to either the pipette, for cis-to-trans experiments, or to the bath, for trans-to-cis experiments. An applied negative bias is used in cis-to-trans experiments and an applied positive bias for trans-to-cis experiments. The sample or electrolyte solution is added to the pipettes using a microfil needle (World Precision Instruments) attached to a 1 mL disposable syringe (Henke-Sass Wolf). Tweezers or a syringe applying negative pressure were used to remove any trapped air bubbles after loading. The bath is filled with the same electrolyte solution as the pipette. The electrodes are then attached to either an Axopatch 200B patch clamp amplifier (Molecular Devices) and a DigiData 1550B digitiser (Molecular Devices) or to a Multiclamp 700B (Molecular Devices) and a DigiData 1440A digitiser (Molecular Devices). Both of these amplifiers were used for the measure of the I-V characteristics of the pipettes for the experiments presented in this thesis. However, for the plasmid results, the Multiclamp was primarily used, due to the flexibility in setting different digital low-pass Bessel filtration

rates, particularly 30 KHz, as this was the value used most often, and the Axopatch uses a 4-pole low-pass Bessel filter, which can only filter at rates of 10 KHz. The low-pass Bessel filter removed signals over the cut off frequency of the filter.[5] The data was filtered and recorded by Clampex 10.7 software (Molecular Devices).

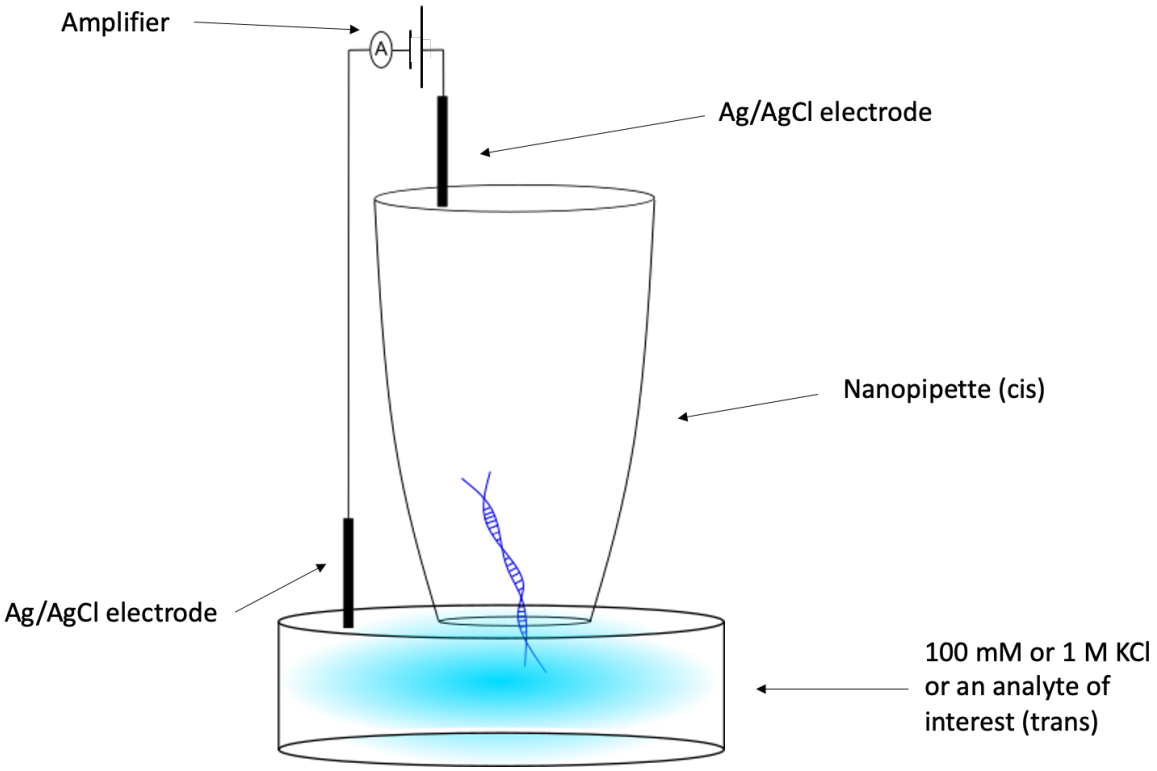


***Figure 2.4**: Schematic for the general nanopore experiments. The DNA or protein could be added to the pipette or the bath and the electrodes were inserted into the bath and the pipette. An applied negative bias is used to drive analytes of interest through the pore from the cis-to-trans while an applied positive bias is used to drive analytes of interests through the pore from the trans-to-cis.*

In these experiments, as the DNA or DNA + Protein, moves through the pore, the current, which is constantly being recorded, fluctuates. For the conditions used, for both cis-to-trans and trans-to-cis experiments, the DNA moving through the pore changes the current for a certain amount of time. A schematic of this is shown in **Figure 2.5**. The length of time that the current is changed is referred to as the dwell time and the change in current is called the peak amplitude. The whole change in current and return to base line as the DNA moves through the pore is called a translocation event or peak. Changes to this peak for samples with added protein are called subpeaks. The baseline value is dependent on several factors, most importantly the electrolyte solution, particularly the ion concentration,

as well as the pipette shape and size. The dwell time and peak amplitude offer information about the molecule moving through the pore, which for unknown samples is very valuable.

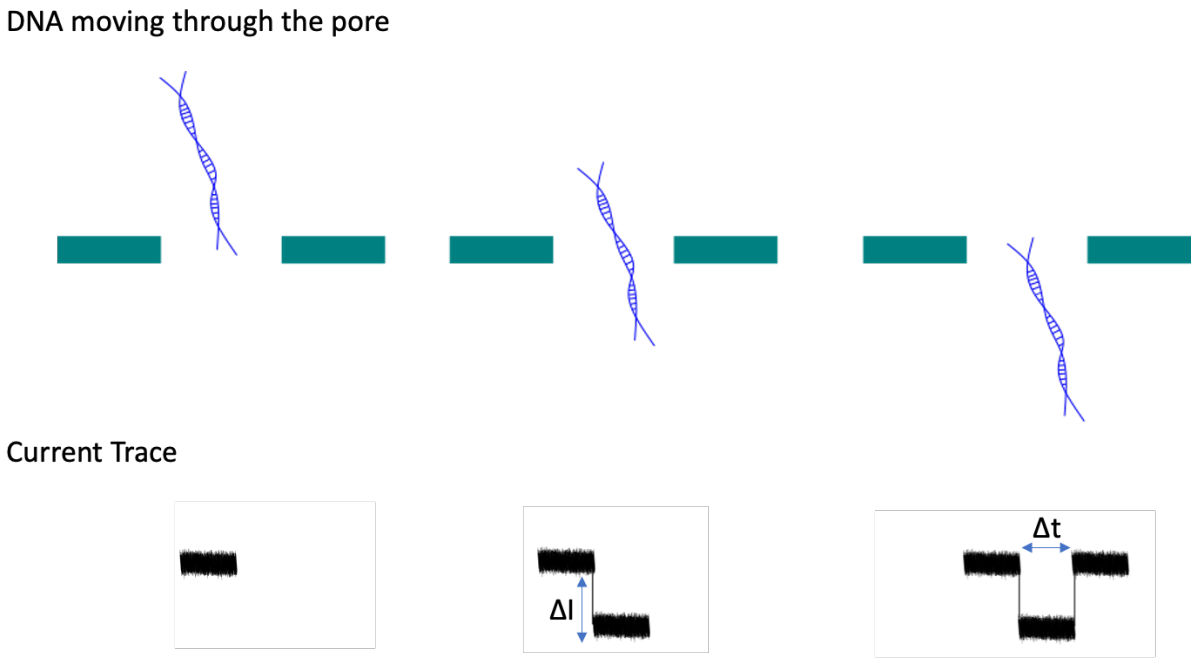


***Figure 2.5**: Schematic of the DNA moving through the pore in the nanopore experimental setup and the subsequent current changes when the DNA moves through the pore. The time the current is changed is called the dwell time and how much the current is affected is called the peak amplitude. After the DNA moves through the pore, the current returns to its baseline.*

2.4 ELECTROLYTE SOLUTIONS

The electrolyte solutions used in these experiments were either 100 mM KCl or 1 M KCl in TE buffer. The pH for the vast majority of the experiments was pH 8.0, but some exploration of other values was conducted, including pHs 5.0, 7.0, 9.0, and 10.0. The pH for these solutions was adjusted using either 1 M HCl or 1 M KOH, so as not to introduce new ions. These solutions were made using ultrapure deionised water and filtered through a 0.2 µm filter (Millipore) to remove any contaminants.

2.5 ELECTRODE FABRICATION

Ag/AgCl electrodes were used in these experiments. They were chosen as they are non-polarisable and have fast charge transfer from the electrode to the solution. They are also inexpensive to manufacture, have a stable potential, and non-toxic components. They,

along with a KCl electrolyte solution, were used for the transport of ions across the pore. The non-polarisable properties of the silver chloride means that the potential difference between the electrode surface and electrolyte solutions is negligible. When a constant voltage, V_{bias}, is applied, a potential difference occurs between the two electrodes. This potential difference drives the redox reactions at the cathode and anode. These redox reactions are shown below, in **Equation 2.5** and **2.6**. It is these reactions, through their generation of Cl^- ions, that stimulate the flow of K^+ and Cl^- ions through the pore, creating the baseline current. The similar ion mobilities of K^+ and Cl^- allow the ions to move in opposite directions at similar velocities. At the anode, Ag is oxidised and at the cathode, AgCl is reduced, forming Cl^- ions.[6]

$$\mathrm{AgCl}_{(s)} + \mathrm{e}^- \rightarrow \mathrm{Ag}_{(s)} + \mathrm{Cl}^- \quad \text{Equation 2.5}$$

$$\mathrm{Ag}_{(s)} + \mathrm{Cl}^- \rightarrow \mathrm{AgCl}_{(s)} + \mathrm{e}^- \quad \text{Equation 2.6}$$

Equation 2.5 and 2.6*: The redox reactions that occur in the electrodes used in the nanopore experiments. AgCl is reduced to form Cl^- ions and Ag is oxidised to form AgCl. Silver wire is used for the electrodes. Equation 2.5 occurs at the cathode(-) and equation 2.6 occurs at the anode(+).*

These electrodes were formed by dipping 0.125 mm silver wire (Goodfellow Cambridge Ltd) connected to a battery into a 1M KCl solution to allow the working electrode to acquire chlorine ions. Each electrode was left in the solution for two minutes to ensure adequate chlorination and similar thickness. The electrodes were then washed with dI water and stored in a 1 M KCl solution to prevent degradation.

2.6 FARADAY CAGE

For all experiments there was at least one Faraday cage surrounding the nanopore set up. The Faraday cage was essential, as otherwise the noise levels would have been too high to see any events at the picoAmpere scale. It is shown in **Figure 2.6**. The Faraday cage was placed on a floating table to prevent mechanical vibrations. Additionally, both the cage and the Multiclamp amplifier had a common ground.

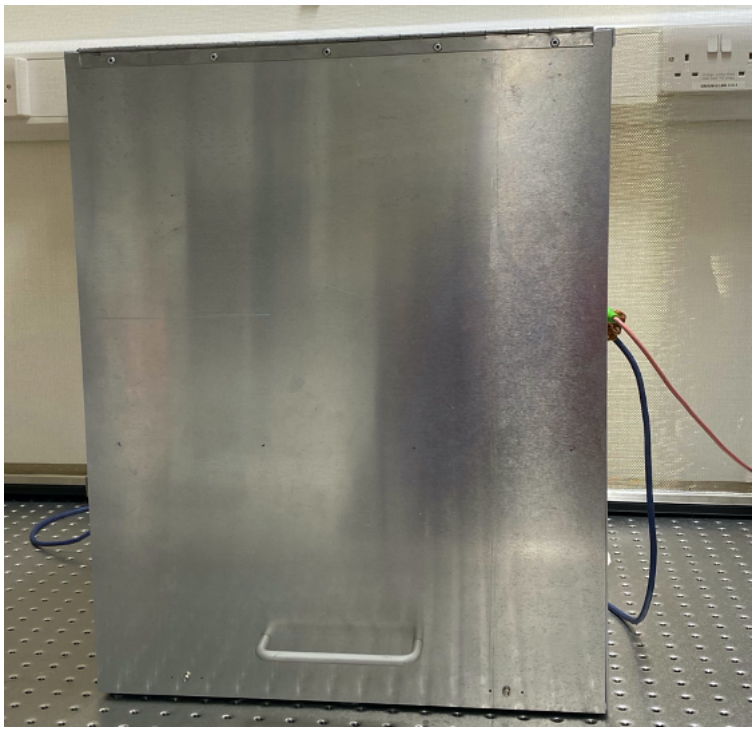


***Figure 2.6**: Faraday Cage used in nanopore experiments.*

2.7 SAMPLE DNA TRANSLOCATIONS

All of the experiments discussed in this research use DNA, which can be added either to the pipette, for cis-to-trans experiments, or to the bath, for trans-to-cis experiments. The electrolyte solution differs for these two experimental conditions. For trans-to-cis, it is necessary to use 1M KCl in TE buffer, but for cis-to-trans 100 mM KCl in TE buffer was used. Negative voltages are required to obtain translocation events for cis-to-trans and positive voltages are applied for trans-to-cis. As voltage is applied, the DNA moves towards the pore via an electrophoretic force, although electroosmotic forces may also play an important role. When the applied voltage is increased, the number of translocations also increases as the electrophoretic force is higher. However, while higher voltages result in higher signal to noise ratios (SNR), it can also lead to the pipette becoming blocked more easily; therefore, most of the experiments in this work were conducted at -300 mV or 300 mV, as this provided a good balance between high SNRs and irregular blocking.

2.7.1 CIS-TO-TRANS

Sample current traces for the cis-to-trans experiments are shown in **Figure 2.7**. Typically, these traces have a small baseline width of around 10 pA, the dwell times tend to be a bit longer than trans-to-cis and the peak amplitude a bit smaller than trans-to-cis. The capture rate compared to trans-to-cis is also higher. All of these can be seen when comparing **Figures 2.7** and **2.8**. As the experiments were conducted in 100 mM KCl in TE buffer, the DNA translocating through the pore increases the negative charge so the current increases. This current enhancement results from the counterions surrounding the negatively charged DNA, forming downward facing events.

***Figure 2.7**: Sample current trace for cis-to-trans with 10 Kbp DNA. DNA concentration is 100 pM and the experimental conditions are a 30 KHz low-pass Bessel filter, 100 KHz sampling rate, 100 mM KCl in TE pH 8.0, and -300 mV. The peak amplitude is typically around 30 pA, and so has good SNR compared to the baseline.*

2.7.2 TRANS-TO-CIS

The trans-to-cis current traces show similar downward facing events. However, as the electrolyte solution is 1 M KCl in TE buffer rather than 100 mM KCl, the DNA lowers the current trace value rather than enhancing it. As the concentration of ions is significantly increased, the additional counterions that are brought with the DNA are no longer enough to enhance the current. Additionally, the number of events for the same concentration differs greatly between both experimental setups, that is the capture rate is much lower than cis-to-trans. A sample trans-to-cis 10 Kbp DNA current trace is shown in **Figure 2.8**. The trans-to-cis generally has a slightly noisier baseline, 20 pA instead of 10 pA, but as the event peak amplitudes are much greater, this would not obscure a significant number of events.

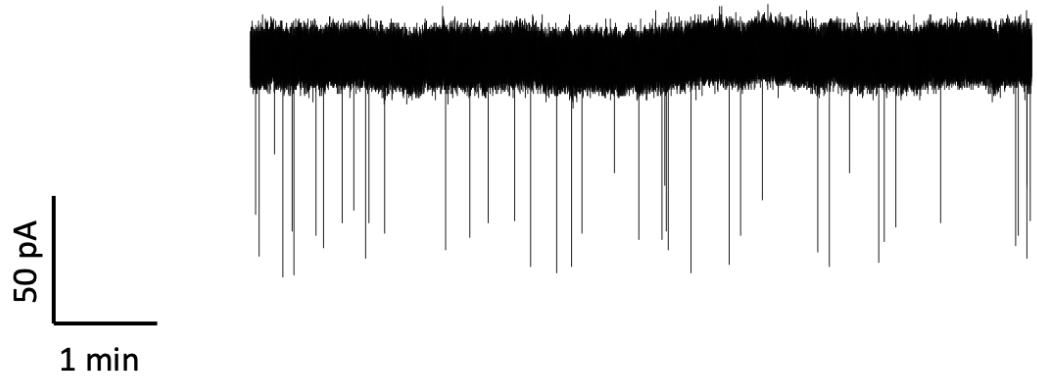


***Figure 2.8**: Sample current trace for trans-to-cis with 10 Kbp DNA. DNA concentration is 100 pM and the experimental conditions are 30 KHz low-pass Bessel filter, 100 KHz sampling rate, 1M KCl in TE buffer, and 300 mV. The peak amplitudes are around 50 pA, so there is good SNR compared to the baseline.*

DNA SAMPLES AND PROTEIN PREPARATION

2.8 OLIGONUCLEOTIDES AND PLASMID

The oligonucleotides were all obtained from Integrated DNA Technologies. They were designed to minimise the number of self-annealing sites and potential hairpins. Additionally, they all had melting temperatures that were less than the temperatures used in the hybridisation protocols. This information was obtained from an oligonucleotide properties calculator published by Northwestern University.[7] Specific sequences are shown in the relevant chapters. All of the sequences were resuspended in biological grade water and the concentration was measured using a Nanodrop 2000C. Samples were aliquoted to minimise freeze-thaw cycles and stored at -20^{o}C. The plasmid (10,201 bp in length) was purchased from ATUM. It was aliquoted to reduce freeze thaw cycles and stored at -20 ^{o}C. 10 Kbp control DNA was purchased from Thermoscientific and stored in the same conditions as the other DNA used in these experiments. In all experiments, the DNA was thawed before use and the concentration was confirmed by spectrophotometry using a Nanodrop 2000C. Additionally after carrier formation, the concentration of the DNA carriers was confirmed in the same manner.

2.9 SOLUTIONS AND REAGENTS

The chemicals used to make the solvents and buffers were Tris base (VWR Chemicals), EDTA (Sigma Aldrich), KCl (VWR Chemicals), and mono and dibasic sodium phosphate (G-Sciences). Biological water was used for all of the carrier modifications and was purchased from Thermoscientific. All other required water was deionised (Sartorius). The pH of the buffers was modified using 1 M KOH or 1 M HCl, both purchased from VWR Chemicals.

2.10 PROTEINS AND ENZYMES

Several proteins and enzymes were used throughout this research. Enzymes that were used to modify the DNA carriers, discussed in detail in further chapters, are T4 ligase (NEB), Blunt/TA ligase master mix (NEB), Nb.BbvCI (NEB), and SAL1 100,000 units/mL (NEB). Proteins that were bound to the carriers included quadrivalent streptavidin (NEB), monovalent streptavidin provided by the Howarth Laboratory at the University of Oxford [8], alkaline phosphatase (Thermoscientific), and human alpha thrombin (Cambridge Biosciences). Biotinylated BSA (Thermoscientific) was also used in some control experiments. Cutsmart buffer and ligase buffer were provided with the enzymes. All enzymes were kept on ice during use to minimise loss of activity, but their relevant buffers were thawed completely. SAL1 in particular seemed very susceptible to temperature changes, such that smaller amounts of it were ordered at a time to prevent significant loss of activity.

2.11 NANODROP

A Nanodrop 2000C was used for all the DNA concentration measurements. The nanodrop operates using two fibre optic cables; the receiving fibre embedded within the pedestal and the source fibre. These two cables are connected by the liquid sample. For all samples, 1 µL of the solvent alone was used as the blank. Before and after each sample or blank the pedestal was wiped clean. 1 µL of sample was used for each measurement and three measurements were performed for each sample. The average of these was then used to calculate the concentration.

The nanodrop measures the absorbances of the samples and then, using the Beer-Lambert equation, it is possible to calculate the concentration. A detailed discussion of

extinction coefficients for the dendrimer DNA is discussed in Chapter 3. For the plasmid experiments, the extinction coefficients did not need to be modified as the final sample was linear DNA. For this nanodrop, the ratio of 260/280 shows the ratio of DNA to RNA, where a ratio of 1.8 is pure DNA and 2.0 is pure RNA. This ratio was examined for all measured samples to ensure that the results were as expected. This detection limit is 2-15,000 ng/µL. For all the carriers, the DNA concentrations were measured before and after modification, and then these concentrations were used to both incubate the carriers with protein and the subsequent dilutions needed for experiments in the nanopore.[9]

2.12 GEL CHARACTERISATION

The main characterisation studies conducted in this research used agarose gel. The gels were made by dissolving agarose (Sigma Aldrich) in TBE (10X TBE buffer purchased from Sigma Aldrich). Gels were run using a BioRad gel electrophoresis system. The gels were primarily used to confirm that the DNA carriers had been modified correctly. Two DNA ladders were used regularly, a 50 bp ladder (NEB) and a 1 Kbp extended ladder (NEB). Each well contained both sample and gel loading dye (NEB). All the gels were stained with SYBR Gold (Life Technologies Ltd or Invitrogen). Specific agarose percentages and gel running conditions are discussed in further chapters.

2.13 PCR THERMAL CYCLER

A PCR Thermal Cycler,TC-3000 (Techne), was used for all of the hybridisation and incubation steps for the DNA carriers. It is possible to set specific programs for each modification step, and the machine was then able to hold the carrier at certain temperatures for specific amounts of time.

2.14 FILTRATION KIT AND GEL EXTRACTION KIT

For the purification steps of both carriers discussed in this research either a filtration kit, Purelink PCR Micro kit (Invitrogen), or a gel extraction kit, Monarch DNA Gel Extraction kit (NEB), were used. The PCR kit has a separation range of 0.1-12 Kbp and was able to be separate quickly; in less than an hour.[10] The gel extraction kit has a range of 50 bp - 25 Kbp,

but took much longer to run.[11] A detailed discussion of both of these purification methods is in later Chapters. The purified samples were then stored at -20 °C.

2.15 BIOANALYSER

The Agilent 2100 Bioanalyser was used for several experiments where the gel electrophoresis resolution was not great enough. An Agilent DNA 12000 kit was used to run these experiments. For all of the experiments, the gel mix was prepared and loaded into the chip. Successful loading was confirmed by the disappearance of capillary lines inside the chip. The marker, ladder, and samples were then loaded and vortexed. Only 1 µL of sample was required for each well. The protocol for all of these experiments was taken from the Agilent DNA 7500 and DNA 12000 Kit Guide.[12] The results were then analysed using the Agilent Bioanalyser software. Results where the marker or ladder did not provide the expected results were disregarded.

DATA ANALYSIS

The current signal and voltages were recorded using Clampex (Molecular Devices) for both the Axopatch and the Multiclamp. The acquired data was filtered at either 10 KHz or 30 KHz on each amplifier respectively.

Data analysis of the current traces and voltages recorded from the nanopipette experiments was performed using a custom Matlab script written by Professor Joshua Edel and Clampfit (Molecular Devices). Clampfit was used to create the I-V curves and sample trace data, while Matlab was used to analyse the translocation events directly. An overview of the general steps of this analysis is shown in **Figure 2.9**.

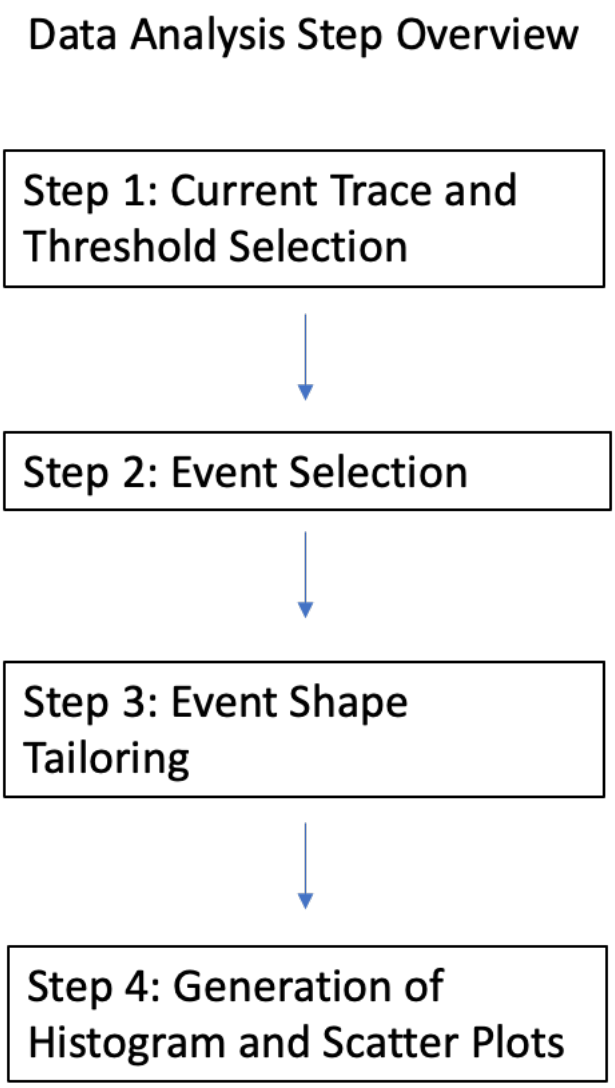


***Figure 2.9**: Flowchart showing the general steps for the data analysis using Matlab.*

The current traces were fit using the code, such that the individual events were correctly selected, shown in **Figure 2.10**. Specific step sizes for this fitting could be varied, so that the current was tracked accurately. A threshold could also be set to remove noise, such that it was possible to input a specific standard deviation and set that as the cut off for individual events above the mean noise level. The mean noise level was defined by tracking and subtracting the baseline signal. The threshold parameters could then be assessed by plotting all of the current points and fitting them with a Poisson probability distribution. A sample threshold fitting is shown in **Figure 2.10** and the accuracy of this fitting was used to adjust the threshold parameters. The threshold parameters were modified such that the script fit the events correctly and to avoid the inclusion of noise in the event count.

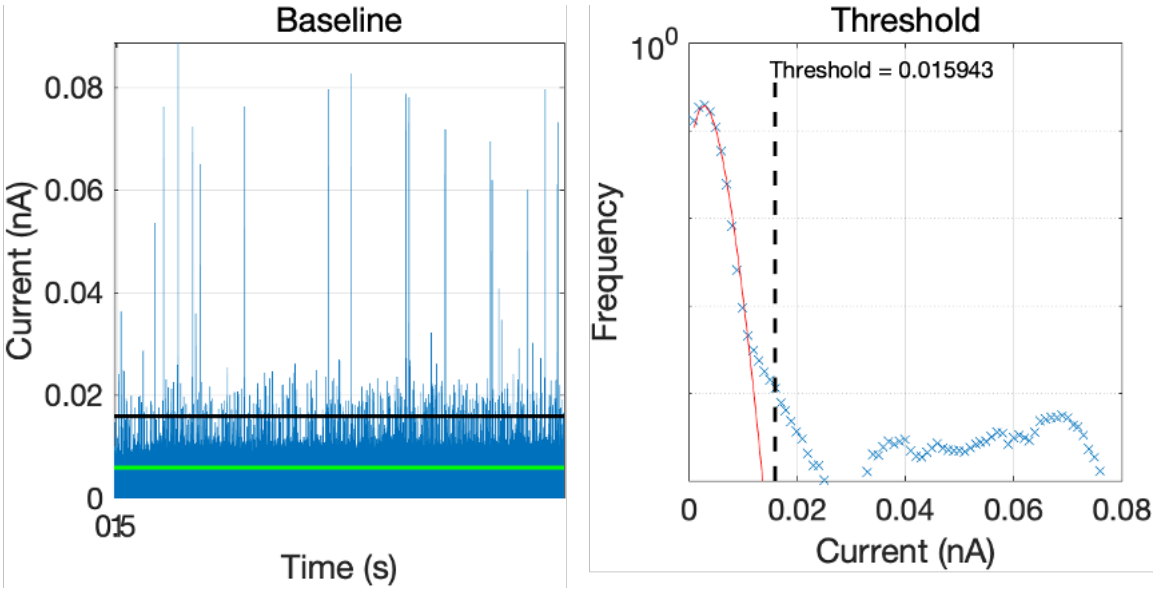


***Figure 2.10**: Sample baseline fitting and threshold histogram with Poisson probability fitting. The green line on the baseline is the step offset which is used to select where the events start, while the black line is the threshold. The Y axis on the threshold graph is logarithmic.*

The dwell time and peak amplitude were defined by subtracting the baseline. The dwell time is defined as difference in the initial time when the peak amplitude was above the baseline and when it returned to baseline levels. The peak amplitude was calculated by subtracting the baseline from the mean measured amplitude within the defined event. These events can either be blockades or enhancements, but they both have significant deviation from the baseline. **Figure 2.11** illustrates the information that can be obtained from a single translocation event.

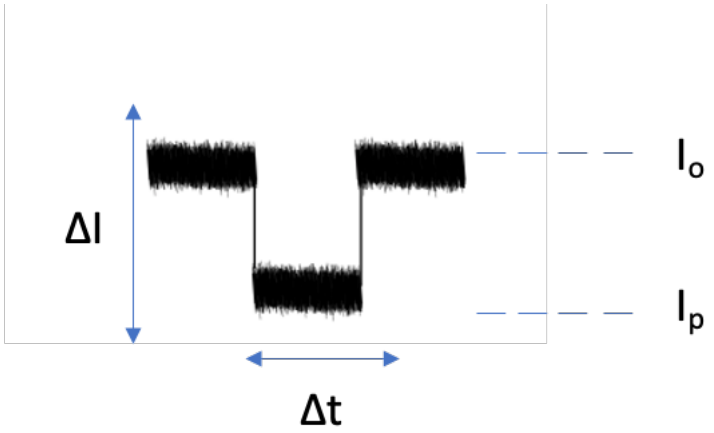


***Figure 2.11**: The information that can be obtained from a typical translocation event. I_o is the current through the open pore, I_p is the current through the blocked pore, ΔI is the change in the current, the peak amplitude, and Δt is the dwell time.*

After event selection, it is possible to look through all identified events individually. This allowed for the removal of certain events; specifically, events that were lower than 0.1 ms when recording with a 10 KHz filter were removed, as this was below the limit of detection, and events longer than 10 ms were removed, as this was considered to be a blocking of the pore. Events that were less than 0.1 ms were also removed with a 30 KHz filter as they were assumed to be DNA bumping the pore. Individual assessment of the events also ensured that the event was selected correctly and fit well. It is possible to adjust several parameters that affect event fitting including the maximum number of levels in each event and the step size used for the fitting to ensure that the event is fit well. A discussion on how this affects the selection of subpeaks is discussed in Chapter 5. However, for all data discussed in this thesis, it was necessary to ensure that the events were selected correctly.

The dwell times and peak amplitudes for all of the selected events were then plotted in a histogram. These histograms present important results for each carrier. It is possible to separate different lengths of DNA from these characteristics.[13] Sample histograms for 10 Kbp cis-to-trans are shown in **Figure 2.12**. Gaussian fitting was also applied to all of these histograms in the Matlab code to determine the mean dwell time and mean peak amplitude. Scatterplots for the dwell time vs current amplitude were also frequently used. These plots made it possible to see trends for specific events in both the dwell time and peak amplitude. The scatter plots also demonstrated how the event populations differed for samples without protein and samples with protein.

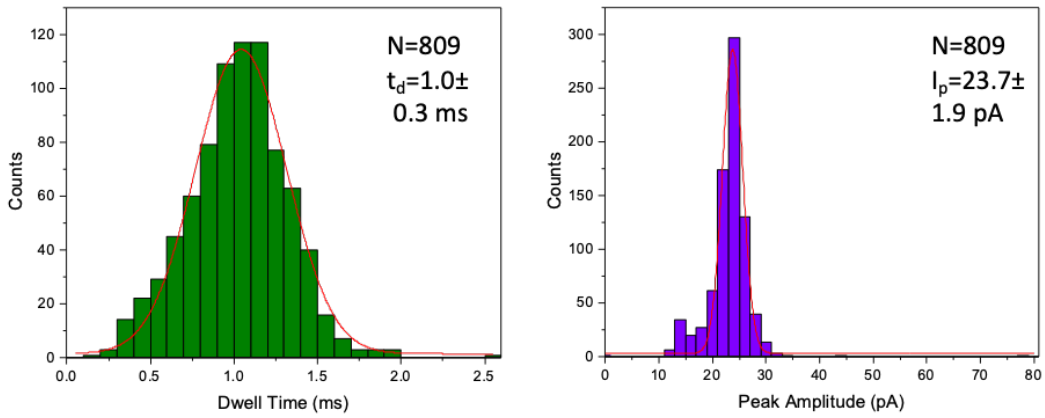


***Figure 2.12**: Sample histograms for 100 pM 10 Kbp DNA recorded at -300 mV, cis-to-trans, 100 mM KCl in TE buffer at pH 8.0, 30 KHz filter, and 100 KHz sampling rate, generated from the Matlab code directly. A Gaussian fitting is used to obtain the mean dwell time and mean peak amplitude.*

For the subpeak analysis, once the events were selected correctly, it was then possible to select for certain characteristics. These include parameters such as dwell time, peak height, subpeak dwell time, and subpeak height. These parameters and the process of automating subpeak selection are discussed in detail in Chapter 5. The automating of the subpeak selection was an important aspect of this research as otherwise all selection had to be performed manually and this, in addition to being time consuming, led to discrepancies in the consistent selection of subpeaks.

For all recordings, samples with fewer than fifty events were discarded. Nanopipette controls were also recorded in the same conditions and analysed in the same manner.

REFERENCES

1. Instruments, S. P-2000 Laser Based Micropipette Puller System Operational Manual.
2. Steinbock, L. J., Lucas, A., Otto, O. & Keyser, U. F. Voltage-driven transport of ions and DNA through nanocapillaries. *Electrophoresis* **33**, 3480–3487 (2012).
3. Gong, X. *et al.* Label-free in-flow detection of single DNA molecules using glass nanopipettes. *Anal. Chem.* **86**, 835–841 (2014).
4. Ying, Y. & Sze, J. Selective Single Molecule Sensing in Nanopore. (2017).
5. Pedone, D., Firnkes, M. & Rant, U. Data analysis of translocation events in nanopore experiments. *Anal. Chem.* **81**, 9689–9694 (2009).
6. Wanunu, M. Nanopores: A journey towards DNA sequencing. *Phys. Life Rev.* **9**, 125–158 (2012).
7. Buehler, Eugene, Cao, Qing, Kibbe, Warren A., Lurie, R. H. Oligo Calc: Oligonucleotide Properties Calculator. *Northwestern.edu* (2022). Available at: http://biotools.nubic.northwestern.edu/OligoCalc.html.
8. Howarth, M. *et al.* A monovalent streptavidin with a single femtomolar biotin binding site. *Nat. Methods* **3**, 267–273 (2006).
9. Thermo-Scientific. NanoDrop 2000 / 2000c Spectrophotometer. (2000).
10. Invitrogen / PureLink™ PCR. PureLink ™ PCR Micro Kit. (2009).
11. NEB. Monarch DNA Gel Extraction Kit.
12. Agilent. DNA 7500 and DNA 12000 Guide.
13. Li, J., Gershow, M., Stein, D., Brandin, E. & Golovchenko, J. A. DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mater.* **2**, 611–615 (2003).

Chapter 3 Dendrimer Carriers

Table of Contents

Synopsis: This chapter discusses the use of DNA dendrimers as DNA carriers in nanopore sensing. Dendrimer formation and their previous uses are examined. Several types of dendrimers, each with different lengths for their respective sticky ends, are investigated. The examined dendrimers included 4 base pair overhangs, 13 base pair overhangs, and 30 base pair overhangs. Some nanopore data for various dendrimer structures is also shown, including data for structures incubated with thrombin. However, it is shown that these structures are not stable enough to be successful DNA carriers.

BACKGROUND

3.1 PREVIOUS USES OF DENDRIMERS

DNA dendrimers have been used in a variety of research settings for a range of different functions, including bioimaging and specific target detection.[1,2] They are formed with one central Y structure with three others Ys attached to form the first generation of the dendrimers. These Y structures bind together using short single stranded sections called sticky ends or overhangs. Then further Y structures can be added in a stoichiometric fashion to form subsequent generations. In previous research, the DNA is hybridised both together with ligase[3–5] or without.[2,6] Typically, the dendrimers with longer sticky ends do not require ligation for successful hybridisation. For example, the dendrimers with a four base overhang designed by Luo *et al.* (2003) require ligation, but ones with a thirteen base overhang designed by Zhang *et al.* (2015), do not require ligation.[2,5] Dendrimers with very short sticky ends, such as ones that are four bases in length, are not stable at room temperature.[7] While the dendrimers can vary in size significantly, the first generation typically ranges from 10-20 nm, and the second from 20-35 nm.[2,8] When the sticky ends are long enough, dendrimers typically have excellent stability, and can be combined easily with sequences of interest, such as aptamers or fluorescently-labelled structures. They are also produced in a straightforward manner with high levels of purity.[2] While dendrimers have been used in many different ways, they have not been used as a DNA carrier in nanopore sensing.

3.2 WHY DENDRIMERS

Dendrimers are DNA structures that are made by combining Y-shaped DNA in a stoichiometric fashion. Each Y-shape is made of three different oligonucleotides hybridised together, with short single stranded sticky ends at the end of each Y branch. In order to

avoid self-hybridisation of the Ys, each sticky end must not be palindromic, so that the dendrimers can only form in one direction.[5] Two different sized generations, G1 and G2, were made. The first generation dendrimer structure (G1) is formed by combining four Y-shapes, Y0 in the centre surrounded by three Y1s, while the larger second generation (G2) is a G1 surrounded by six Y2s. This shape allows for a significant number of binding sites with a relatively small DNA structure; G1 allows for three and G2 allows for six. Therefore, multiplexed sensing should be possible. Dendrimers, as they are made from DNA, can be modified with existing enzymes and can function well in the physiological environment. They also are inherently biocompatible and can be modified to target specific molecules.[9] The easy customisability, the large numbers of binding sites, and the theoretical ease to form these structures are why this shape was explored as a DNA carrier.

3.3 AIM OF THE CARRIER

The DNA dendrimer is a particularly useful shape as a DNA carrier for several reasons. It is completely customisable, and thus can have different aptamers or other binding mechanisms attached easily. Additionally, several of these can be attached at one time, so multiplexed sensing can be performed. The size of the dendrimer also is large enough that it can be seen as it passes through the pore, but not so large that it often blocks the pore. The specific aim of this research is to show that the dendrimer can be used as a DNA carrier. That is, it can bind to proteins specifically and these dendrimer and protein events are able to be identified from dendrimer only events.

DENDRIMER FORMATION

3.4 DENDRIMER STRUCTURE- 4 BASE STICKY ENDS

The first dendrimer structure attempted used a four base overhang and was modelled on the one made by Luo *et al*. (2003). These form dendrimers with base pair sizes of 38 bp for the Y0, 164 bp for the G1, and 440 bp for the G2, using only double stranded base pairs for the sizing. Three four base pair length overhang oligonucleotides, Y0a, Y0b, Y0c, were hybridised together to form a Y0 structure. The oligonucleotides were heated to 95°C for 2 min and then cooled to 65°C and incubated for 5 min. Then, they were annealed

at 60^{o}C for 2 min, and further annealing at 60^{o}C for 0.5 min was performed. The annealing steps were repeated 40 times. Afterwards the mixture was stored at 4^{o}C. Each oligonucleotide was dissolved in annealing buffer, 10 mM Tris 1 mM EDTA 50 mM NaCl (pH 8.0). Y1a, Y1b, and Y1c were combined using the same hybridisation process to form Y1. The sequences for these oligonucleotides are shown in **Table 3.1**. **Figure 3.1** shows a diagram of the formation of the dendrimers.

Y0a 5'- CGAT T GGA TCC GCA TGA CAT TCG CCG TAA G-3'
Y0b 5'- CGAT CTT ACG GCG AAT G AC CGA ATC AGC CT-3'
Y0C 5'- GTCA AGG CTG ATT CGG T TC ATG CGG ATC CA-3'

Y1a 5'- TGAC T GGA TCC GCA TGA CAT TCG CCG TAA G-3
Y1b 5'- TGAC CTT ACG GCG AAT G AC CGA ATC AGC CT-3'
Y1c 5'- TGAC AGG CTG ATT CGG T TC ATG CGG ATC CA-3'

Table 3.1*: Schematic showing the oligonucleotide sequences for the 4 base sticky ended oligonucleotides. The oligonucleotides were incubated together to form Y structures with 4 base sticky ends at the terminus of each Y branch. These Y structures were then bound together to form a first generation dendrimer.*

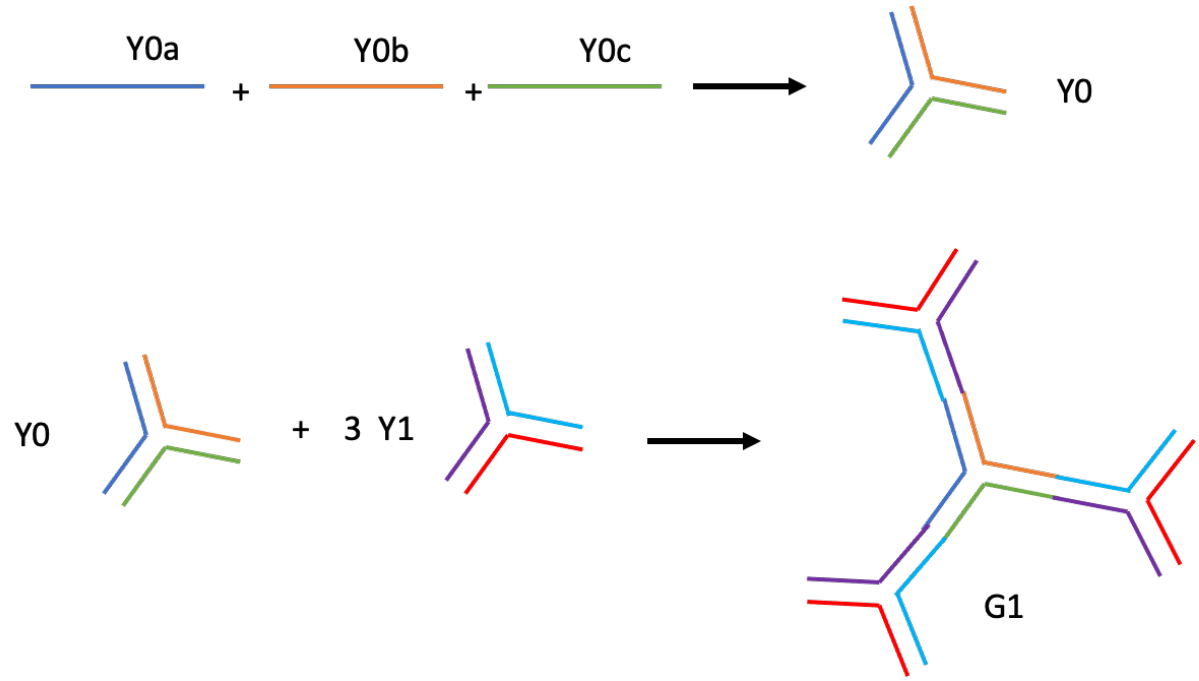


***Figure 3.1**: Diagram of the formation of the first generation dendrimers with a four base overhang.*

At first the oligonucleotides were combined at a lower concentration than described by Luo *et al.* (2003), at 50 μM. They were combined in two different methods, one pot or Y structure method. In the one pot method, all of the oligonucleotides are combined in the correct stoichiometric ratios, such that there are three times as many Y1s as Y0s. The oligonucleotides are hybridised together and the mixture is incubated for two hours at room temperature. In the Y structure method, the oligonucleotides are hybridised separately so Y0 and Y1 are formed. The Ys are then combined in a ratio of 3:1 and then incubated for two hours at room temperature. The results of these two methods are shown in **Figure 3.2**, and do not seem to have a vast difference, as the G1, the first generation of dendrimer, did not form well for either. The bands for the Ys are similar to the bands for the G1s suggesting that the G1s are not forming.

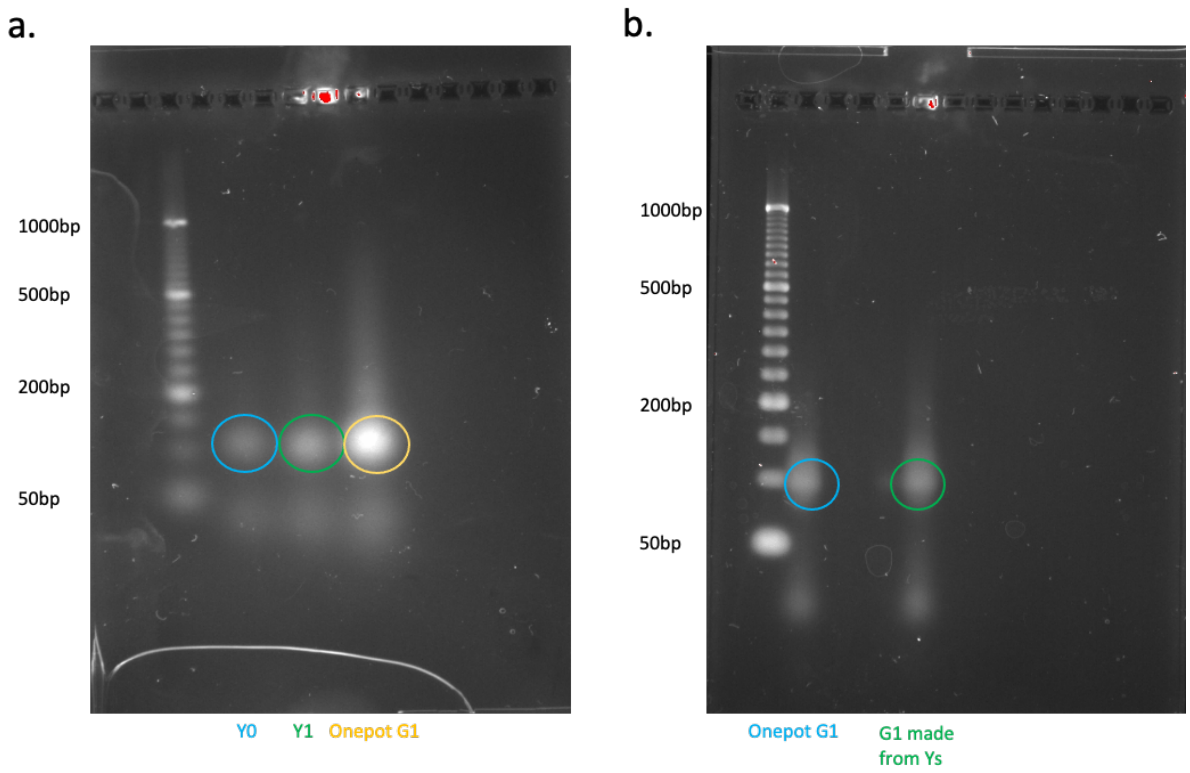


***Figure 3.2**: 3% agarose gel with G1s and Ys, where the gels were run for 2 hrs at 75 V on ice. In gel A, Y0 is shown in well 1 and Y1 is shown in well 2.. The G1 was made either by combining all of the oligonucleotides in a one pot formation, well 3 in gel A and well 1 in gel B, where they are added all together in the correct stoichiometric ratio and then hybridised and incubated together. While the other G1 was made by combining Y structures that had been made from oligonucleotides hybridised together, well 2 in gel B. The circles mark the bands of interest.*

As the G1 did not form very well using the previously discussed methods, the exact protocol, including higher concentrations and ligase, from Luo *et al.* (2003), was performed. New oligonucleotides were used so that the higher concentration of 5 mM for each oligonucleotide could be achieved. There were some issues with replicating this protocol with the exact concentrations, as they were so high the oligonucleotides would not dissolve in the solution; this made combining the oligonucleotides in the correct ratios difficult. T4 DNA ligase was used and the Ys were incubated with it at room temperature for 2 hours. **Figure 3.3** shows the results of this protocol. G1x is the first generation made with the one pot method and G1 with combining the Y structure method. Neither method seemed to work very well, as the bands that show up for the G1s are the same as the ones that show up for the Ys. Thus, dendrimers with longer sticky ends, 13 bases, were explored.

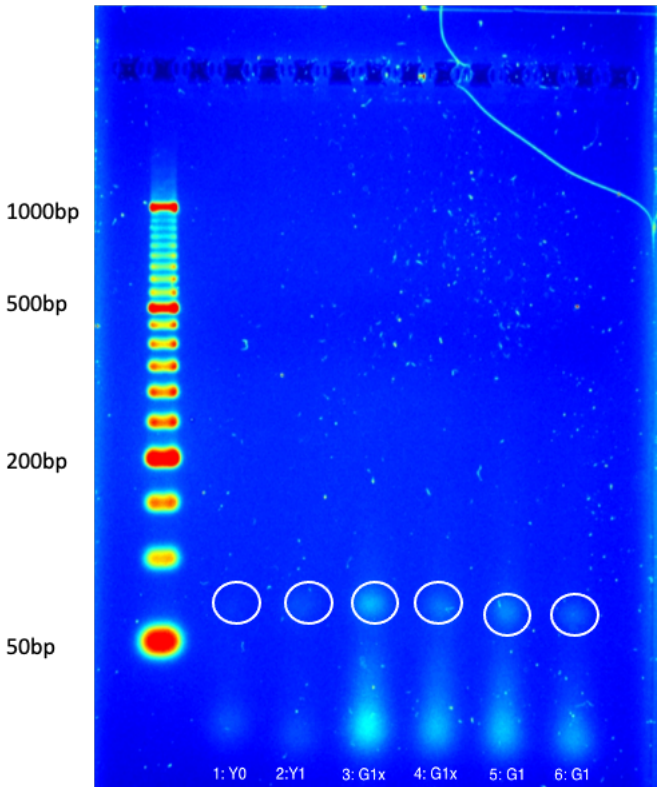


***Figure 3.3**: 3% agarose gel for the dendrimer structures. The gel was run for 2 hrs at 75 V on ice. The first generation dendrimers were formed with in one pot (G1), wells 5 and 6, or by combining the Ys (G1x), wells 3 and 4. The Ys, well 1 and 2, were ligated together and each oligonucleotide only has a four base overhang. The circles mark the bands of interest.*

3.5 DENDRIMER FORMATION- 13 BASE STICKY ENDS

After the dendrimers with 4 base long sticky ends were determined to not be stable enough, 13 base sticky ends modelled on ones used by Zhang *et al.* (2015) were used. The oligonucleotides were combined in the same way as with the 4 base sticky ended ones to form the Y structures and then the G1s, however some aspects of the protocol were different. Each oligonucleotide, purchased from IDT was resuspended in dI water at a concentration of 100 µM. They were diluted to 20 µM in a phosphate buffer, 50 mM phosphate 100mM Na^+ pH 8.0. 10 µL of each oligonucleotide was combined for a total reaction volume of 30 µL for each Y, i.e. Y0a, Y0b, and Y0c were combined to form Y0. The Y structures were hybridised by heating the mixture to 95°C for 2 min and then cooling them to 4°C at a rate of 1°C per min. Y0 and Y1 were then incubated together in a ratio of 1:3

respectively at room temperature for 2 hours. The oligonucleotide sequences are shown in **Table 3.2**. Additionally, for the 4 base sticky ended dendrimers, only a 15 base long thrombin aptamer was explored, but for the 13 base one a longer 29 base thrombin aptamer was used, as this had a much higher binding rate. The 29 base thrombin aptamer (a2) has a K_d of 0.5 nmol/L, while the K_d of the 15 base aptamer (a1) is 10-150 nmol/L.[10]

Y0a	**5'- GAC CGA TGG ATG A CCT GTC TGC CTA ATG TGC GTC GTA AG-3'**
Y0b	**5'- GAC CGA TGG ATG A CTT ACG ACG CAC AAG GAG ATC ATG AG-3'**
Y0c	**5'- GAC CGA TGG ATG A CTC ATG ATC TCC TTT AGG CAG ACA GG-3'**
Y1a	**5'- GAA GCC ACT CTG A CC TGT CTG CCT AAT GTG CGT CGT AAG-3'**
Y1b	**5'- GAA GCC ACT CTG A CT TAC GAC GCA CAA GGA GAT CAT GAG-3'**
Y1c	**5'- TCA TCC ATC GGT C CT CAT GAT CTC CTT TAG GCA GAC AGG-3'**
Y2a	**5'- GAC ACA CTG AGG T CC TGT CTG CCT AAT GTG CGT CGT AAG-3'**
Y2b	**5'- GAC ACA CTG AGG T CT TAC GAC GCA CAA GGA GAT CAT GAG-3'**
Y2c	**5'- TCA GAG TGG CTT C CT CAT GAT CTC CTT TAG GCA GAC AGG-3'**
Y1aapt	**5'- GGT TGG TGT GGT TGG CCT GTC TGC CTA ATG TGC GTC GTA AG-3'**
29b	**5'- AGT CCG TGG TAG GGC AGG TTG GGG TGA CT C CTG TCT GCC TAA TGT GCG TCG TAA G-3'**

***Table 3.2**: Oligonucleotide sequences for the 13 base sticky ended oligonucleotides. The Ys were formed in a similar manner to the previous shorter oligonucleotides, but the sticky ends that bind to form the dendrimers were much longer at 13 bases.*

Schematics of how these dendrimer structures are formed and what they look like are shown in **Figure 3.4**. Both the longer and shorter thrombin aptamers are depicted. Additionally, they demonstrate how multiple proteins could bind to each dendrimer and how the larger generations of dendrimers are formed. **Figure 3.4** shows up to the second generation, G2, which is able to bind six thrombins at a time. While the schematics only show the 2D rendition of the structures, in 3D they have a spherical shape[2], so it is unlikely that there could be significant steric effects that could hinder binding for multiple thrombins.

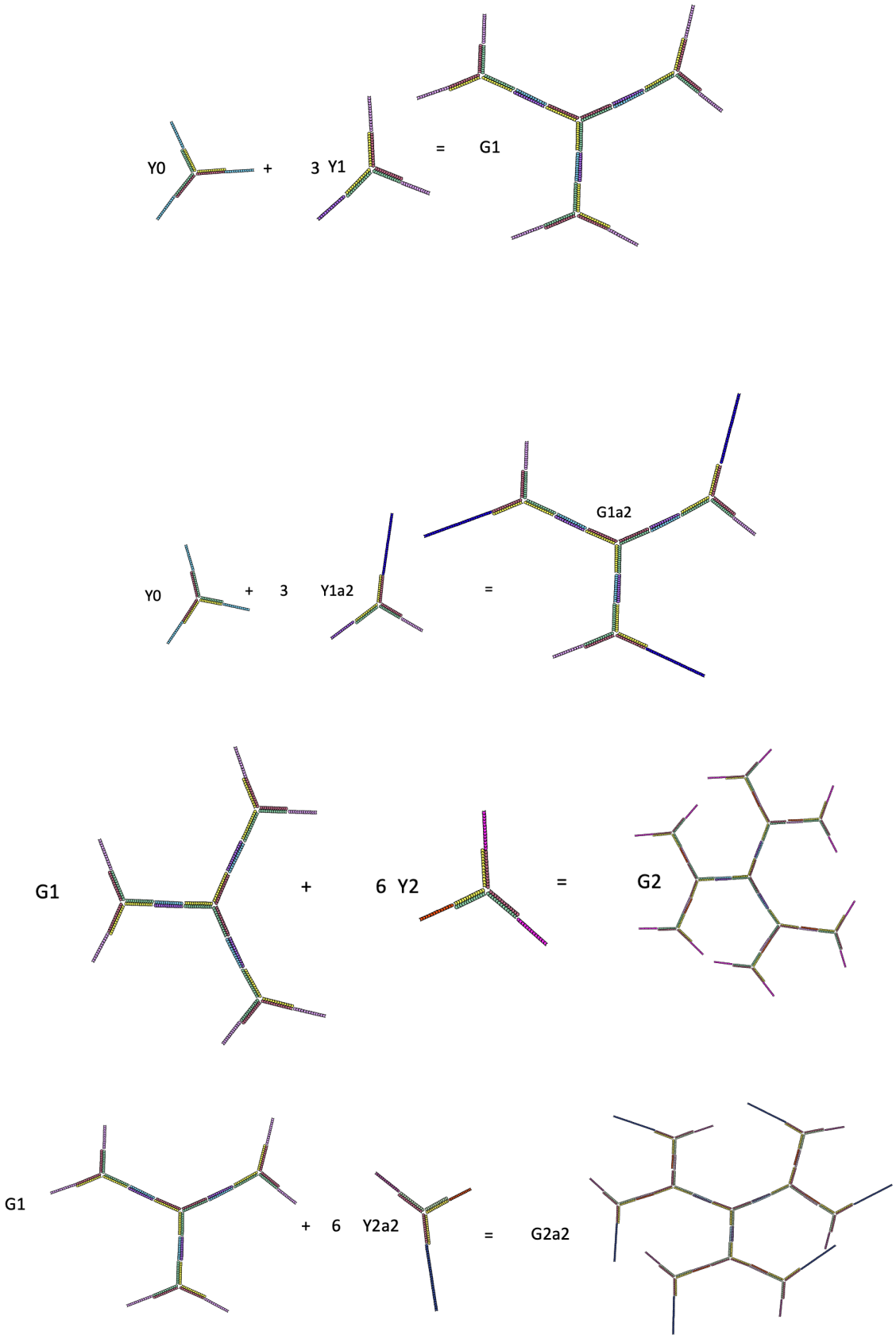


***Figure 3.4**: Schematic of the formation of the first and second generation dendrimers with a 13 base overhang. The corresponding stochiometric ratios for the Y structures are depicted*

as well. The 29 base thrombin aptamer is shown as attached in G1a2, where each first generation structure has three aptamer binding sites. It is shown in G2a2 as well, where each second generation structure has six aptamer binding sites.

The first generation with the 13 bases was made by combining the Y structures, not the one pot method, and without the addition of ligase. The hybridisation protocol was the same as previously described, and even without ligase, the formation of the G1 was much better than with the 4 base sticky ends. The concentration of each oligonucleotide was 20 µM so the previously described issue of the oligonucleotides not dissolving properly was not a problem for this protocol. The successful formation is shown in **Figure 3.5**, where there is a clear difference between the bands for the Y and for the G1. The lower band in the circle could be G1 that does not fully form. Thus, it seems likely that the issue with the earlier dendrimer attempts was the sticky end length and not the protocol.

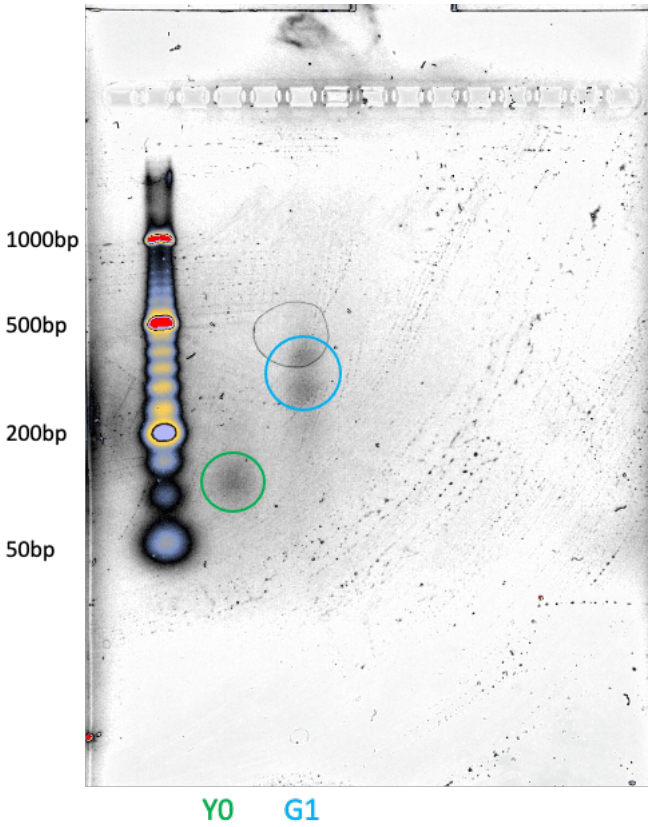


***Figure 3.5**: 3% agarose gel showing Y0, well 1, and G1, well 2, where the gel was run for 2 hrs at 75 V on ice. The circles mark the band where the structure is expected to be.*

Once it was shown that the G1 was capable of forming with the 13 base overhang, the protocol needed to be optimised. The G1 still did not seem to have a high rate of

completion, although it was clear that the Y structures were forming, so the hybridisation of the Ys was examined. The hybridisation time was varied from 2 hours. The Ys were hybridised for 1, 2, and 3 hours and also overnight. However, as shown in **Figure 3.6**, 1 hour was clearly not enough time, and if left for longer, it seemed like it was possible for the structure to fall apart. As the band for 2 hours was still the strongest, even though all the Ys did not form into G1, it was deemed to most efficient way of forming the G1s.

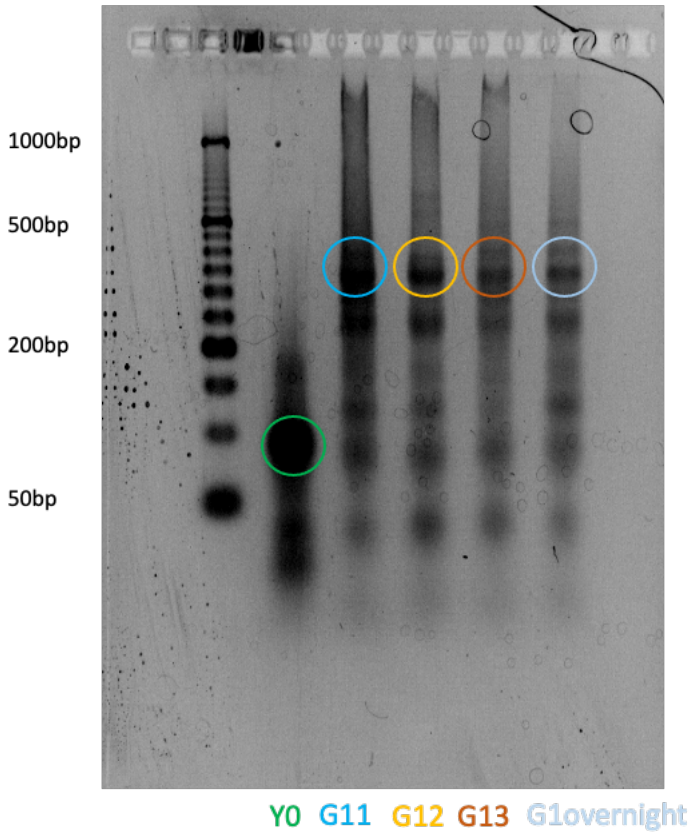


Figure 3.6: *3% agarose gel for a comparison of the length of incubation time for the Y structures after hybridisation was varied for the first generation. The gel was run for 2 hrs at 75 V on ice. Y0 is well 1. G11, well 2, was incubated for one hour, G12, well 3, for 2 hrs, G13, well 4, for 3 hrs, and G1overnight, well 5, was incubated overnight. The circles mark the band where the structure is expected to be.*

As the G1 was able to be formed successfully, the next steps were making the G2 and the dendrimers with the aptamers. Two aptamers for thrombin were used, the 15 base one and the 29 base one. The Y and G structures with the 15 base aptamer are referred to as Y1a1 and G1a1 respectively and for the 29 base one Y1a2 and G1a2. This naming convention is applied to the second generation dendrimers as well. The same protocol was used to make the Y structures and then the various dendrimer generations. For the first

generation with aptamers, the Y1s were made with one branch, one oligonucleotide, with the aptamer. Therefore each Y1 had one aptamer so the first generation would be able to bind to three thrombins at once. For the second generation, the G1 was combined with six Y2s. The Y2s could again have aptamers added, where each one had one aptamer, such that the G2 with the aptamers was able to bind to six thrombins at once. The flexibility in tailoring the Y structures is an essential positive feature of the dendrimers. **Figure 3.7** shows the gel results for both G1 and G2 with aptamers. Both generations were able to form and while it is not possible to tell from the gel if the aptamer was incorporated correctly, as the dendrimer formed, it was assumed that the aptamer was added successfully. Both of the gels shown demonstrate that the dendrimers do not form completely, as there are other bands, but it does seem that a significant percentage of them do form the dendrimer correctly. The Ys are estimated to be 78 bp, while G1 is 351 bp, and G2 819 bp. The sizes in base pairs only consider the double stranded portions of the structures. The diameters are about 13 nm for G1 and 22 nm for G2.[2] The degradation in the structures and the multiple bands is likely to be caused by the heat generated while the gel is running. Although, the dilution of the sample required for the gel electrophoresis could also lead to dissociation of the Y-structures.

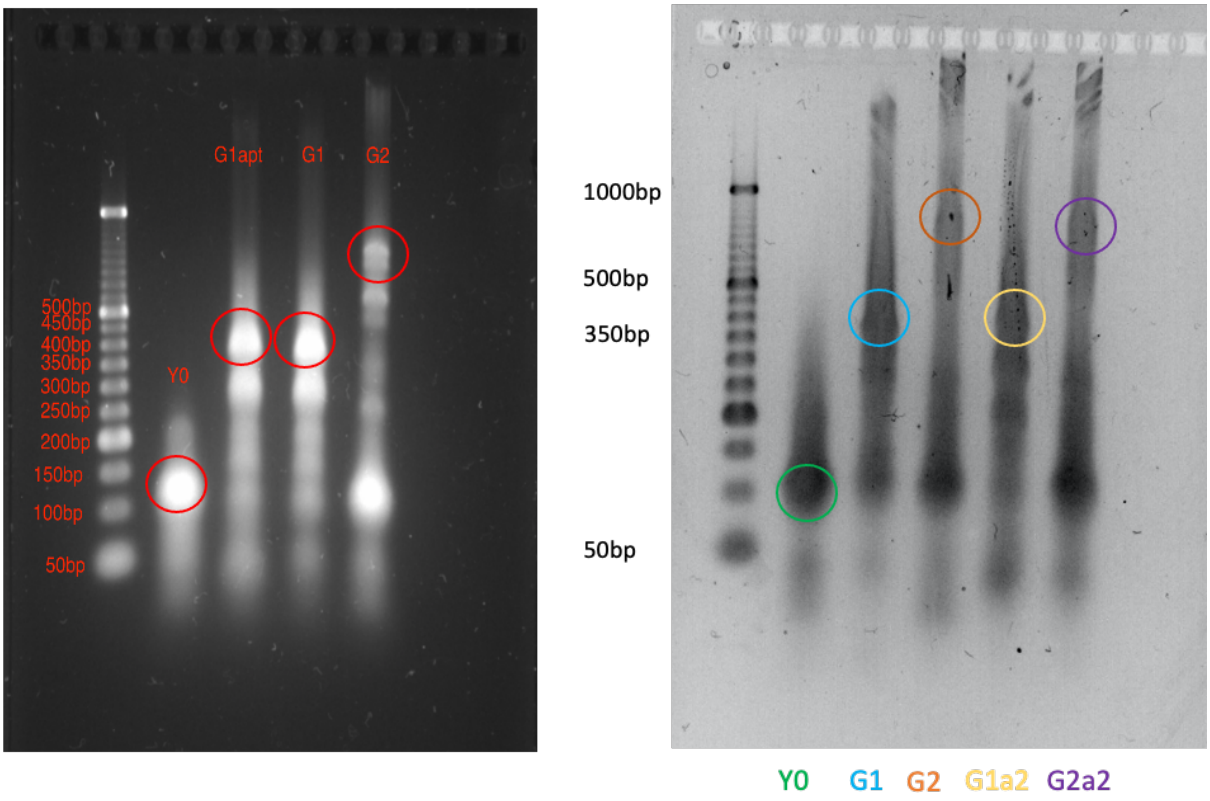


***Figure 3.7**: 3% agarose gels comparing the formation of the different dendrimer generations with the 13 base overhang. The gels were run for 2 hrs at 75 V on ice. For the gel on the left, the 15 base thrombin aptamer was added to one of the oligonucleotides in the Y1 structure. In well 2 and 3, the Y0 to Y1 ratio was 1:3, while in wells 4 and 5, the Y0 to Y2 ratio was 1:6. The formation of G2 is also shown. The circles mark the band where the structure is expected to be.*

In order to see if the completion rate could be increased, the concentrations of the oligonucleotides and the subsequent structures were examined. If the concentrations of each oligonucleotide were significantly different, this could lead to a lower completion rate for certain Y structures, as well as a significant amount of leftover oligonucleotide that could inhibit hybridisation. Therefore, some exploration of equimolar hybridisation versus following the protocol of the Zhang *et al*. (2015) was performed; as the oligonucleotides were all assumed to be the same concentration, the same volume was added for each, however it is possible that their concentration varied slightly. In the equimolar version the exact concentrations of each oligonucleotide were used, such that the exact same number of moles for each oligonucleotide was added to make the Y structures. The results of this are shown in **Figure 3.8**. However, the equimolar procedure has been abandoned due to the extra calculation needed and time required, as it did not provide significantly better results.

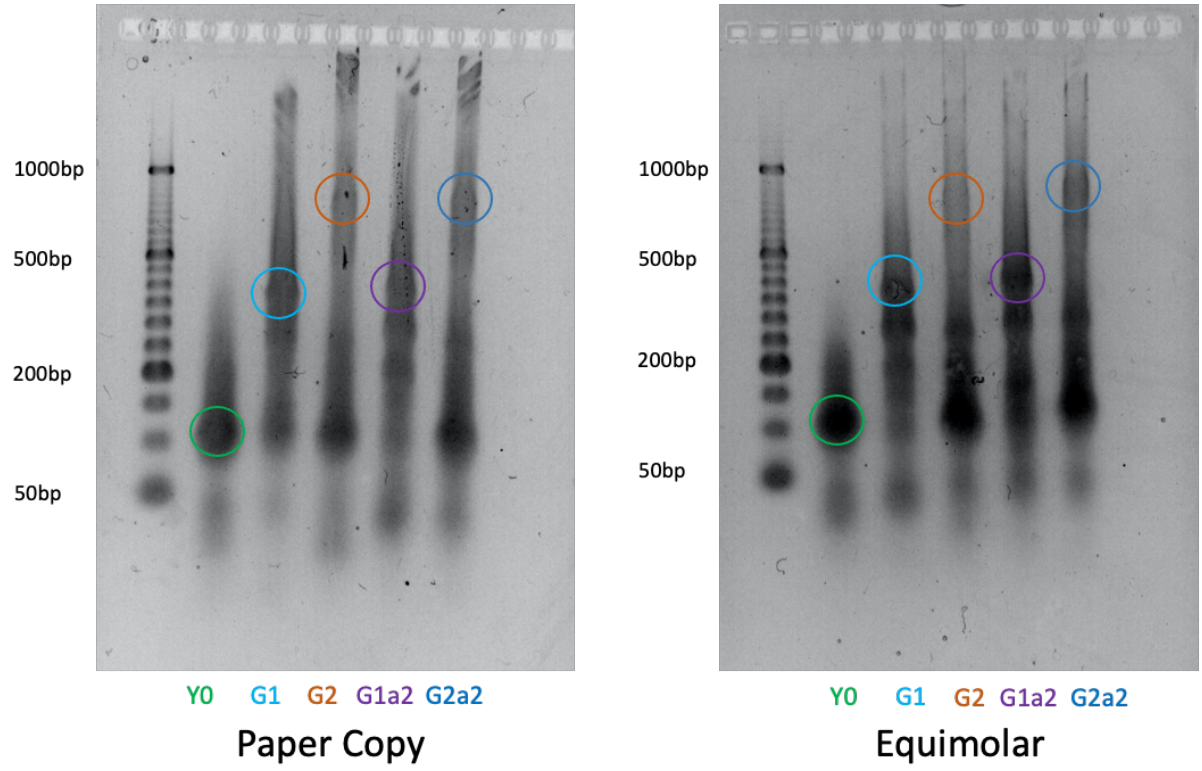


***Figure 3.8**: 3% agarose gel comparions of the paper protocol to the equimolar protocol. The gel was run for 2 hrs at 75 V on ice. The circles mark the band where the structure is expected to be.*

There was some difficulty in determining the concentration of the larger structures, the Ys and the Gs. The concentrations for the all the structures were calculated using a Nanodrop 2000c spectrophotometer. This instrument measures the absorbance and then, as long as the extinction coefficient and pathlength are known, the concentration can be calculated using Beer's law. While the extinction coefficients of the oligonucleotides were known, they were provided by IDT, calculating the extinction coefficients of the non-linear structures was more complex. They were calculated using an extinction coefficient calculator[11] and several assumptions. As the Nanodrop 2000C spectrophotometer measures the asborbance at 260 nm, the extinction coeffients were calculated at 260 nm. In the Y structures, the extinction coefficients were calculated by assuming that only the double-stranded DNA was important and that it was linear. That is, by taking the total double-stranded sequence from all three of the arms of the Y. A similar process was applied for the larger structures as well. Y0 for the dendrimers was considered totally double stranded and for G2 so was Y1. For G1, the Y1 portions were included three times as there are three Y1s in

each G1. For each G2, the Y1 portions were included three times and the double-stranded sections of Y2 were included six times. The extinction coefficients for all of these structures are shown in **Table 3.3**. The concentrations calculated using these extinction coefficients were close to the expected values, so this method of calculting extinction coefficients was deemed appropiate.

Oligomer/Dendrimer Structure	**Extinction Coefficient at 260 nm (L/(mol*cm))**
Y0a	373,000
Y0b	396,200
Y0c	380,000
Y1a	364,400
Y1b	387,600
Y1c	363,500
Y2a	369,500
Y2b	392,700
Y2c	367,300
Y1aapt	381,200
29b	522,900
Y0	640,551
Y1	640,551
Y2	640,551
Y1a1 (aptamer 1, shorter)	640,551
Y1a2(aptamer 2, longer)	640,551
G1	3,892,677
G1a1	3,892,677
G1a2	3,892,677
G2	10,283,645
G2a1	10,283,645
G2a2	10,283,645

Table 3.3*: This table shows the estimated extinction coefficients for each oligonucleotide, Y, and dendrimer structure. The oligonucleotide extinction coefficients were taken from the manufacturer, Integrated DNA technologies (IDT). For the Y and dendrimer structures, the extinction coefficient was calculated with an assumption of linearity and only considering double-stranded sections of the structure, thus excluding the overhangs. As the resulting concentrations were close to the expected value, the assumptions were considered valid.*

3.5 EXPLORATION OF SALT BINDING

As all of the nanopore experiments were conducted in salt conditions, several stability studies were conducted. Y0 and G1 were diluted in 100 mM KCl in TE buffer (pH 8.0) and left for several days. The results of this study are shown in **Figure 3.9**. This gel

suggests that the structures are relatively stable in the buffer, as there is not significant decay after several days incubation in it at 4°C.

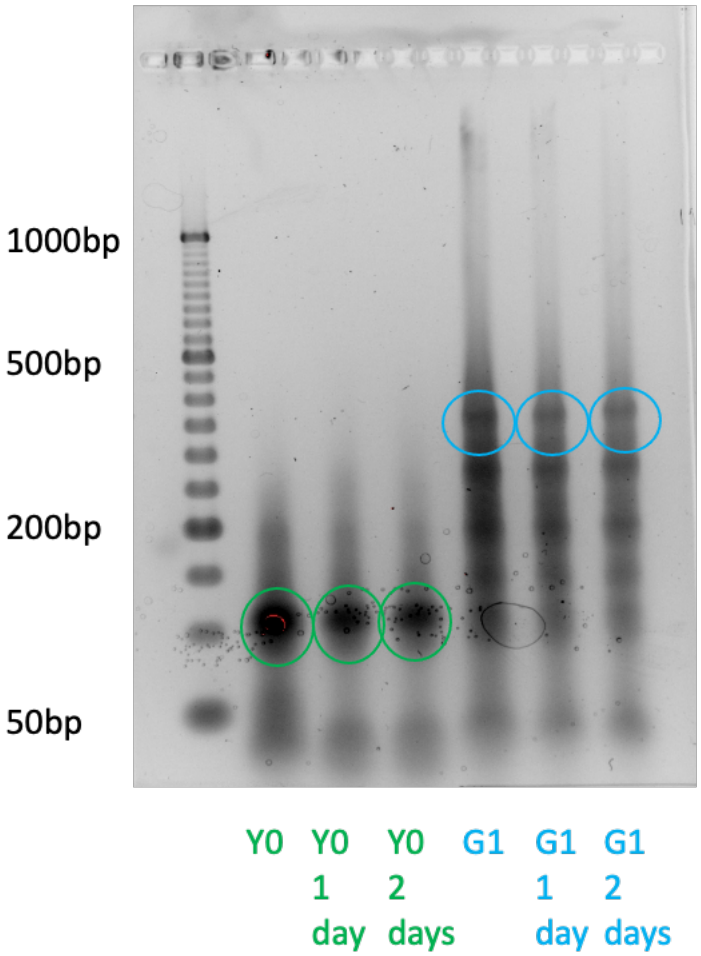


***Figure 3.9**: A stability study run in a 3% agarose gel showing that when kept at 4°C the Y structures and G1 structures are relatively stable for several days in 100 mM KCl in TE pH 8.0. The gel was run for 2 hrs at 75 V on ice. The G1 had been incubated for two hours when made. Wells 1-3 are the Y0 structures. Wells 4-6 are the G1s. The circles mark the band where the structure is expected to be.*

The structures were then incubated with various salt concentrations to see if they would be stable over a range of conditions. Each was incubated with TE buffer pH 8.0 with KCl concentrations ranging from 0-1 M. The structures incubated were Y0, G1a2, and G2. The results of this study are shown in **Figure 3.10**. It does seem like with the addition of salt, there is some degradation back to the Y structures for both G1a2 and G2, as the band for the Ys is brighter with the addition of salt. However, the Y0 does not seem so affected. It is also not possible to tell the exact concentration that starts the degradation as it seems that any concentration has a bit of reversion back to the Y structures. As such, it seems necessary

to perform the nanopore experiments in 100 mM KCl, rather than a higher salt concentration.

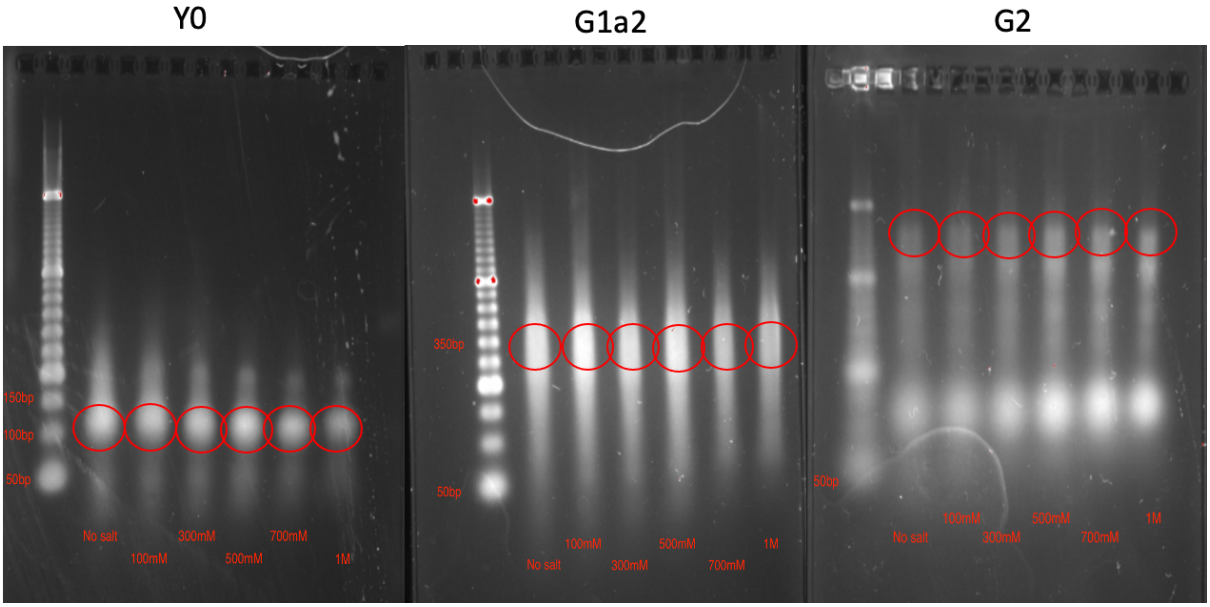


***Figure 3.10**: Stability study of the dendrimer and Y structures in various salt concentration in TE buffer pH 8.0 in 3% agarose gel. The gel was run for 2 hrs at 75 V on ice. For all three of the gels, well 1 is the structure in a solution without salt, well 2 is the structure in a solution at 100 mM salt, well 3 is the structure in a solution at 300 mM salt, well 4 is the structure in a solution at 500 mM salt, well 5 is the structure in a solution at 700 mM salt, and well 6 is the structure in a solution at 1 M salt. The circles mark the band where the structure is expected to be. Measurements of the intensity for all three gels at the bands of interest show some degradation of all structures in salt solutions, although at 100 mM this is minimal.*

3.6 PURFICATION METHODS AND PROTEIN BINDING

As there were many leftover oligonucleotides after the Y structure formation, it was necessary to remove them. An abundance of single stranded oligonucleotides could possibly affect the nanopore experiments. A micro PCR filtration kit was chosen as it was able to remove smaller DNA fragments without affecting the larger dendrimer structures. The filtration kit also was not as time consuming as gel extraction and worked similarly well. **Figure 3.11** shows the results of this filtration. Unfortunately, as very few gels were run for this experiment, the results are not especially clean. However, it is possible to see that when comparing both the G2 and G2 filtered (G2f) and G1a1 and G1a1 filtered (G1a1f) that there are no more or far fewer leftover oligonucleotides. The band near the top of the gel for both

filtered samples also suggests that the filtration was able to work, but that the sample run in this gel was too concentrated to move through it easily. It is possible to see some light bands around the length expected for G1a1f. Therefore, it was concluded that filtering removed the oligonucleotides effectively and was a necessary purification step.

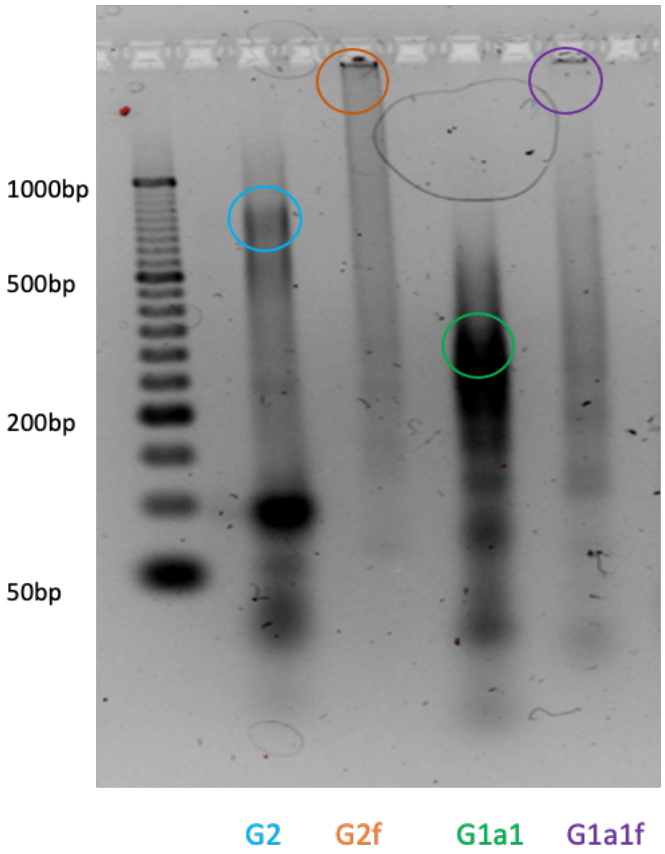


***Figure 3.11**: 3% agarose gel showing the results of filtering with a micro PCR filtration kit. The gel was run for 2 hrs at 75 V on ice. Well 1 is G2, well 2 G2 filtered, well 3 G1a1, and well 4 G1a1 filtered. The circles mark the band of interest for each structure.*

As these dendrimers had a thrombin aptamer added, it was also necessary to confirm that the thrombin was able to bind successfully to the dendrimer before running the nanopore experiments. Several ratios of dendrimer to thrombin were explored, 1:3, 1:6, and 1:9. The first generation dendrimer G1a2, so the sample with the more effective aptamer, was used. Each G1a2 had three aptamers to which the thrombin could bind. The results of this control are shown in **Figure 3.12**. In the gel, there is an additional dark band around 1000 bp, significantly higher than the dendrimer without protein, for all three ratios, suggesting that the thrombin is able to bind to the dendrimer.

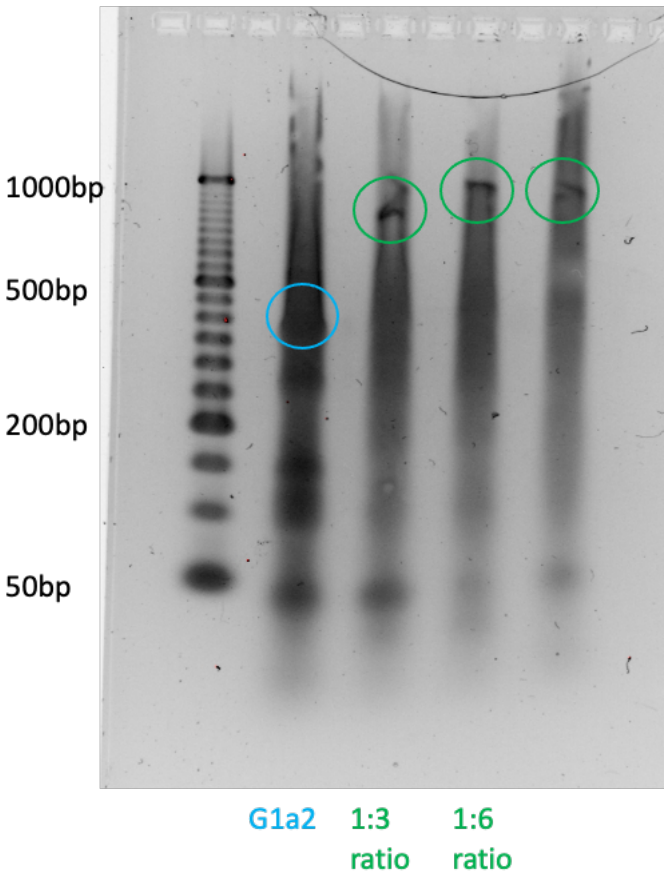


***Figure 3.12**: 3% agarose gel showing the successful binding of thrombin to the first generation dendrimer G1a2. The gel was run for 2 hrs at 75 V on ice. Well 1 is G1a2, well 2 G1a2 with a 1:3 ratio of dendrimer to thrombin, well 3 G1a2 with a 1:3 ratio of dendrimer to thrombin. The circles mark the band of interest for each structure.*

3.7 SWITCH TO LONGER OVERHANGS

Prior to full investigation of the 13 base overhang dendrimers in the nanopore, it was determined that at the low concentration needed for the nanopore, around 300-600 pM, the 13 base overhang would not be long enough to form stable structures. That is, the K_d for the 13 base overhang is on the picomolar scale so a stronger bond would be needed to form the dendrimers. Some initial nanopore results for the 13 base overhang will be shown in the next section.

As the 13 base overhang was determined to not be long enough, 30 base overhang oligonucleotides were used. The sequences for these oligonucleotides are shown in **Table 3.4**. The protocol, including purification by filtration and the incubation times were not changed for these longer overhangs. The hybridisation temperature was still high enough to break any bonds between the sticky ends or oligonucleotides. The only change was the

longer sticky ends, as they had a lower dissociation constant. These dendrimers should be stable at the low picomolar concentrations required for nanopore experiments.

Y0a 5'- GAC CGA TGG ATG TTG ACG GAA AAT GGC GCT CCT GTC TGC CTA ATG TGC GTC GTA AG-3'
Y0b 5'- GAC CGA TGG ATG TTG ACG GAA AAT GGC GCT CTT ACG ACG CAC AAG GAG ATC ATG AG-3'
Y0c 5'- GAC CGA TGG ATG TTG ACG GAA AAT GGC GCT CTC ATG ATC TCC TTT AGG CAG ACA GG-3'

Y1a 5'- GTT GCC ACT CTG ACC AAT GCC TAC CAA CGT CCT GTC TGC CTA ATG TGC GTC GTA AG-3'
Y1b 5'- GTT GCC ACT CTG ACC AAT GCC TAC CAA CGT CTT ACG ACG CAC AAG GAG ATC ATG AG-3'
Y1c 5'- AGC GCC ATT TTC CGT CAA CAT CCA TCG GTC CT CAT GAT CTC CTT TAG GCA GAC AGG-3'

Y2c 5'-ACG TTG GTA GGC ATT GGT CAG AGT GGC AAC CT CAT GAT CTC CTT TAG GCA GAC AGG-3'

Table 3.4*: New 30 base overhang oligonucleotide sequences. These are the oligonucleotides that were needed in order to ensure that any sticky end bound to form G1 and G2 was 30 bases long. For Y2, only the Y2c oligonucleotide binds to the G1 sticky ends so it was not necessary to have longer sticky ends for the other two oligonucleotides. Additionally, the oligonucleotides with thrombin aptamers were not able to bind to other dendrimer structures either, so it was not necessary to change these oligonucleotides.*

The longer sticky ends did seem to help with increasing the amount of each dendrimer produced through the protocol, particularly for G2. As shown in **Figure 3.13**, even though there still is a lot of leftover oligonucleotide, the G2 structures have much clearer bands than compared to the 13 base overhangs. The G1s also have far fewer leftover oligonucleotides. This gel was run without filtering out the oligonucleotides and the same hybridisation protocol as previously mentioned. The Y structures were incubated at room temperature for 2 hours to form each dendrimer structure.

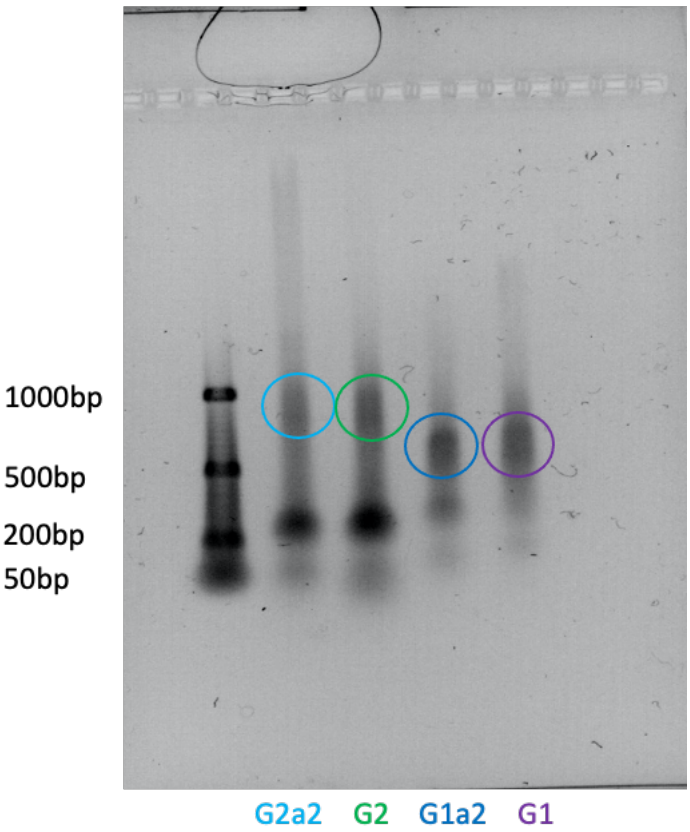


***Figure 3.13**: 3% agarose gel for the 30 base overhang structures. The gel was run for 1.5 hrs at 75 V on ice. Well 1 is G2a2, well 2 G2, well 3 G1a2, and well 4 G1. The circles mark the band of interest for each structure.*

NANOPORE DENDRIMER RESULTS

3.8 Y STRUCTURE RESULTS

As a control for the nanopore results, in addition to the buffer alone which yielded no events, the Y0 structures were run on the nanopore for both the 30 base overhang and the 13 base overhang. **Figure 3.14** shows some sample traces from these experiments. The 13 bp Y0 is in maroon and the 30 bp in green. The concentrations of both are 300 pM and the experimental conditions for the recordings are 100 KHz sampling rate, 10 KHz filter, and 100 mM KCl TE pH 8.0, -300 mV. The dendrimers were inserted into the pipette tip, the cis side, and flowed through the pore into the trans side. The average pipette diameter was around 15 nm. The lack of events suggests that the Y structure is too small to show up in the pipettes, and as such with the larger structures, it is unlikely that the Y structures would be present in the current traces.

***Figure 3.14**: Sample traces for Y0 for both 13 base pair overhangs and 30 base pair overhangs oligonucleotides. The 13 base Y is shown in maroon and the 30 base in green. The recording conditions are cis-to-trans, 100 KHz sampling rate, 10 KHz filter, -300 mV, and 100 mM KCl, pH 8.0.*

3.9 1ST GENERATION RESULTS

The G1 for both length sticky ends were run on the nanopore as well. The same conditions were used as for the Y structures, although the concentration of the 30 base overhang was 600 pM rather than 300 pM. Sample traces are shown in **Figure 3.15**. The 13 base sticky end is maroon and the 30 base sticky end is green. The sample traces do not show many events despite the relatively high concentrations of the G1. Additionally, many of the events were very short, quite near 0.1 µs which is the limit of the recording. The shortness of the events might stem from the G1s not being stable enough, so there are some that are only partially formed or once they are in the salt conditions, as shown in the previous salt studies, many of the dendrimers simply fall apart. For the 13 base overhang, the background is quite noisy, and this could stem from left over Y structures or partially formed Gs blocking the pore, as the shorter sticky end structures are even more likely to fall apart. These two factors lead to very few events.

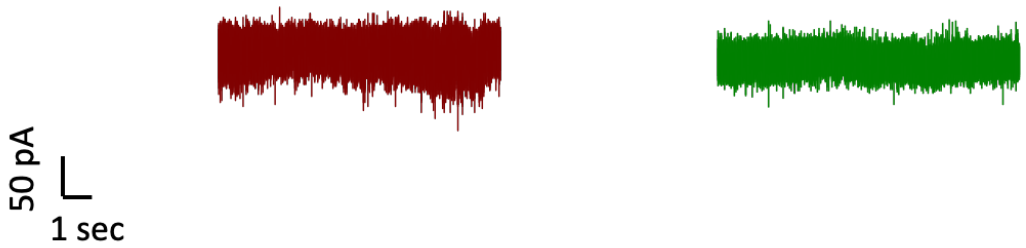


***Figure 3.15**: Sample traces for G1 for both the 13 base sticky ends and the 30 base. The 13 base is maroon and the 30 base green. The recording conditions were 100 mM KCl in TE pH 8.0, 100 KHz sampling rate, 10 KHz filter, cis-to-trans, -300 mV, and 300 pM for 13 base and 600 pM for 30 base.*

There were similar results for the first generation with the longer thrombin aptamer, G1a2. Sample traces with similar recording conditions as for the G1 and the Ys were used. However, the concentration for both the 13 base and 30 base was 300 pM. This data is shown in **Figure 3.16**. The sample traces again did not have many events and were similarly noisy. Additionally, the dwell time skews heavily towards the lower limit of the recording. This and the low number of events suggests that the structure is not forming stably.

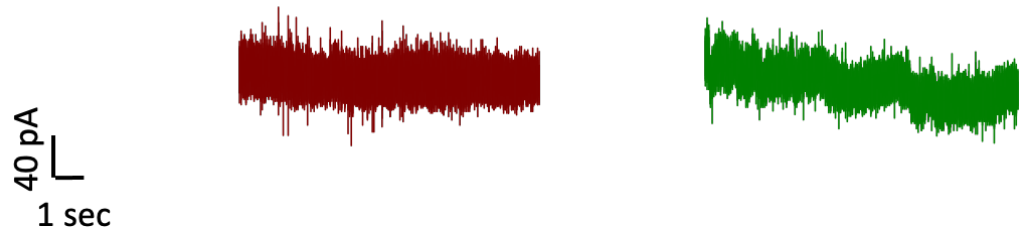


***Figure 3.16**: Sample traces for G1a2 for both the 13 base sticky ends and the 30 base. The 13 base is maroon and the 30 base green. The recording conditions were 100 mM KCl in TE pH 8.0, 100 KHz sampling rate, 10 KHz filter, cis-to-trans, -300 mV, and 300 pM.*

3.10 2ND GENERATION RESULTS

Second generation results were also recorded on the nanopore. The same conditions as used for the Ys and the G1s were used, although the concentration for both G2s, 13 and 30 base overhangs, were 600 pM. The sample traces are shown in **Figure 3.17**. The 13 base is maroon and the 30 base is green. There are not many events in the current traces and the background noise level is not very different compared to the G1s. Additionally, the dwell times are also very low for both samples. They are quite similar to the G1 length, although it would be expected that they would have longer dwell times due to their larger size. This could be explained by the structure not being very stable for both sticky end lengths, as they would fall apart to the smaller sizes; thus having low dwell times that were similar to the smaller structures.

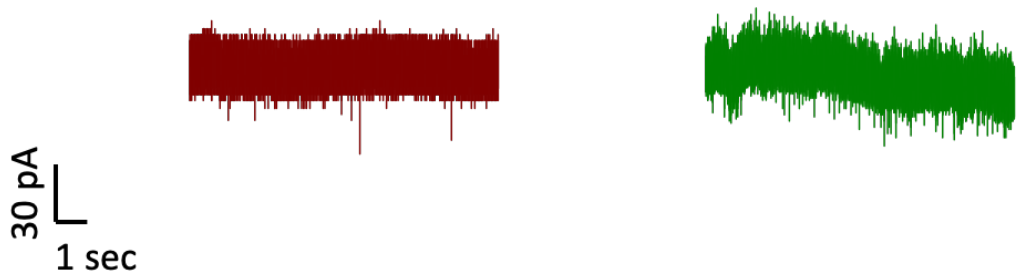


***Figure 3.17**: Sample traces for G2 for both the 13 base sticky ends and the 30 base. The 13 base is maroon and the 30 base green. The recording conditions were 100 mM KCl in TE pH 8.0, 100 KHz sampling rate, 10 KHz filter, cis-to-trans, -300 mV, and 600 pM.*

The second generation with the longer thrombin aptamer, G2a2, had even fewer events. The recording conditions were the same as for the Ys, although for G2a2 with 13 base overhangs, the sample trace was recorded at -200 mV instead of -300 mV. Also the concentrations of both samples was 600 pM. The current traces are shown in **Figure 3.18**, with the 13 base in maroon and the 30 base in green. For the 30 base, which typically had more events, the dwell time still is very short, and definitely not significantly longer than the dwell time of the G1s. All of this suggests that the G2a2s are not stable and therefore not able to be recorded in the nanopore successfully.

***Figure 3.18**: Sample traces for G2a2 for both the 13 base sticky ends and the 30 base. The 13 base is maroon and the 30 base green. The recording conditions were 100 mM KCl in TE pH 8.0, 100 KHz sampling rate, 10 KHz filter, cis-to-trans, -200 mV for the 13 base and -300 mV for the 30 base, and 600 pM.*

3.11 BINDING TO THROMBIN RESULTS

Some experiments with thrombin were also performed for both the 13 base and 30 base first generation dendrimers with the longer thrombin aptamer. The thrombin was incubated with the dendrimers for 45 min in phosphate buffer at room temperature. Sample current traces for these experiments are shown in **Figure 3.19**. The experimental conditions were cis-to-trans 100 KHz sampling rate, 10 KHz filter, 100 mM KCl TE pH 8.0, and -300 mV. The thrombin was recorded alone at 600 pM and no events appeared over a ten second period. For the 13 bp and 30 bp, the dendrimers were 600 pM and then incubated with thrombin in a ratio of 10:1. Neither of these two recordings had many events, even though there was less background noise than in some of the other recordings. This lack of events is reflective of the recordings of the dendrimers with thrombin for both the 13 base and 30 base ones. Several concentrations, including 1X, 30X, 50X, 70X, and 90X, and voltages were attempted for both G1a2s, but there were not traces with any significant numbers of events. As such, it could be concluded that the dendrimers are not stable enough and therefore fall apart even when the aptamers are bound to the thrombin.

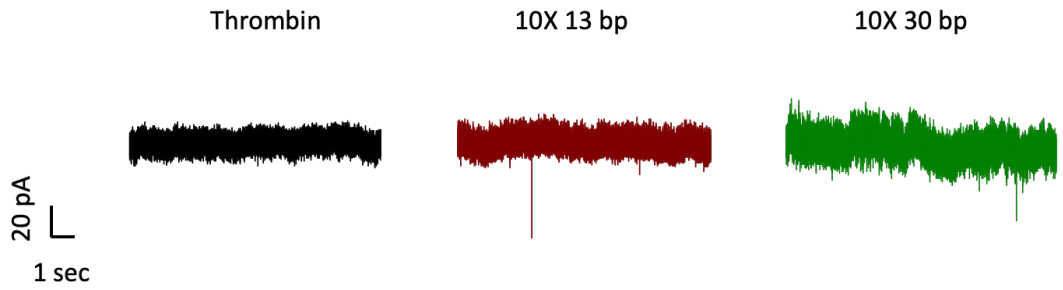


***Figure 3.19**: Sample traces for thrombin, 10:1 thrombin to G1a2 with 13 bases, and 10:1 thrombin to G1a2 with 30 bases. The recording conditions were cis-to-trans 100 KHz sampling rate, 10 KHz filter, 100 mM KCl TE pH 8.0, and -300 mV. The thrombin alone had a concentration of 600 pM and the other dendrimer samples had a dendrimer concentration of 600 pM.*

3.12 CONCLUSION: WHY DENDRIMERS ARE NOT EFFECTIVE CARRIERS

Unfortunately, while it seems it is possible to make these dendrimers, particularly the G1s, it does not seem like they are stable enough to be recorded in the nanopore. The 13 base sticky end dendrimers likely had problems as they have a binding constant that meant that at the low concentrations required by the nanopore, the structures were not formed in significant numbers. For the 30 base sticky ends, the salt conditions, would not allow the G1s and G2s to remain fully formed. This also affected the 13 base structures. Further for the 30 base sticky ends, they were unable to remain fully formed despite having a higher melting temperature for the structures so they are mostly fully formed at picomolar concentrations. As such, this led to a lack of events and significant background noise, probably caused by these structures falling apart. Therefore, it is reasonable to conclude that these DNA dendrimers do not perform well as DNA carriers for nanopores.

REFERENCES

1. El-Mahdy, A. F. M., Shibata, T., Kabashima, T. & Kai, M. Dendrimer-like polymeric DNAs as chemiluminescence probes for amplified detection of telomere DNA on a solid-phase membrane. *Chem. Commun.* **50**, 859–61 (2014).
2. Zhang, H. *et al.* A controllable aptamer-based self-assembled DNA dendrimer for high affinity targeting, bioimaging and drug delivery. *Sci. Rep.* **5**, 1–8 (2015).
3. Um, S. H., Lee, J. B., Kwon, S. Y., Li, Y. & Luo, D. Dendrimer-like DNA-based fluorescence nanobarcodes. *Nat. Protoc.* (2006). doi:10.1038/nprot.2006.141
4. Freedman, K. O. *et al.* Diffusion of single star-branched dendrimer-like DNA. *J. Phys. Chem. B* **109**, 9839–9842 (2005).
5. Yougen, L. I. *et al.* Controlled assembly of dendrimer-like DNA. *Nat. Mater.* **3**, 38–42 (2004).
6. Mohri, K. *et al.* Self-Assembling DNA Dendrimer for Effective Delivery of Immunostimulatory CpG DNA to Immune Cells. *Biomacromolecules* **16**, 1095–1101 (2015).
7. Ch??vez-Guti??rrez, L. *et al.* The mechanism of ??-Secretase dysfunction in familial Alzheimer disease. *EMBO J.* **31**, 2261–2274 (2012).
8. Qu, Y. *et al.* Self-Assembled DNA Dendrimer Nanoparticle for Efficient Delivery of Immunostimulatory CpG Motifs. *ACS Appl. Mater. Interfaces* (2017). doi:10.1021/acsami.7b05890
9. Ying, Y. L., Wang, H. Y., Sutherland, T. C. & Long, Y. T. Monitoring of an ATP-binding aptamer and its conformational changes using an α-hemolysin nanopore. *Small* **7**, 87–94 (2011).
10. Xu, Y., Yang, L., Ye, X., He, P. & Fang, Y. An aptamer-based protein biosensor by detecting the amplified impedance signal. *Electroanalysis* **18**, 1449–1456 (2006).
11. Buehler, Eugene, Cao, Qing, Kibbe, Warren A., Lurie, R. H. Oligo Calc: Oligonucleotide Properties Calculator. *Northwestern.edu* (2022). Available at: http://biotools.nubic.northwestern.edu/OligoCalc.html.

Chapter 4 Plasmid Carrier Formation

Table of Contents

Synopsis: This chapter outlines why the plasmid was chosen as a carrier and the various adjustments that were done to the protocol to modify the plasmid. The typical nanopore behaviour of the plasmid was also shown, as well as a comparison to 10 Kbp behaviour. The effect of pH, both as a gradient and by varying the pH of both the sample and the bath was discussed. It also demonstrates how confirmation of the successful modification of the plasmid and incorporation of the oligonucleotides was established.

BACKGROUND

4.1 WHY PLASMIDS

Plasmids have great potential as DNA carriers for several reasons. They are very customisable, can be expressed on a relatively large scale, and are obtainable in a fully formed format. Additionally, as they are double-stranded DNA, they are very stable and are easily modified, due to the plethora of enzymes available for the manipulation of double-stranded DNA. As such, they can be combined easily with aptamers and/or fluorescent-labelled structures. Also, it is possible to utilise the same plasmid backbone and switch out aptamers so that various proteins can be bound. Finally, due to their length, the one used in this research is 10 Kbp and is shown in **Figure 4.1**, they have the potential to bind multiple proteins with the one carrier and therefore could be used for multiplexed sensing.

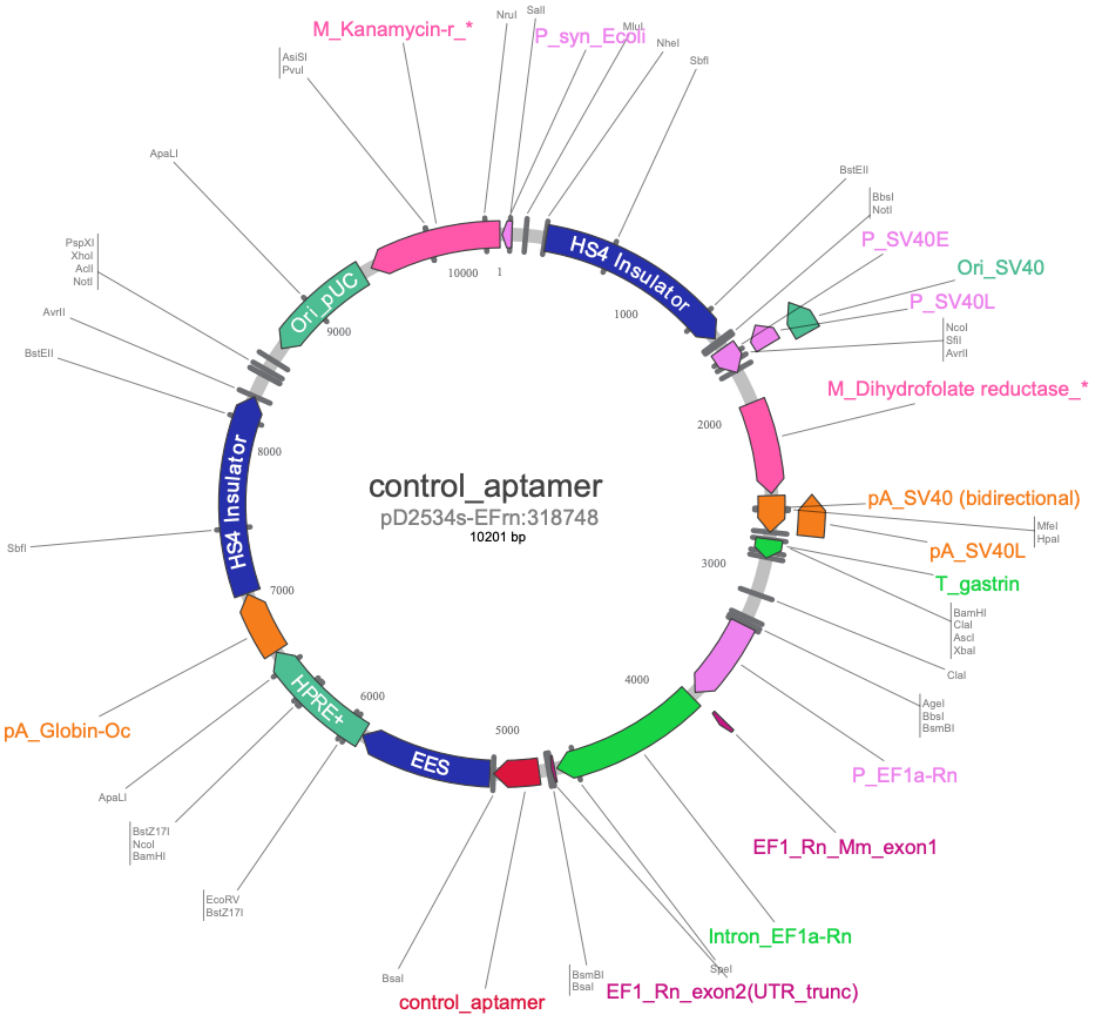


***Figure 4.1**: 10 Kbp plasmid used in this research, purchased from Atum and designed by Dr. Paolo Cadinu.*

4.2 PREVIOUS EXAMPLES OF PLASMIDS

While many types of double-stranded DNA have been used as DNA carriers in nanopore sensing, to the best knowledge of the experimenter, a plasmid has never been used as a DNA carrier. Compared to some other types of DNA carriers previously used, a plasmid requires less modification and offers the potential for more multiplexed sensing.

4.3 AIM OF THIS CARRIER

The aim of utilising a plasmid for this research is to demonstrate that plasmids can be modified so they can function as DNA carriers in a nanopore sensor and bind specifically to target protein molecules. This is demonstrated by the distinctive signals for the plasmid carrier with and without attached protein. Lastly, the carrier's potential for multiplexed sensing is also shown.

FORMATION OF PLASMID STRUCTURE

4.4 PLASMID FORMATION METHOD

The 10 Kbp plasmid was purchased from Atum and designed by Dr. Paolo Cadinu. It was designed so that a short single-stranded section, 66 bases, could be cut out using the restriction enzyme Nb.BbvCI. Additionally, it has a cleavage site for the restriction enzyme, SAL1, which is able to linearise the plasmid. The process of forming the DNA carrier from the plasmid requires several steps: step one is nicking, step two hybridisation, step three ligation, step four linearisation/cutting followed by purification, and step five is the incubation with the protein. An overview of what happens to the plasmid in each step is shown in **Figure 4.2**. While these are the steps used in the final protocol for the plasmid, these modifications were an area that was actively explored and the protocol did vary quite a bit.

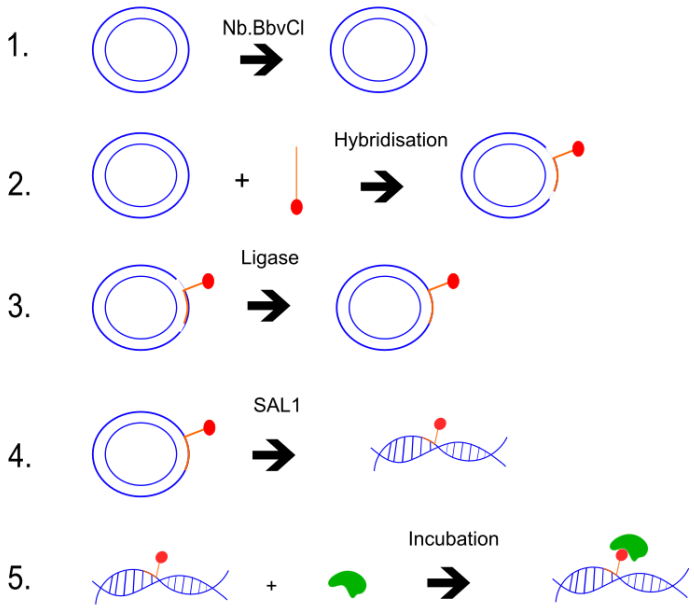


***Figure 4.2**: Overview of the modification of the plasmid. Step one shows the nicking by Nb.BbvCI. Step two shows the competitive hybridisation. Step three the ligation. Step four the linearisation and purification. The final step, five, is the incubation with the protein.*

The following is the final protocol for the plasmid. The steps are 1. nicking by Nb.BbvCI., 2. hybridisation, 3. Ligation, 4. the linearisation and purification, and 5. Incubation with protein. For step one, Nb.BbvCI (NEB) nicks one strand of the plasmid in two places so that the strand can be replaced via competitive hybridisation. 0.5 µL of the plasmid (~326 nM) is combined with 5 µL of 1X Cutsmart buffer, 2 µL of Nb.BbvCI, and 42.5 µL of nuclease-free biological water for a total reaction volume of 50 µL. The mixture is held at 37°C for one hour and then heated to 80°C for twenty minutes to denature the enzyme. In step two, the hybridisation, the nicked section of the plasmid is replaced, via competition, with another oligonucleotide. This oligonucleotide has the exact same sequence of 66 bases, but also has an additional few bases plus either one or two biotins. The oligonucleotide is diluted down to a 3:1 ratio, such that there are three oligonucleotides for every plasmid based on the original amount of plasmid added, in phosphate buffer (50 mM phosphate, 100mM Na^+, pH 8.0). The oligonucleotide and the reaction mixture are then combined and heated to 75°C for two minutes and cooled to 4°C at a rate of 1°C/min. To maximise the

amount of oligonucleotide incorporation, the plasmid then undergoes a ligation step, step three. This is performed with a Blunt/TA ligase master mix from NEB. 50 μL of this master mix are combined with 50 μL of the reaction mixture. It is incubated at room temperature for twenty minutes. Step four is the linearisation step. In order to improve the nanopore results it is necessary to linearise the plasmid, so its behaviour is as similar to 10 Kbp DNA as possible. The SAL1 (NEB) cuts across both strands of the plasmid, such that the added biotin from the oligonucleotide ends up roughly in the middle of the plasmid. This is done by adding 0.4 μL SAL1 to each 50 μL of reaction mixture. This mix is then incubated at 37^{o}C for an hour and then 65^{o}C for twenty minutes to denature the enzyme. The reaction mixture is then purified with a Purelink PCR Microfiltration kit to remove the enzymes and extra oligonucleotides. The filtration kit is able remove DNA less than 100 bp and more than 12 Kbp. Finally, step five is the incubation of the plasmid carrier with protein. This is performed in the phosphate buffer (50 mM phosphate, 100 mM Na^{+}, pH 8.0), and the carrier and protein are incubated at room temperature for an hour.

The protein used for the proof of concept experiments was monovalent streptavidin rather than naturally occurring quadrivalent streptavidin. The monovalent streptavidin is a modified quadrivalent streptavidin (the wild type) where three of the binding sites are inactive, and thus the size is similar with only a small decrease in the strength of binding. [1] This is shown in **Figure 4.3**. This was to reduce the risk of multiple carriers binding to one streptavidin.

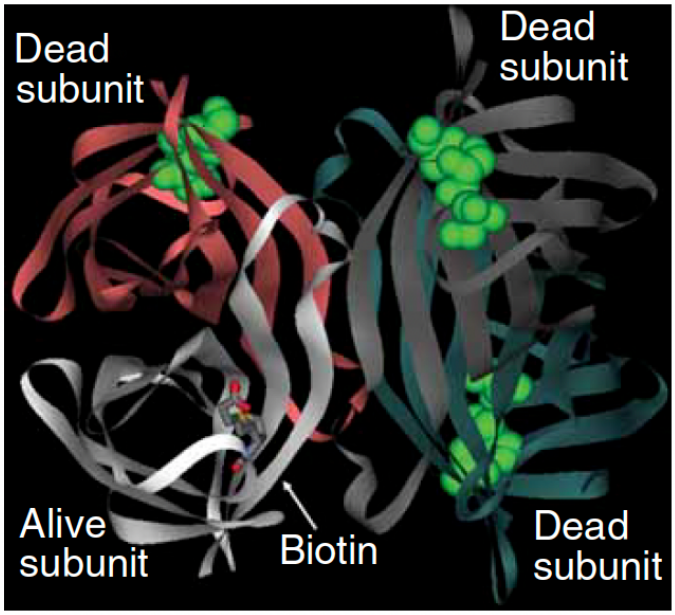


***Figure 4.3**: Modified streptavidin molecule, where three of the binding sites are inactivated such that it becomes monovalent streptavidin. The monovalent streptavidin has a similar K_d to quadrivalent streptavidin. The monovalent streptavidin was developed by the Howarth Laboratory. Adapted from Reference 1.*[1]

4.5 OLIGONUCLEOTIDE SEQUENCES

Three oligonucleotides purchased from IDT were used in these experiments. The oligonucleotides all had the original 66 base sequence, which is nicked by Nb.BbvCI: 5'-TGA GGT TTG GGC GGC GAC CTG GCT CAA GCG AGT GGA AAA AGT TAG AAG CTT AAA AAC TTA CGC AGC-3'. The potential hairpin turns and self-annealing sites were minimised as much as possible so that the oligonucleotides would not bind to each other or themselves, thus increasing the amount that are successfully hybridised.

For the first oligonucleotide, BO, six As were added to the 5' side, as well as biotin, so that the biotin sticks out from the plasmid by about 2 nm.[2] The final sequence is 5'-biotin-AAA AAA TGA GGT TTG GGC GGC GAC CTG GCT CAA GCG AGT GGA AAA AGT TAG AAG CTT AAA AAC TTA CGC AGC-3'. The sequence has three potential hairpin turns and fifteen potential annealing sites (all of which are six bases long) and the additional six As do not affect these numbers.[3]

Figure 4.4 shows that this oligonucleotide is able to bind to biotin successfully. The agarose gel concentration for both of these gels is 2%, which allows good visualisation of

both the unbound and bound oligonucleotide. The Varying Protein Concentration gel keeps a constant amount of oligonucleotide while gradually increasing the amount of protein. The highest ratio clearly visible is 50X, and as there is only one band, this suggests that all of the BO has been bound to monovalent streptavidin. As the protein concentration is not very high and the oligonucleotide one is, it was necessary to dilute the oligonucleotide to get the higher ratios so the concentration of the oligonucleotide might not be high enough to show up in the gel. As the streptavidin is quite big, 56 KDa for the quadrivalent streptavidin and a similar size for the monovalent streptavidin, the oligonucleotide is not able to travel as far in the gel when it is bound.[4] The gel varying oligonucleotide concentration shows the opposite results. At first when the ratio of oligonucleotide to monovalent streptavidin is very low, most of the oligonucleotide is bound, but as it increases significantly, less and less is so multiple bands are again visible.

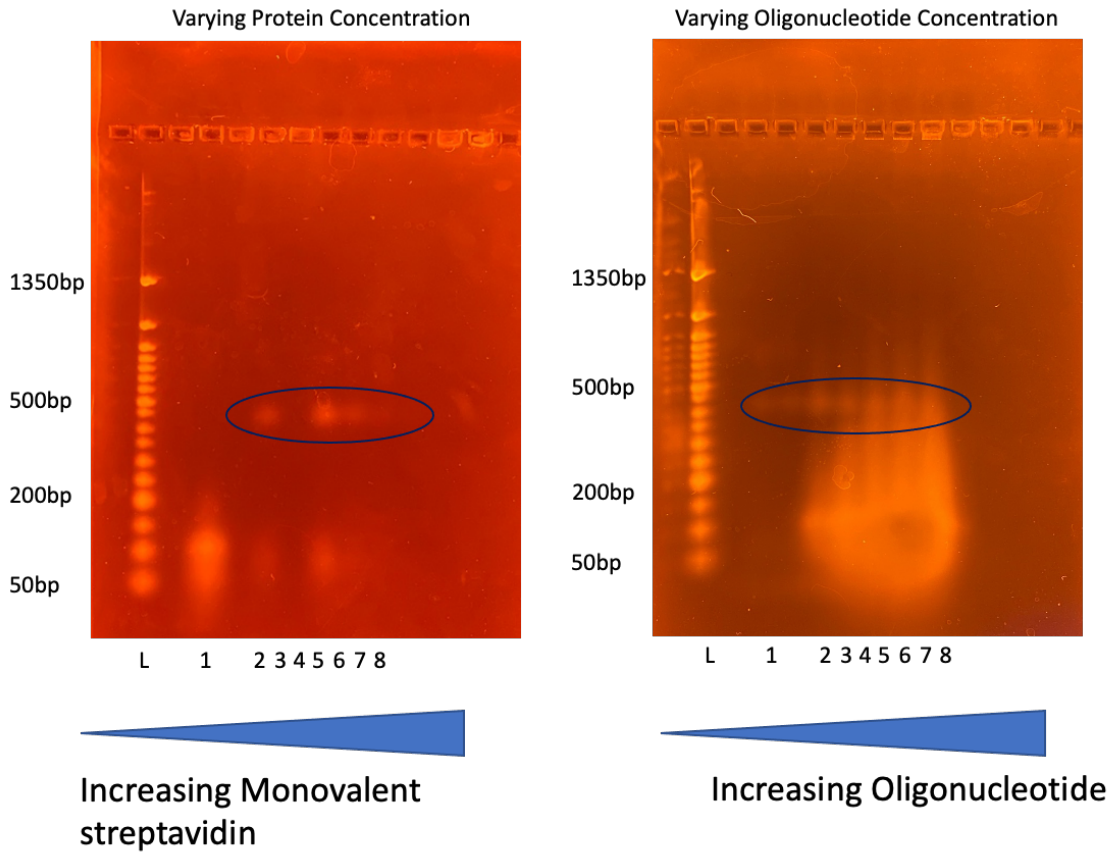


***Figure 4.4**: Both gels are 2% agarose and the ladder goes from 50 bp to 1350 bp. For varying protein concentration the wells are 1. BO, 2. 0.5X (monovalent streptavidin to oligonucleotide), 3. 1X, 4. 10X, 5. 50X, 6. 100X, 7. 250X, and 8. 1000X.. The wells for the varying oligonucleotide concentration are 1. Monovalent streptavidin, 2. 0.5X (oligonucleotide to monovalent streptavidin), 3. 1X, 4. 10X, 5. 50X, 6. 100X, 7. 250X, 8. 1000X, 9. BO only. The bands where binding occurred are marked with a circle.*

The second oligonucleotide (BO2) has two biotins added, one on the 5’ end and the other on the 3’ end. As there are two biotins, twelve additional bases have been added so that each biotin can stick out. Additionally, the 66 base sequence is long enough that steric effects should not prevent multiple streptavidin binding; each quadrivalent streptavidin is about 5 nm in diameter, while the length of the oligonucleotide is approximately 22 nm.[5] The oligonucleotide sequence is 5’-biotin-AAA AAA TGA GGT TTG GGC GGC GAC CTG GCT CAA GCG AGT GGA AAA AGT TAG AAG CTT AAA AAC TTA CGC AGC AAT AAA-biotin-3’. As and Ts were used as the additional bases so that the melting temperature is still less than 75^{o}C, as this is the temperature used in the hybridisation step. The sequence with the additional twelve bases has four potential hairpin turns and seventeen potential annealing sites, all of which are six bases long, which is one and two more respectively than the 66 base sequence on its own.[3]

Figure 4.5 again demonstrates that the oligonucleotide is able to bind to monovalent streptavidin. However, unlike with the single biotin, as BO2 has two biotins, it is able to bind two of them. This is apparent in the Varying Protein Concentration gel, where the monovalent streptavidin concentration is gradually increased as it is possible to see two bands much higher up in the gel than where the oligonucleotide alone appears. The oligonucleotide alone appears slightly higher up than one would expect, as it is in between the 100 and 150 bp bands, and is only 78 bp long. However, perhaps the additional biotin could explain why it is shifted slightly higher. Importantly, it is shifted significantly higher with the addition of protein. The Varying Oligonucleotides Concentration gel, where the BO2 concentration is gradually increased does not show the binding quite as nicely as the other gel, but it still suggests that there is binding occurring.

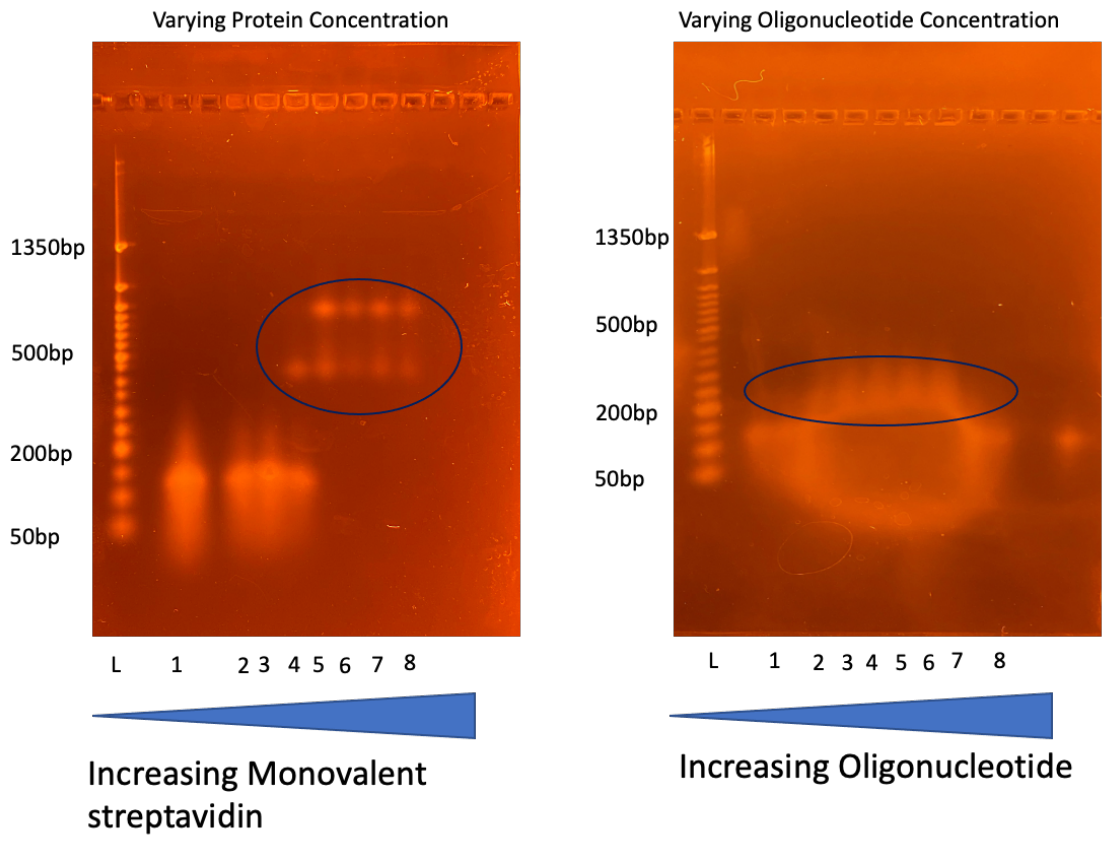


***Figure 4.5**: The gel was made with 2% agarose and the ladder ranges from 50 bp to 1350 bp. For the Varying Protein Concentration gel the bands are 1. BO2, 2. 0.5X (monovalent streptavidin to oligonucleotide), 3. 1X, 4. 10X, 5. 50X, 6. 100X, 7. 250X, 8. 1000X. For the Varying Oligonucleotides Concentration gel, the wells are similar, with 1. 0.5X (oligonucleotide to monovalent streptavidin), 2. 1X, 3. 10X, 4. 50X, 5. 100X, 6. 250X, 7. 1000X, and 8. BO2. The bands where binding occurred are marked with a circle.*

4.6 INITIAL PROTOCOL AND INITIAL MODIFICATIONS

The protocol mentioned previously in this chapter is the final protocol. However, some significant changes have been applied to reach that protocol. Initially, the plasmid was incubated with 2 μL of both restriction enzymes, SAL1 and Nb.BbvCI, for one hour at 37°C and then heated to 80°C in order to denature both enzymes. The total reaction volume was 50 μL with 0.5 μL plasmid, 5 μL 1X Cutsmart buffer, and 40.5 μL biological water making up the remaining reaction volume. Afterwards, the sample was purified using a Purelink PCR microfiltration kit. **Figure 4.6** shows what the plasmid looks like after being nicked and cut at the same time. Well 1 is the uncut plasmid (PC), well 2 is the linearised plasmid (PCSAL1), and well 3 is the linearised, nicked, and filtered plasmid (PCSAL1NbBVfil). The SAL1 does

seem to be able to linearise the plasmid, as shown by the band at 10 Kbp, but it does not seem to be very efficient.

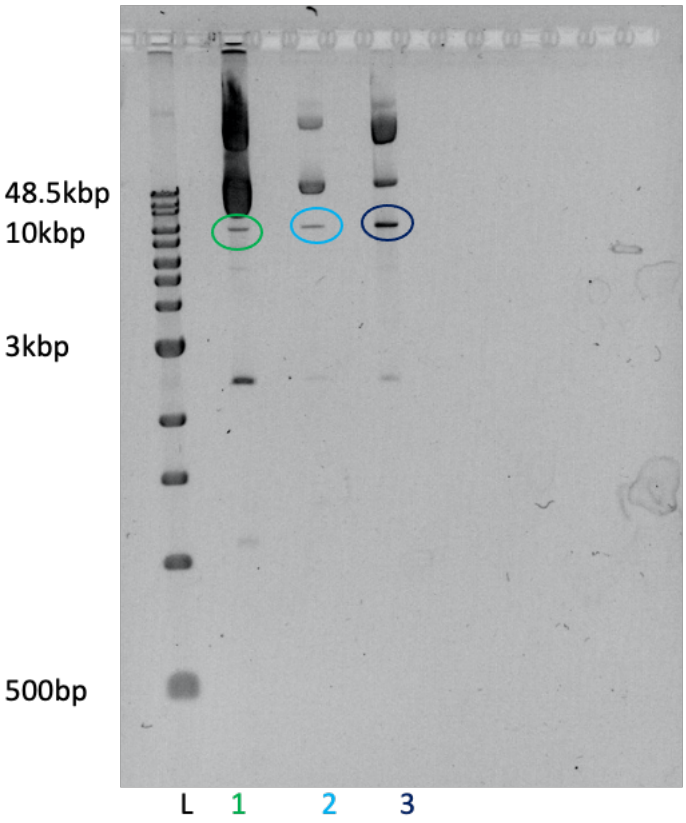


***Figure 4.6**: 1% agarose gel showing the uncut plasmid (PC) (1), the linearised plasmid (PCSAL1) (2), and the nicked and linearised plasmid with filtration (PCSAL1NbBVfil) (3). The ladder ranges from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

As the SAL1 did not seem to be very efficient at linearising the plasmid, the amount was increased to 3 µL with the same incubation conditions; the amount of water was reduced so that the total volume was still 50 µL. The gel in **Figure 4.7** shows the results of this change. It seems that there is less uncut plasmid, but there is still a band present above 10 Kbp for all the wells. The wells again are well 1, the uncut plasmid (PC), well 2, the linearised plasmid (PCSAL1), and well 3,the linearised and nicked plasmid (PCSAL1NbBVfil).

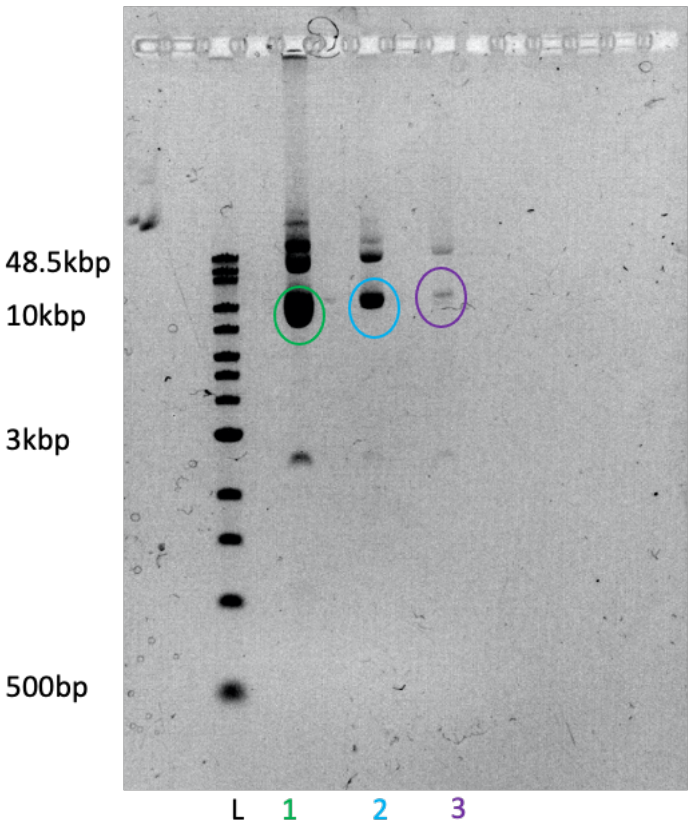


***Figure 4.7**: 1% agarose gel showing the uncut plasmid (PC) (1), the linearised plasmid (PCSAL1) (2), and the nicked and linearised plasmid with filtration (PCSAL1NbBVfil) (3). The ladder ranges from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

In addition to incubating both restriction enzymes with the plasmid at the same time, the protocol initially included an extra filtration step. After the incubation, the plasmid was filtered before being hybridised. After the filtration, the concentration of the plasmid was measured and this number was used to calculate how much oligonucleotide was needed. The hybridisation protocol was heating to 95°C for 1 min and then cooling down to 4°C at a rate of 1°C/min. Also, the hybridisation buffer was 50 mM NaCl 10 mM Tris and 1 mM EDTA, called annealing buffer. The oligonucleotides, BO, were added to the linearised and nicked plasmid in a ratio of 100:1. As the second filtering could not remove the uncut plasmid, a gel extraction instead of filtering was performed as the second purification step so that the sample would be as pure as possible.

Thus the steps for modified plasmid shown in **Figure 4.8** are incubated with the restriction enzymes (2 μL of each enzyme), filtration, hybridisation with annealing buffer, and gel extraction with a QIAquick gel kit. Each well has the same PCSAL1NbBVBOfil sample. While the SAL1 did not cut all of the plasmid, it was possible to extract only the cut plasmid

from the gel by using the band at 10 Kbp. The excess oligonucleotide from the hybridisation process is shown at the bottom of the gel.

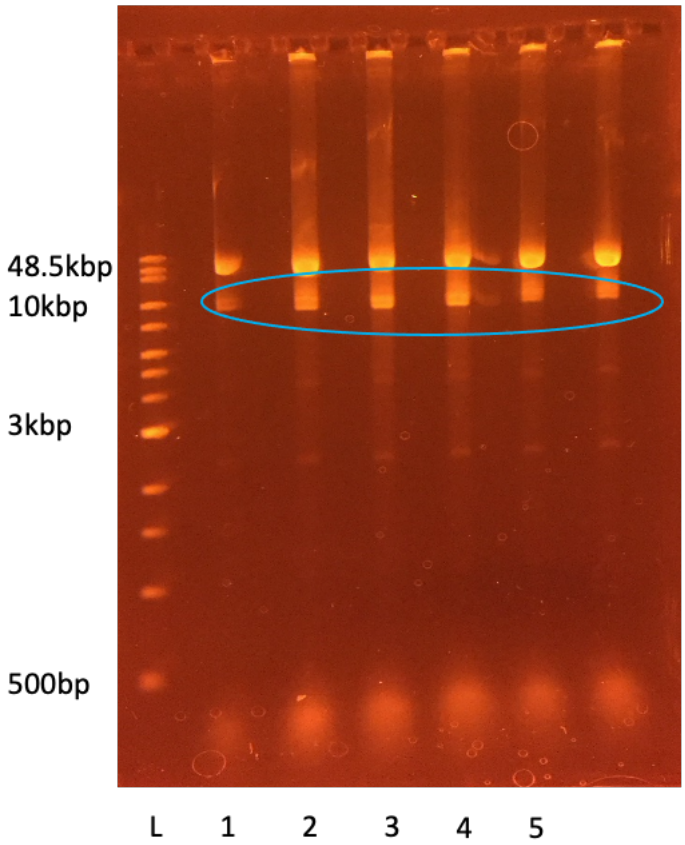


***Figure 4.8**: 0.7% agarose gel for gel extraction of PCSAL1NbBVBOfil. All the wells contain the same sample and the ladder extends from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

This protocol was then slightly modified by using a different gel extraction kit, the Monarch DNA gel extraction kit. This change was needed as the QIAquick gel kit has a range of 70 bp-10 Kbp, while the Monarch kit has a range from 50 bp-25 Kbp.[6,7] As the plasmid is slightly over 10 Kbp with the added oligonucleotide, the Monarch kit was deemed to be the better purification choice.

4.7 FURTHER PLASMID PROTOCOL MODIFICATIONS

To improve the efficiency of the enzymes and the incorporation of the oligonucleotide into the plasmid several protocol steps were changed. This included an addition of a ligation step and removing the filtration step before hybridisation to reduce sample loss. As ligation was now included the restriction enzymes were incubated with the plasmid at different times, such that the ligase could not repair the linearisation from the

SAL1. The new order was nicking with Nb.BbvCI, hybridisation with BO, ligation with T4 ligase, linearisation with SAL1, and then gel extraction. The hybridisation step was also changed as the oligonucleotide was diluted in a phosphate buffer (50 mM phosphate, 100 mM Na^+, pH 8.0), so that the EDTA in the annealing buffer could not affect the ligation step.[8] The ligation step involved incubating 100 μL of reaction volume with 10 μL 10X ligase buffer with 2 μL of T4 ligase at 16°C for twelve hours and then at 65°C for ten minutes. The separating of the nicking and cutting steps also allowed for denaturing the SAL1 at 65°C instead of at 80°C. **Figure 4.9** shows what the plasmid looks like before gel extraction; all of the wells contain the same sample. The separation of the gel steps and the ligation do seem to have increased the efficiency of the linearisation, as the band at 10 Kbp is much more prominent than the other bands.

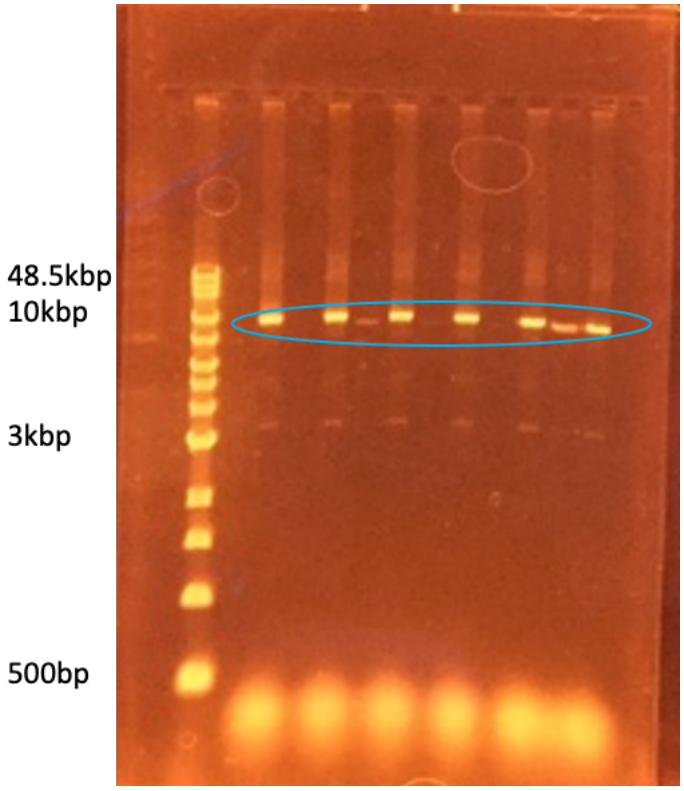


***Figure 4.9**: This 0.7% agarose gel with wells with PC2ELigBOA. As before the ladder is from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

To increase the hybridisation efficiency, the temperature of the protocol was varied. It initially held the plasmid at 95°C for 1 min before lowering the temperature. However, in order to reduce issues with parts of the plasmid separating and then not coming back together properly, the temperature was lowered to 85°C and then finally 75°C. It was not

possible to go any lower than 75°C as the melting temperature (Tm) for the oligonucleotide is only slightly lower than this temperature. **Figure 4.10** shows that the amount of left over oligonucleotide seems to decrease even at 85°C, suggesting that the oligonucleotide binding is more efficient.

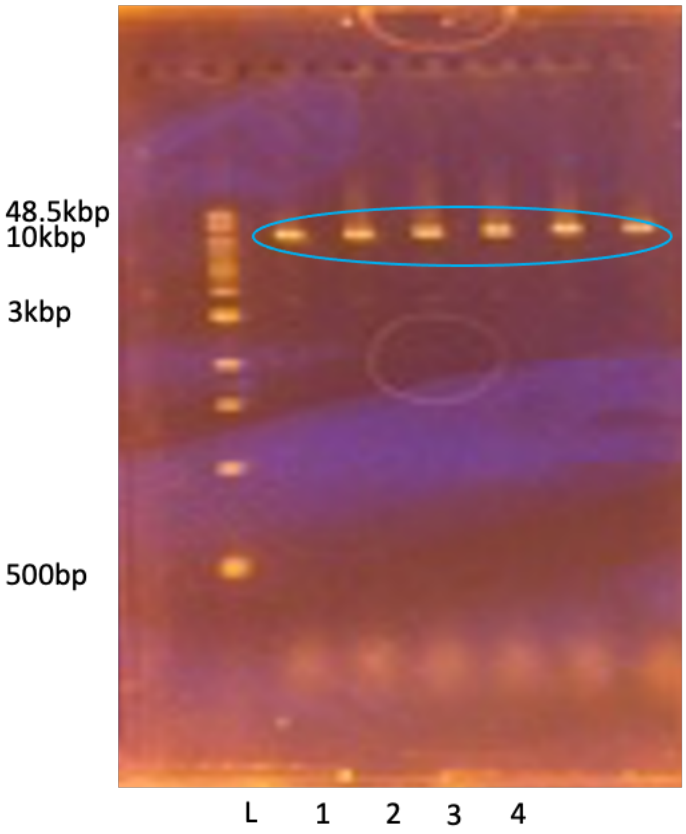


***Figure 4.10**: This 0.7% agarose gel shows the PC2ELigBOA before it has been extracted from the gel. All the wells are the same sample, and the hybridisation temperature was lowered to 85°C. The ladder as before ranges from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

Another hybridisation buffer, 150 mM NaCl, 10 mM $MgCl_2$, TE buffer, pH 7.4, was also explored, but this did not seem to have a big effect on the hybridisation. Similarly, another buffer for incubating the monovalent streptavidin and plasmid was explored, in addition to the phosphate buffer, but this buffer, 140 mM NaCl, 20 mM $MgCl_2$, 20 mM PBS pH 7.4, also did not seem to have a significant effect.

Further protocol changes included switching back to filtration as the purification method. Gel extraction takes significantly longer than filtration, and as it does not seem like there is an important difference in filtered versus extracted samples, as shown in **Figure 4.11**, filtration was used as the purification method. In the Purification Method Comparison

gel, well 1 is the un-purified sample, well 2 the filtered sample, and well 3 the extracted sample. No significant difference is observed in bands between the filtered and extracted samples. Additionally, as each well was filled with the sample volume of sample, there seems to be a higher concentration in the filtered sample. A similar lack of difference is shown in the Purification Comparison with Monovalent streptavidin gel, where well 1 is the filtered sample, well 2 the extracted sample, well 3 the 1X filtered sample, and well 4 the 1X extracted sample. However, on this gel it is possible to see a slight band for the filtered samples further down, but as it is quite faint and the filtration both decreases the loss of sample and is faster, filtration was used as the purification method.

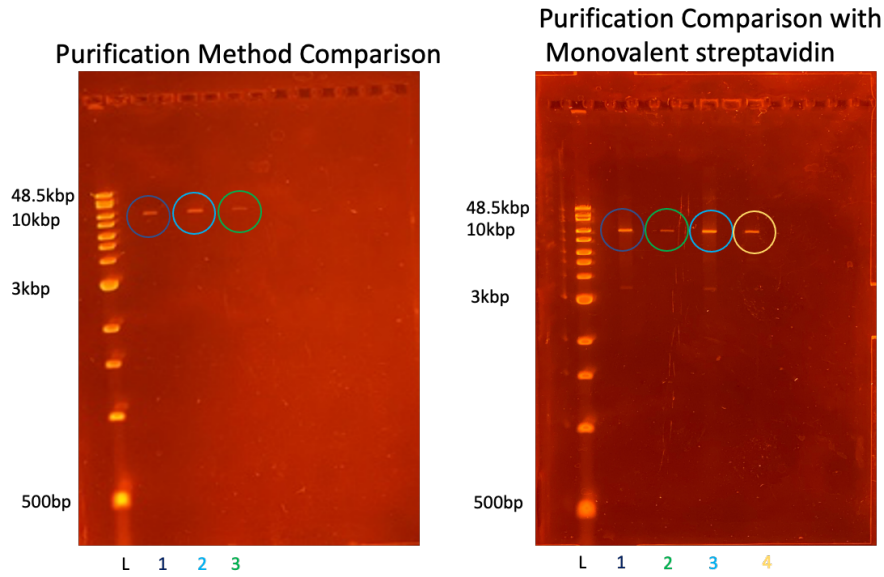


***Figure 4.11**: The 1% agarose gels show an comparison of the purification methods used for the plasmid carrier, filtration and gel extraction. The Purification Method Comparison gel compares the un-purified sample, well 1, the filtered sample, well 2, and the extracted sample, well 3, while the Purification Comparison with Monovalent streptavidin gel compares the filtered sample, well 1, the extracted sample, well 2, with the 1Xfiltered sample (1:1 ratio of monovalent streptavidin to plasmid), well 3, and the 1Xextracted sample, well 4. Both gels use a 1 Kbp extended ladder that ranges from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

Gels showing what the outcome of the all of the different purification steps are shown in **Figure 4.12**. This complete protocol is called Protocol A. Well 1 for both gels is the uncut plasmid, PC, well 2 is the nicked plasmid, PCNbBV, well 3 is the nicked and hybridised plasmid, PCNbBVBOA or PCNbBVBO2A, well 4 is the nicked, hybridised, and ligated plasmid, PCNbBVLigBOA or PCNbBVLigBO2A, well 5 is the nicked, hybridised, ligated, and cut plasmid, PC2ELigBOA or PC2ELigBO2A, and well 6 is the nicked, hybridised, ligated, cut, and

filtered plasmid, PC2ELigBOfilA or PC2ELigBO2filA. This naming convention is summarised in **Table 4.1**.

Modification Steps	Naming Convention
Uncut Plasmid	PC
Nicked Plasmid	PCNbBV
Nicked and Hybridised Plasmid with Biotin Oligonucleotides	PCNbBVBOA
Nicked and Hybridised Plasmid with Two Biotins Oligonucleotides	PCNbBVBO2A
Nicked, Hybridised, and Ligated Plasmid with Biotin Oligonucleotides	PCNbBVLigBOA
Nicked, Hybridised, and Ligated Plasmid with Two Biotins Oligonucleotides	PCNbBVLigBO2A
Nicked, Hybridised, Ligated, and Cut Plasmid with Biotin Oligonucleotides	PC2ELigBOA
Nicked, Hybridised, Ligated, and Cut Plasmid with Two Biotins Oligonucleotides	PC2ELigBO2A
Nicked, Hybridised, Ligated, Cut, and Filtered Plasmid with Biotin Oligonucleotides	PC2ELigBOfilA
Nicked, Hybridised, Ligated, Cut, and Filtered Plasmid with Two Biotins Oligonucleotides	PC2ELigBO2filA
Plasmid that undergoes all modification step conditions without any added enzyme or oligonucleotides	PCstepsA
Nicked Plasmid that undergoes all modification step conditions with only the nicking enzyme added	PCNbBVstepsA
Nicked and Ligated Plasmid that undergoes all modification step conditions with the nicking and ligating enzymes added	PCNbBVLigstepsA
Nicked, Ligated, Cut, and Filtered Plasmid that undergoes all modification step conditions	PC2ELigfilA

Table 4.1*: Table listing all of the plasmid naming conventions for the different modification steps and controls.*

For wells 5 and 6, it is possible to see that the SAL1 was able to cut the plasmid successfully, as this is when the band at 10 Kbp shows up. Furthermore, particularly comparing wells 5 and 6 for the BO gel, the filtered sample is more concentrated than the unfiltered sample, as the band is brighter, as well as, shown more clearly with the BO2

sample, a tighter band. These gels show that the SAL1 is able to cut the modified plasmid efficiently, as wells 5 and 6 have a band at 10 Kbp. Additionally, they suggest that the filtration is able to both remove unwanted excess material, as shown by the tighter band in the BO2 gel, and concentrate the sample, as shown by the brighter band in the BO gel.

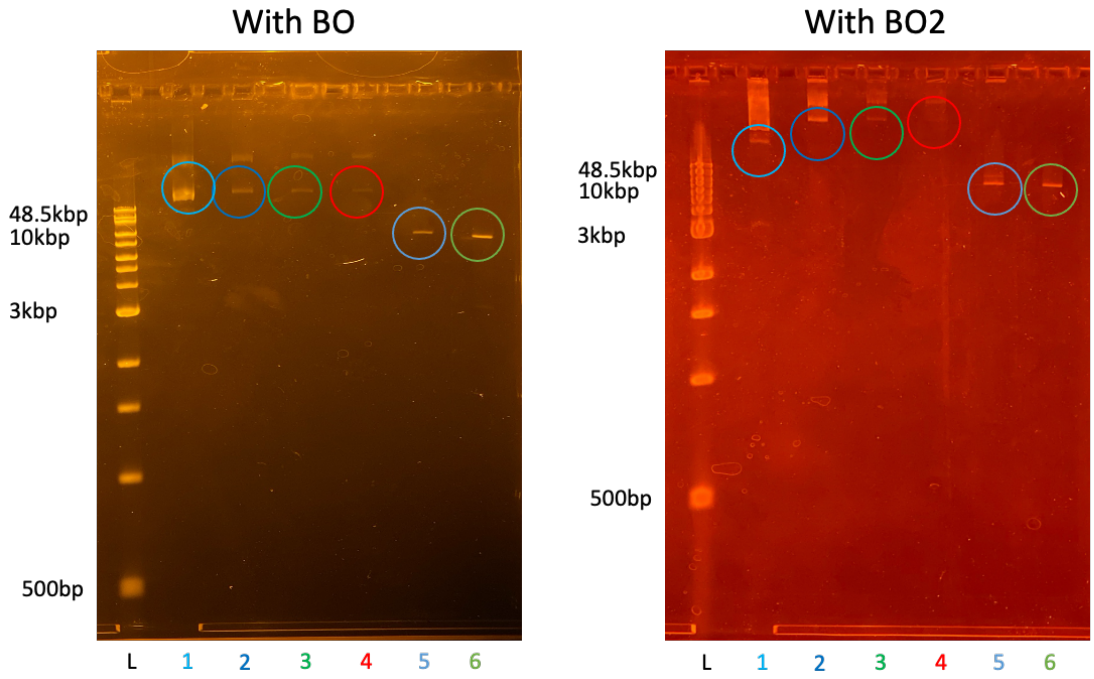


***Figure 4.12**: 1% agarose gels showing the results of the all of the modification steps of the plasmid, Protocol A, for both the carrier with the BO and BO2 oligonucleotides. The ladder used for both gels ranges from 500 bp to 48.5 Kbp. The wells are well 1, uncut plasmid, PC, well 2, nicked plasmid, PCNbBV, well 3, nicked and hybridised plasmid, PCNbBVBOA or PCNbBVBO2A, well 4, nicked, hybridised, and ligated plasmid, PCNbBVLigBOA or PCNbBVLigBO2A, well 5, nicked, hybridised, ligated, and cut plasmid, PC2ELigBOA or PC2ELigBO2A, and well 6, nicked, hybridised, ligated, cut, and filtered plasmid, PC2ELigBOfilA or PC2ELigBO2filA. The band of interest in each step is shown with a circle.*

The binding of monovalent streptavidin to this modified plasmid carrier was also explored. Several gels were run to see if the binding of the monovalent streptavidin to the plasmid carrier was successful. However, unfortunately, as shown in **Figure 4.13**, it does not seem possible to distinguish the binding in agarose gel. **Figure 4.13** is a gel of the binding curve of PC2ELigBOfilA, such that well 1 is PC2ELigBOfilA, well 2 is 0.5X (ratio of monovalent streptavidin to plasmid), well 3 is 1X, well 4 is 10X, well 5 is 50X, well 6 is 100X, well 7 is 250X, and well 8 is 1000X. While it is possible to see the plasmid carrier band at 10 Kbp, it does not seem like there is an additional band for the bound samples. The additional bands also seem to be present in the filtered sample. This conclusion follows as it is possible to see

the binding with the oligonucleotides themselves only in a 2% gel and even then the change according the ladder is only by a few hundred base pairs, such that when the plasmid is 10 Kbp long, the change with the monovalent streptavidin would not be expected to show up in the gel.

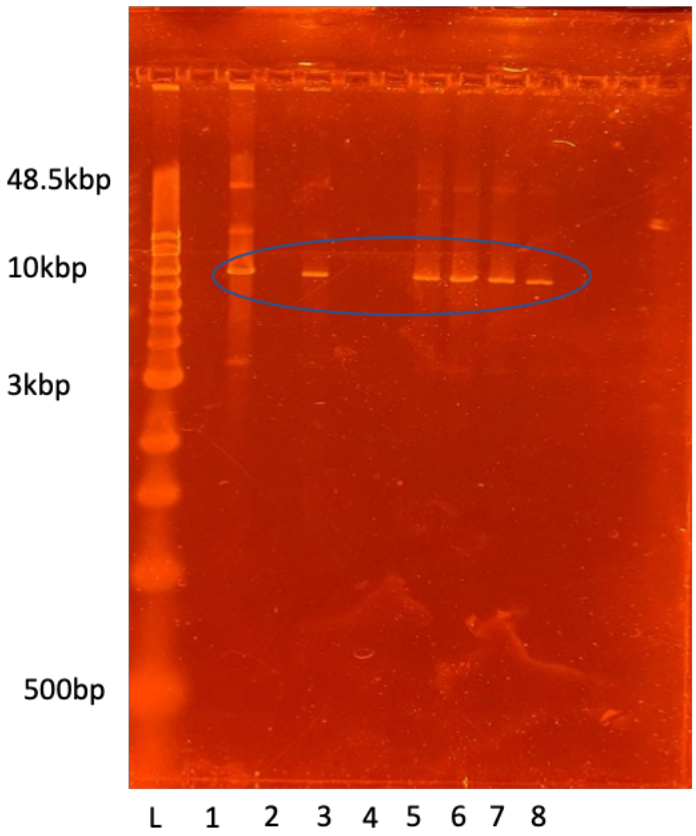


***Figure 4.13**: A 1% agarose gel showing the binding curve of PC2ELigBOfilA. Well 1 is PC2ELigBOfilA, well 2 is 0.5X (ratio of monovalent streptavidin to plasmid), well 3 is 1X, well 4 is 10X, well 5 is 50X, well 6 is 100X, well 7 is 250X, and well 8 is 1000X. The ladder is 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

4.8 FINAL PROTOCOL MODIFICATIONS: A NEW OLIGONUCLEOTIDES AND LIGATION METHOD

To increase the efficiency of the hybridisation further a new oligonucleotide with phosphorylation was used. The third oligonucleotide (BOph) has the biotin on the 3' end, so that the 5' end can be phosphorylated to facilitate the ligation: this phosphorylation is necessary when the oligonucleotide is being ligated.[9] The sequence is 5'-phosphorylated-TGA GGT TTG GGC GGC GAC CTG GCT CAA GCG AGT GGA AAA AGT TAG AAG CTT AAA AAC TTA CGC AGC AAT AAA-biotin-3'. The additional six bases are AAT AAA, which is the additional six on the 3' end used in the two biotin oligonucleotide. The sequence has four

potential hairpin turns and sixteen potential self-annealing sites, all of which are six bases long.[3] **Figure 1** in the **Appendix I** shows that this oligonucleotide is also able to successfully bind to monovalent streptavidin.

In addition to changing the oligonucleotide, the ligation protocol was adjusted as well. The main impetus behind this change was due to the changing lab schedule with the Covid-19 pandemic. This ligation protocol, Protocol B, used a ligation master mix that allowed for a much shorter incubation time with the ligase. The initial nicking step was not changed. For the hybridisation step, the ratio of oligonucleotide to plasmid was lowered to 3:1 from 100:1, as this master mix should greatly increase the efficiency of the ligation, but the reaction mixture was still heated to 75^{o}C and then cooled to 4^{o}C. 50 μL of the hybridised plasmid was then incubated with 50 μL of the master mix at room temperature for twenty minutes, rather than the overnight incubation. The cutting and purification steps were not changed. This is the protocol described in the Plasmid Formation Method section. **Figure 4.14** is a gel showing these plasmid structures. While the BO2 oligonucleotide was not changed, as it is not possible to add phosphorylation due to the biotin being on both ends of the oligonucleotide, the new ligation was performed with this oligonucleotide. Well 1 is the PC2ELigBOphfilB, well 2 is the PC2ELigBO2filB, well 3 is 1XPC2ELigBOphfilB, and well 4 is 1XPC2ELigBO2filB. The Protocol B samples were incubated with the monovalent streptavidin in the same way, at room temperature for an hour in a phosphate buffer. **Figure 4.14** suggests that plasmids were still modified correctly, they were definitely cut properly, and that due to the lower ratio of oligonucleotide, there was a lot less leftover oligonucleotide so that there was not an apparent band in the oligonucleotide, like in the previous protocol, Protocol A. Additionally, it was still not possible to see a shift from the binding of the monovalent streptavidin.

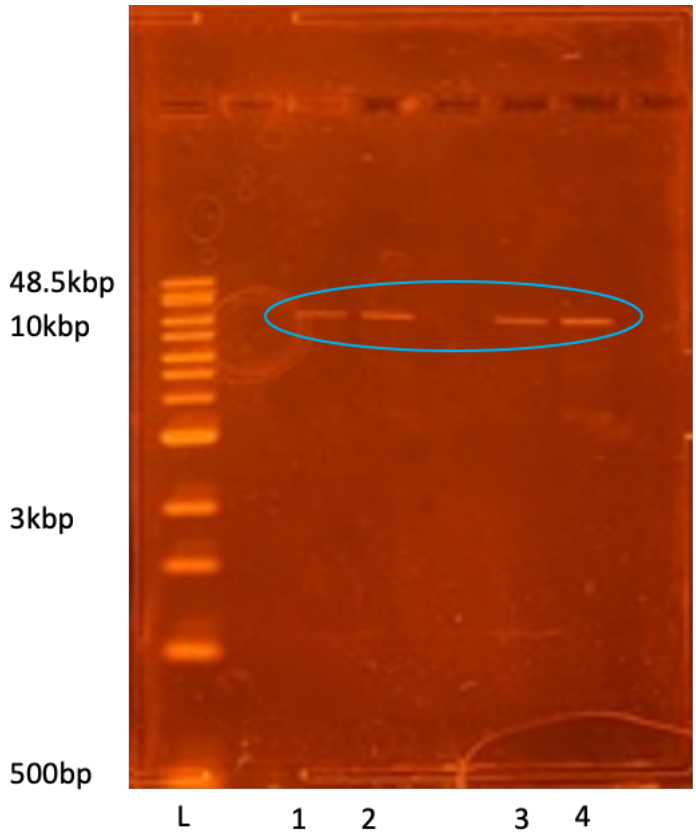


***Figure 4.14**: 1% agarose gel showing Protocol B for the plasmid. Well 1 is the PC2ELigBOphfilB, well 2 is the PC2ELigBO2filB, well 3 is 1XPC2ELigBOphfilB, and well 4 is 1XPC2ELigBO2filB. The ladder ranged from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

Figure 4.15 shows the results of each modification step for the final protocol, Protocol B. Well 1 is the uncut plasmid (PC), well 2 is the nicked plasmid (PCNbBV), well 3 is the nicked and hybridised plasmid (PCNbBVBOphB or PCNbBVBO2B), well 4 is the nicked, hybridised, and ligated plasmid (PCNbBVLigBOphB or PCNbBVLigBO2B), well 5 is the nicked, hybridised, ligated, and cut plasmid (PC2ELigBOphB or PC2ELigBO2B), and well 6 is the filtered plasmid (PC2ELigBOphfilB or PC2ELigBO2filB). The naming conventions for these modification steps are summarised in **Table 4.2**.

Modification Steps and Controls	Naming Convention
Uncut Plasmid	PC
Nicked Plasmid	PCNbBV
Nicked and Hybridised Plasmid with Phosphorylated Biotin Oligonucleotides	PCNbBVBOphB
Nicked and Hybridised Plasmid with Two Biotins Oligonucleotides	PCNbBVBO2B
Nicked, Hybridised, and Ligated Plasmid with Biotin Oligonucleotides	PCNbBVLigBOphB
Nicked, Hybridised, and Ligated Plasmid with Two Biotins Oligonucleotides	PCNbBVLigBO2B
Nicked, Hybridised, Ligated, and Cut Plasmid with Biotin Oligonucleotides	PC2ELigBOphB
Nicked, Hybridised, Ligated, and Cut Plasmid with Two Biotins Oligonucleotides	PC2ELigBO2B
Nicked, Hybridised, Ligated, Cut, and Filtered Plasmid with Biotin Oligonucleotides	PC2ELigBOphfilB
Nicked, Hybridised, Ligated, Cut, and Filtered Plasmid with Two Biotins Oligonucleotides	PC2ELigBO2filB
Plasmid that undergoes all modification step conditions without any added enzyme or oligonucleotides	PCstepsB
Nicked Plasmid that undergoes all modification step conditions with only the nicking enzyme added	PCNbBVstepsB
Nicked and Ligated Plasmid that undergoes all modification step conditions with the nicking and ligating enzymes added	PCNbBVLigstepsB
Nicked, Ligated, Cut, and Filtered Plasmid that undergoes all modification step conditions	PC2ELigfilB

Table 4.2*: Table showing all of the naming conventions for the different plasmid structures and controls that were formed using Protocol B.*

For the PC2ELigBOphfilB gel, the SAL1 seems to have cut the plasmid very efficiently, as there is only one band not at 10 Kbp after the cutting step, and the filtration is able to remove this band. Additionally, the results of each step looks similar to the previous protocol, suggesting that it is still able to modify the plasmid correctly. For the PC2ELigBO2filB gel, the SAL1 was not able to cut as efficiently, as there are extra bands left after the filtration step. However, the SAL1 can degrade easily, so the use of a new SAL1

does seem to avoid this problem. This particular sample was not cut very efficiently, but the rest of the steps do seem quite similar to the previous protocol suggesting that Protocol B works for the BO2 oligonucleotide as well.

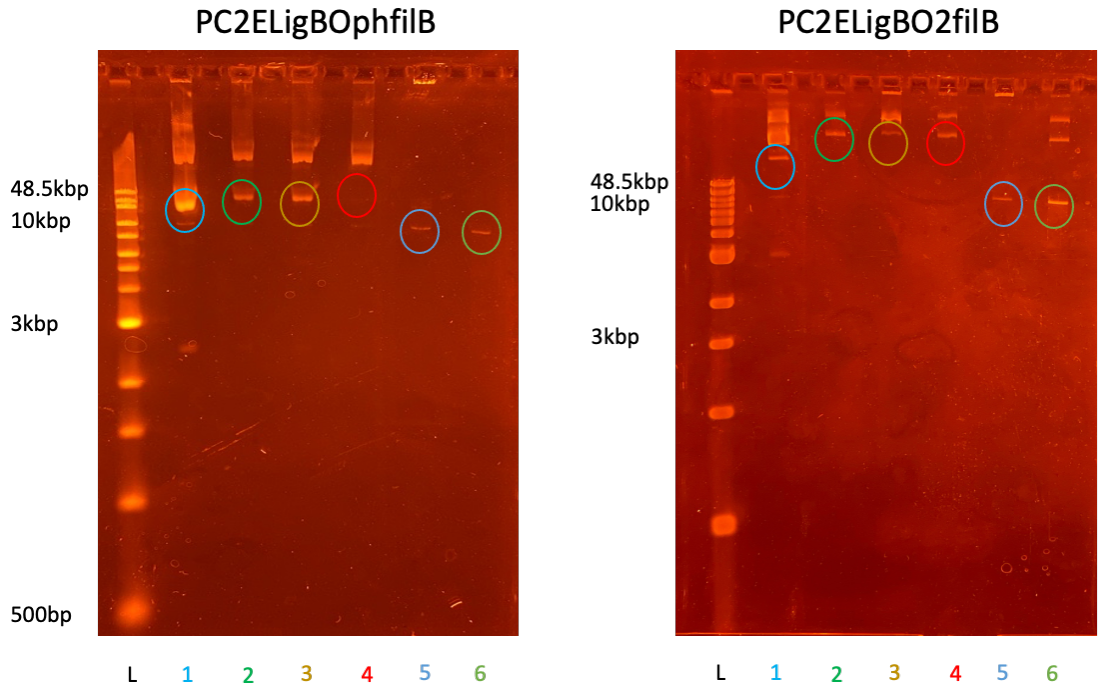


***Figure 4.15**: 1% agarose gels showing all of the steps in Protocol B. Well 1 is the uncut plasmid (PC), well 2 is the nicked plasmid (PCNbBV), well 3 is the nicked and hybridised plasmid (PCNbBVBOphB or PCNbBVBO2B), well 4 is the nicked, hybridised, and ligated plasmid (PCNbBVLigBOphB or PCNbBVLigBO2B), well 5 is the nicked, hybridised, ligated, and cut plasmid (PC2ELigBOphB or PC2ELigBO2B), and well 6 is the filtered plasmid (PC2ELigBOphfilB or PC2ELigBO2filB). The ladder used for these gels ranged from 500 bp to 48.5 Kbp. The band of interest in each step is shown with a circle.*

A comparison of the filtered and unfiltered samples was also done. **Figure 4.16** shows this gel. As shown in this gel, both the plasmid with the BOph and the BO2 need purification, but there are no longer a large number of excess oligonucleotides as the ratio is now 3:1. The filtering seems able to remove some of the uncut plasmid or any large clumps of plasmid that cannot easily move throughout the gel, as there is only one dark band, at 10kbp, for both PC2ELigBOphfilB and PC2ELigBO2filB. In **Figure 4.16**, well 1 is PC2ELigBOphB, well 2 is PC2ELigBO2B, well 3 is PC2ELigBOphfilB, and well 4 is PC2ELigBO2filB.

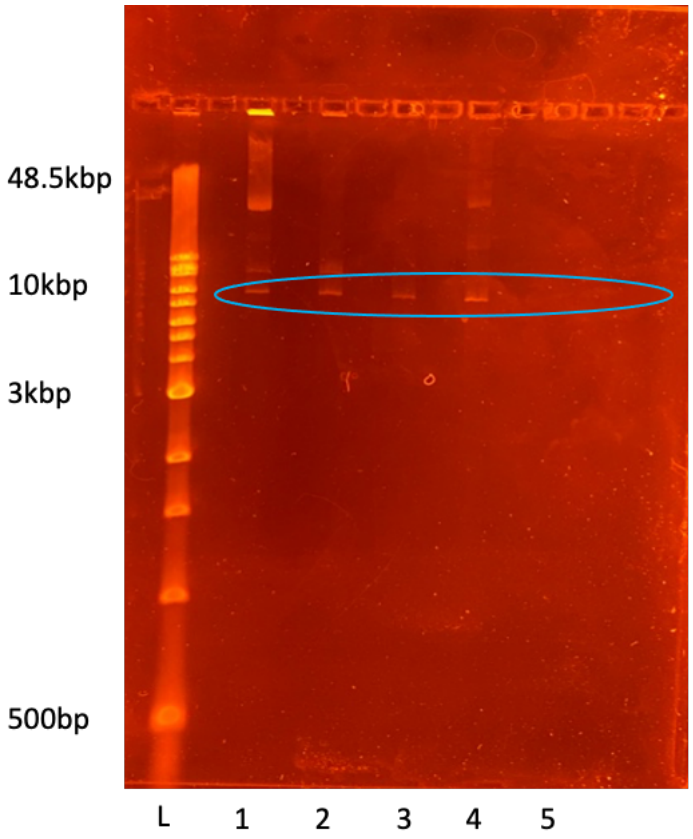


***Figure 4.16**: This 1% agarose gel compares the samples made with Protocol B before and after filtration. Well 1 is PC2ELigBOphB, well 2 is PC2ELigBO2B, well 3 is PC2ELigBOphfilB, and well 4 is PC2ELigBO2filB. The ladder is from 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

These samples, PC2ELigBOphfilB and PC2ELigBO2filB, were also incubated with monovalent streptavidin and no shift shows up in these gels, shown in **Figure 2** in **Appendix I**.

BEHAVIOUR OF THE PLASMID IN THE NANOPORE

4.9 VOLTAGE STUDIES

Voltage studies are when the current trace for specific sample is measured at a variety of voltages in one pipette. They were performed to confirm that the nanopore was able to detect plasmid translocations. **Figure 4.17** shows the behaviour of PC2ELigBOphfilB in one pipette over a variety of voltages. The sample was loaded into the pipette, cis-to-trans, and events were recorded at -200 mV, -300 mV, and -400 mV. The buffer was 100 mM KCl in TE pH 8, and the sample concentration was 100 pM. As shown in the figure, the dwell time, t_d, decreased as the magnitude of the voltage increased, a trend that is expected, and

the peak amplitude, I_p, increased as the magnitude of the voltage increased, which was also expected. It is also possible to see that the peak amplitude increases in the ten second sample traces for each voltage. The behaviours are what would be expected for double-stranded DNA travelling through the nanopore, so the plasmid carrier behaves similarly to other double-stranded DNA. The current traces and events were analysed using a custom Matlab script, and all had the same event parameters applied.

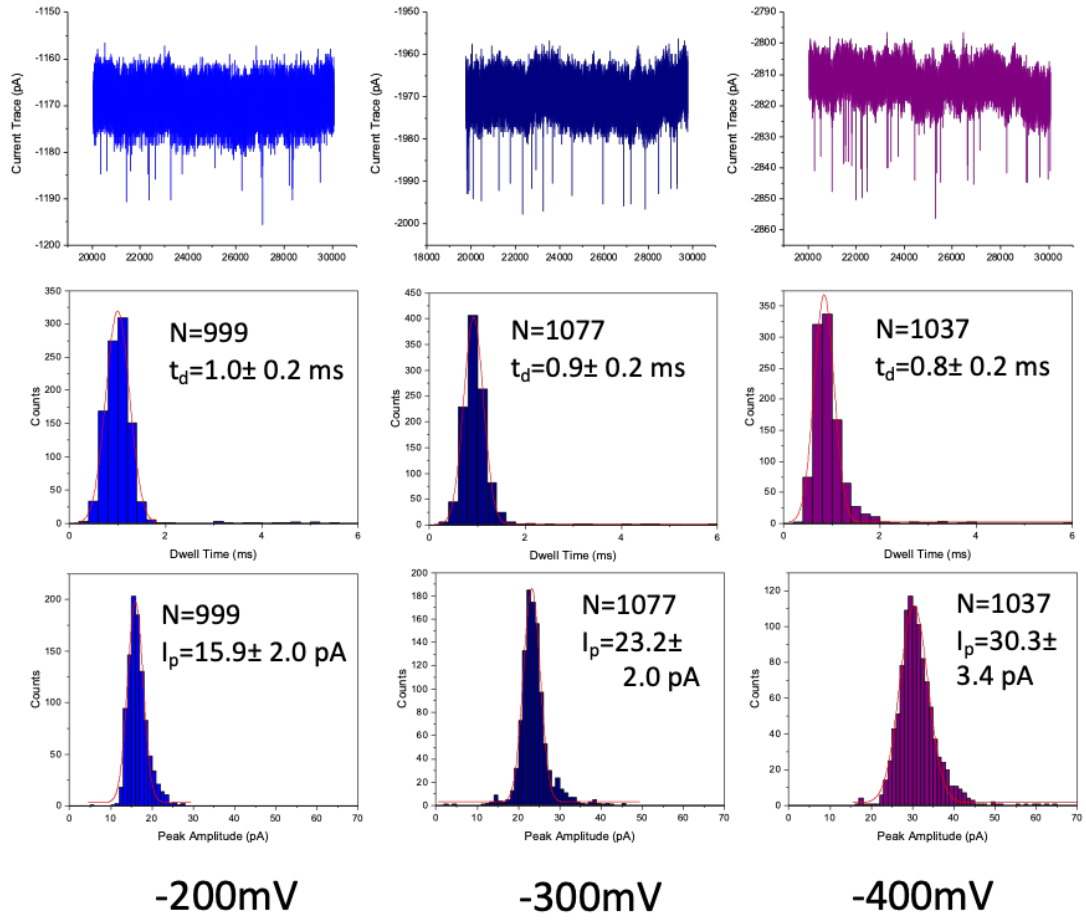


***Figure 4.17**: Voltage study of the dwell time, t_d, and peak amplitude, I_p, for PC2ELigBOphfilB where the sample was loaded into the pipette, cis-to-trans. The buffer used for these experiments was 100 mMKCl in TE buffer pH 8.0 and the sample concentration was 100 pM. A 30 KHz filter and a sampling rate of 100 KHz were used. A Gaussian fit was used to determine the average values. A one way ANOVA test was performed and both the dwell time and peak amplitude were significantly different.*

These voltage studies were performed for all the samples, which included PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, and PC2ELigBO2filA**. Figure 4.18** shows graphs of voltage vs dwell time and voltage vs peak amplitude for all of these samples. They were performed in 100 mM KCl TE buffer pH 8.0 and cis-to-trans. The recording conditions

were also the same, a 100 KHz sampling rate, a 30 KHz filter, and a sample concentration of 100 pM. The data for these voltage studies came from at least three different pipettes per sample and a weighted average was used to determine the t_d and the I_p, as well as the errors. The peak amplitude trend of increasing while increasing the voltage is very clear in all four samples and at all the voltages. The dwell time trend is not so clear, although it suggests a decreasing dwell time as voltage increases. Perhaps as the decrease in the dwell time is not as large as the increase for the peak amplitude, this trend is more susceptible to slight changes in pipette shape or size. Changing in pipette size might have a particularly high effect as they can result in interactions between the plasmid and the pore wall, which would slow down the dwell time. These interactions could also affect the confirmation of the plasmid as it passes through the pore.

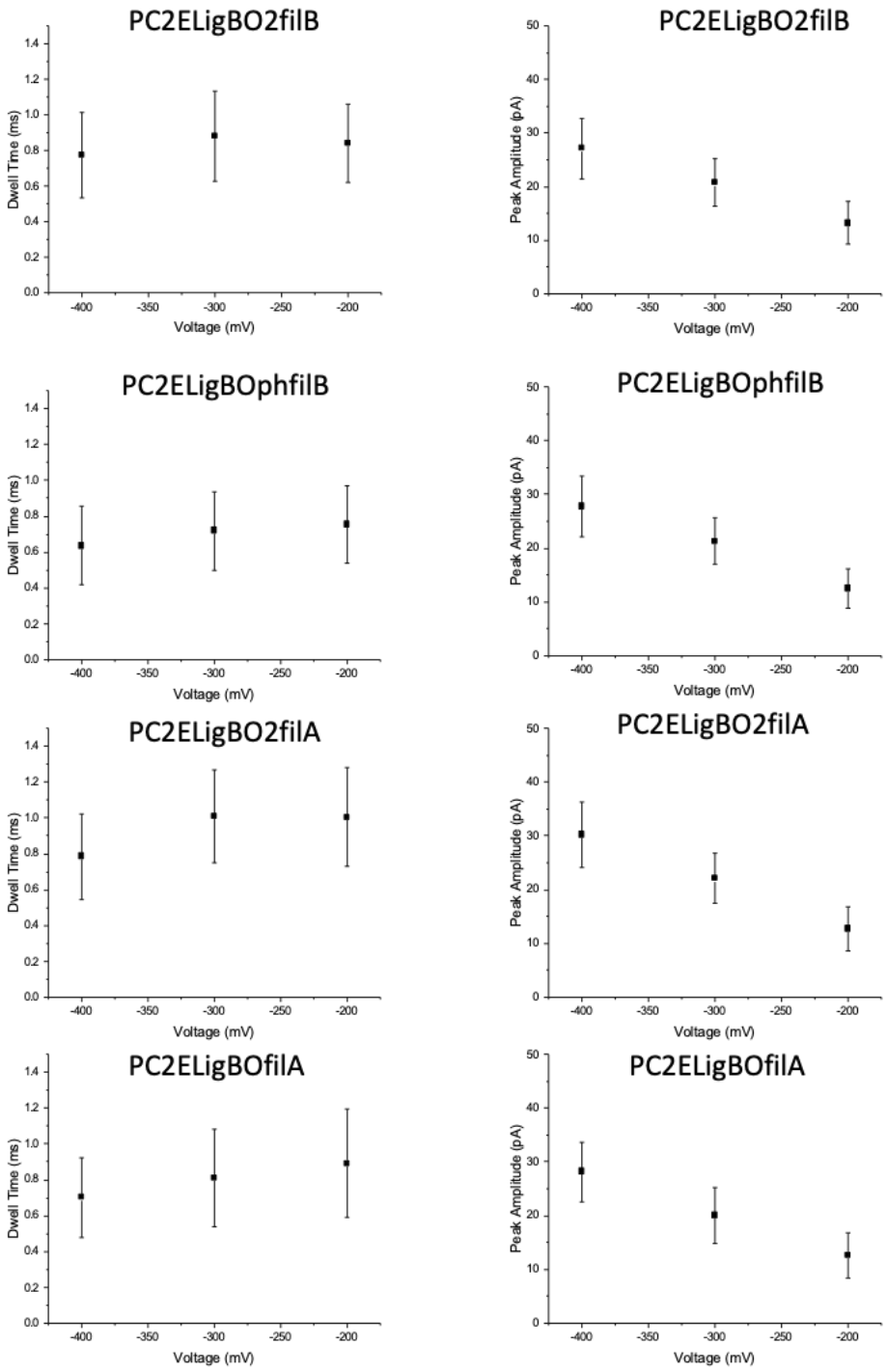


Figure 4.18*: Voltage study graphs for PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, PC2ELigBO2filA. The experiments were performed cis-to-trans and in a 100 mM KCl TE buffer pH 8.0. The recording conditions were 100 KHz sampling rate and 30 KHz filter, and these were the same for all experiments. The sample concentration was 100 pM. The data for each sample comes from taking a weighted average of results from at least three different pipettes. The data was analysed using a custom Matlab script and with the same event conditions.*

As well as exploring cis-to-trans behaviour, trans-to-cis data was recorded for these samples. For trans-to-cis, the capture rate tends to be dramatically lower for the same

concentration. As there are fewer events, the trends are not sometimes as clear as for cis-to-trans. The buffer used was 1 M KCl in TE pH 8.0, rather than 100 mM KCl. Also as the sample is loaded into the bath rather than the pipette, the sign of the voltage is reversed, such that for these voltage studies events were recorded at 200 mV, 300 mV, and 400 mV. **Figure 4.19** shows a similar comparison as **Figure 4.17**, where results from one pipette are shown. Both the histograms and the ten second current traces at each voltage suggest similar trends to the cis-to-trans; that as the magnitude of the voltage increases dwell time decreases, and as magnitude of the voltage increases peak amplitude increases. It is also easier to see the increasing number of events in a ten second period, thus showing the trend of the increasing capture rate as the voltage increases. This trend is not so apparent for cis-to-trans, as there are so many events in a ten second period; it is not immediately visible without analysing the current trace.

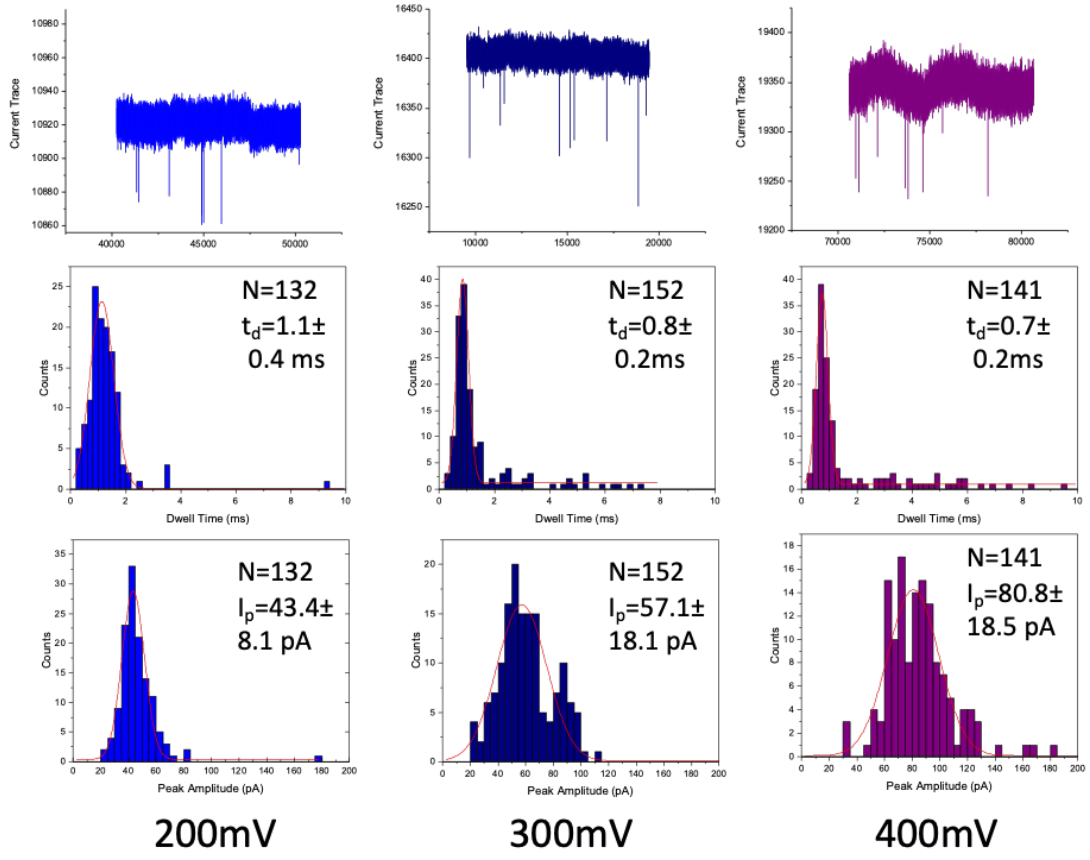


***Figure 4.19**: Trans-to-cis voltage study for the dwell time, t_d, and peak amplitude, I_p, of PC2ELigBO2filA.. The voltage ranged from 200-400 mV and this data was all recorded in one pipette. The buffer used for these experiments was 1M KCl in TE buffer pH 8.0 and the sample concentration was 100 pM. The sampling rate was 100 KHz and the filter was 30 KHz. A Gaussian fit was used to determine the average values. A one way ANOVA test was performed level and both the dwell time and peak amplitude were significantly different.*

The voltage study trends for all four samples, PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, PC2ELigBO2filA are shown in **Figure 4.20**. Unlike the cis-to-trans data, the dwell time trend and peak amplitude trend is very clear and is as expected for the trans-to-cis data. The dwell time decreases clearly for all the samples as the voltage increases and, similarly, the peak amplitude increases as voltage increases for all of them. The voltage studies use a weighted average of data from at least three pipettes for the dwell time, dwell time error, peak amplitude, and peak amplitude error. The concentration is 100 pM and the buffer used is 1M KCl in TE pH 8.0.

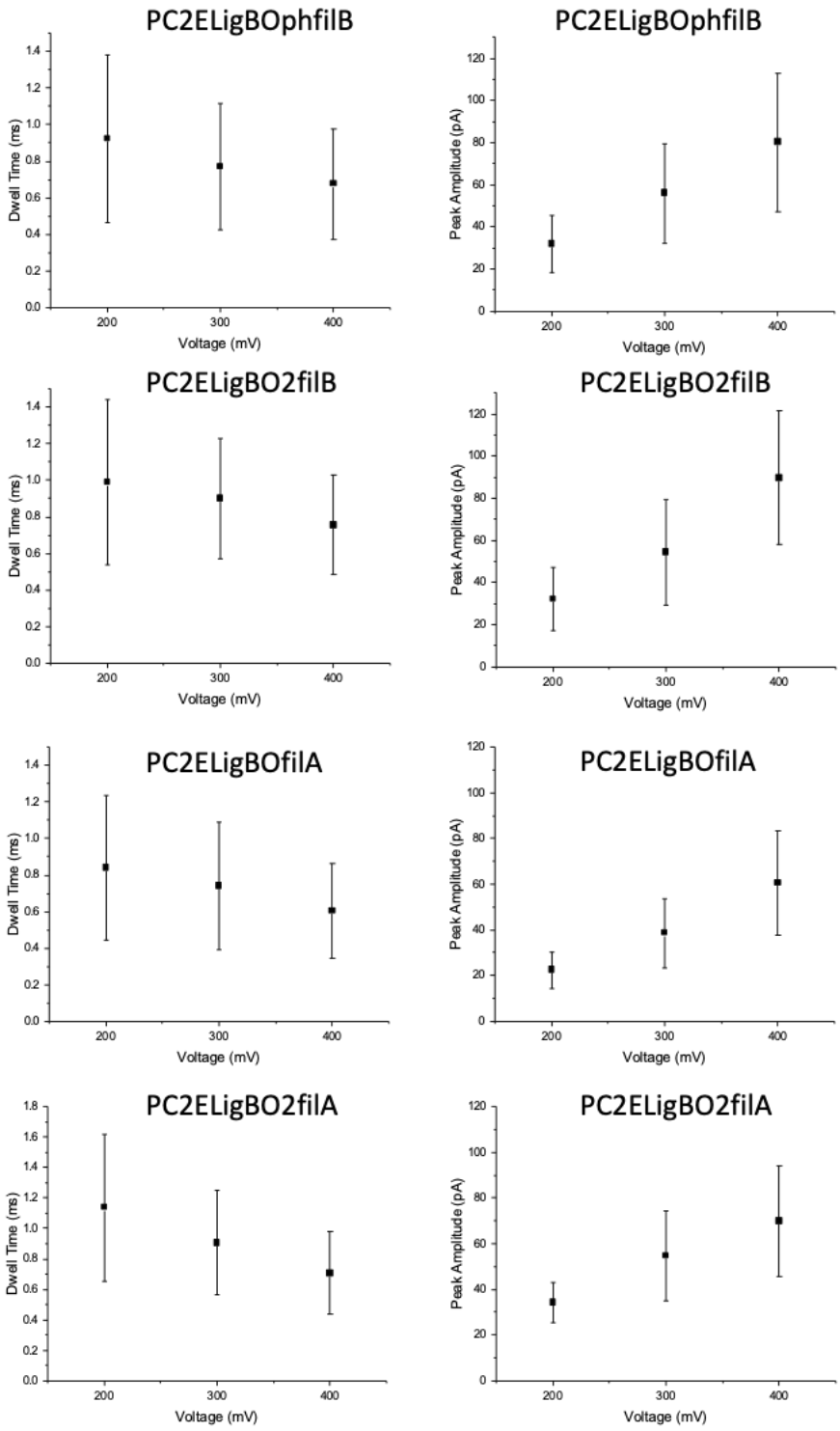


Figure 4.20*: Voltage studies for PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, PC2ELigBO2filA, for trans-to-cis. The data was taken from at least three different pipettes for each sample and a weighted average was used for the dwell time, dwell time error, peak amplitude, and peak amplitude error. The sample concentration was 100 pM and the buffer used was 1M KCl in TE pH 8.0. The sampling rate was 100 KHz and the filter was 30 KHz. The voltage range was 200-400 mV.*

The pipettes used for all of these voltage studies were characterised and their I-Vs are shown in **Figure 4.21**. There was some variation between the pipettes, particularly for some trans-to-cis ones. However, as this voltage study data was recorded over a period of nine months this variation is expected. The pipette puller already has some variation and during this time the protocol was adjusted slightly, such that some pipettes were slightly larger or differently shaped than others. The diameter of the pipettes used in the cis-to-trans studies was 14.4-21.7 nm and for trans-to-cis it was 8.7-19.5 nm. These diameters were calculated using the method described by Gong *et al.* (2014).[10] The use of weighted averages should help reduce the error that comes from comparing pipettes that vary in shape and size, and as evidenced by the clear trends shown in the voltage studies, it seems it does.

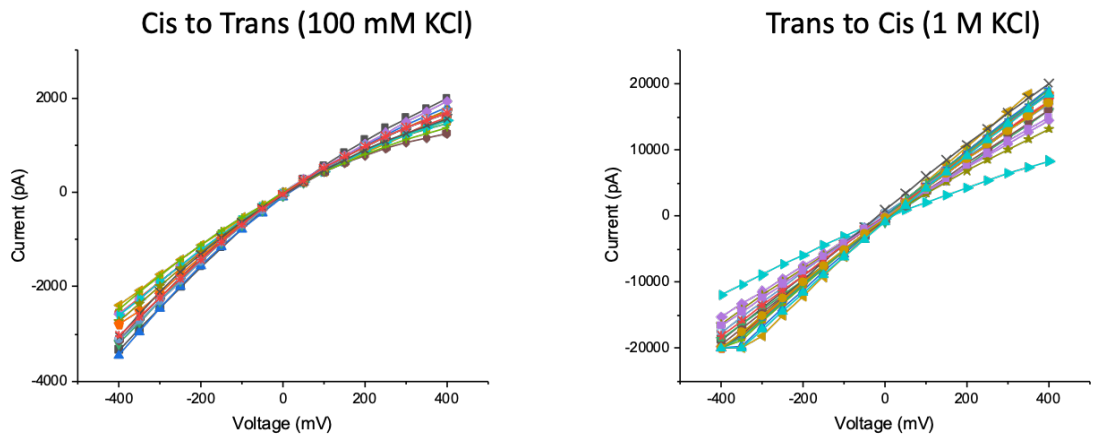


***Figure 4.21**: I-Vs for pipettes used in all of the voltage studies. The I-V was recorded first and in the exact same conditions as the voltage study. The pipette diameters varied from 14.4-21.7 nm for cis-to-trans and 8.7-19.5 nm for trans-to-cis.*

4.10 COMPARISON TO 10 KBP DNA FOR CIS-TO-TRANS

In addition to the voltages studies, the plasmid carrier was also compared to 10kbp DNA to determine that it behaved in a similar way to linear 10 Kbp DNA. Dwell times and peak amplitudes were compared for PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, and PC2ELigBO2filA with 10 Kbp DNA. The 10 Kbp DNA was recorded in the same pipette for each sample. The pipette was washed with dI water in between samples. It is important to compare these histograms, the dwell time ones are shown in **Figure 4.22** and the peak amplitude ones in **Figure 4.23**, for measurements made in the same pipette as pipette size and shape can have a significant effect on how the samples move through it. Overall the

average dwell times and average peak amplitudes are quite similar. The average dwell times are typically less than 0.1 ms apart and there is typically less than 1 pA difference for the peak amplitudes. Each sample had a least 100 events, although many had more. While this data comes from individual pipettes, the similarity in behaviour between the 10 Kbp DNA and the plasmid carrier is a good reflection of typical results. Thus, it is reasonable to conclude that the plasmid carrier, even with various different oligonucleotide additions, behaves similarly to 10 Kbp DNA.

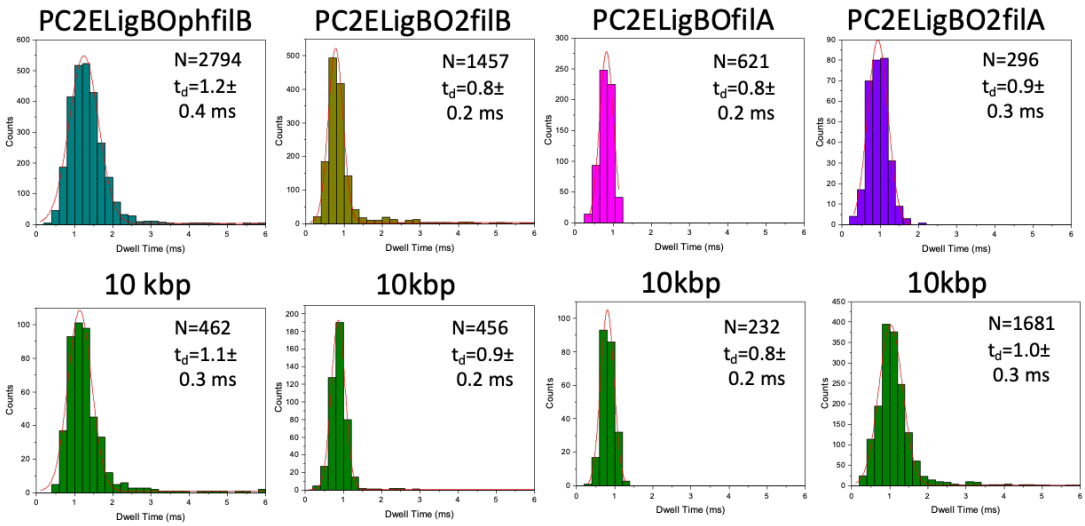


***Figure 4.22**: A comparison of dwell times for the plasmid carrier and 10 Kbp DNA. All of the data was cis-to-trans and the sample concentration was always 100 pM. The experimental conditions were the same with a buffer of 100 mM KCl in TE pH 8.0, a 100 KHz sampling rate, a 30 KHz filter, and the voltage was -300 mV. A Gaussian fit was used to determine the average values.*

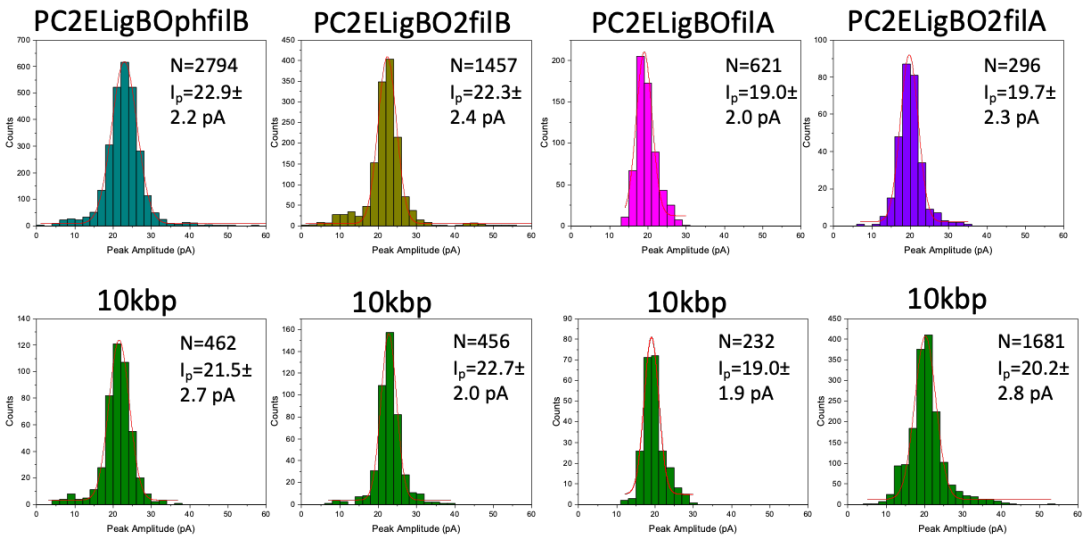


***Figure 4.23**: A comparison of peak amplitudes for the plasmid carrier and 10 Kbp DNA. All of the data was recorded cis-to-trans and the sample concentration was always 100 pM. The experimental conditions were the same with a buffer of 100 mM KCl in TE pH 8.0, a 100 KHz sampling rate, a 30 KHz filter, and the voltage was -300 mV. A Gaussian fit was used to determine the average values.*

VARIATION OF pH

4.11 POTENTIAL EFFECT OF pH

The pH of the buffer can have a significant effect on the results of the nanopore. While the pH is unlikely to change the overall shape of the double-stranded DNA plasmid, it is possible that some secondary structures could be affected. This effect has been observed for single-stranded DNA in nanopores.[11] Additionally, some I-V rectification with pH changes have been observed. There can also be changes to the DNA dwell times and peak amplitudes as the surface charge density can be affected by the change in pH.[12] As pipette walls are slightly negative, this change in pH can have a significant effect on the plasmid translocations. Models that include a consideration of the H^+ and OH^- ions show that they can affect the conductance of the nanopore.[13]

There are two ways to explore this pH change. The first is by doing a pH gradient, such that the bath and the sample have a different pH. The second is by changing the pH of both the buffer and the sample from pH 8.0, as experiments in this research have typically be conducted at this pH.

4.12 pH GRADIENT

The first of the pH experiments conducted were varying the pH of the sample such that there was a pH gradient. These experiments were always performed cis-to-trans so the sample was diluted into 100 mM KCl TE buffers of various pHs, with a range of pH 5.0 to pH 10.0. The goal of these experiments was to see if there was any significant effect of the pH on the binding of the monovalent streptavidin. Thus the plasmid carrier, PC2ELigBOphfilB, was incubated in phosphate buffer with monovalent streptavidin at a ratio of 1:1. The sample with protein, 1X, and the PC2ELigBOphfilB were both diluted with 100 mM KCl TE buffers of pH 5.0, 7.0, 8.0, 9.0, and 10.0. As the incubation occurs before the dilution and the biotin-streptavidin bond is so strong across a variety of conditions, it was expected that the biotin and the monovalent streptavidin would be able to remain bound. The bath was kept constant at pH 8.0. **Figure 4.24** shows the average dwell times and peak amplitudes for these samples across the range of pHs. Unfortunately, there does not seem to be a significant difference, although the ligated often has a lower dwell time and a higher peak amplitude. This may stem from issues establishing a pH gradient in a nanopipette, as the dimensions of the pore may be too large. Thus, it does not seem that forming a pH gradient helps with determining which events have protein bound and which do not, so for this reason, as well as the desire to keeps the samples as close to biological conditions as possible, the pH gradient was not used again to study the protein binding to the plasmid carrier.

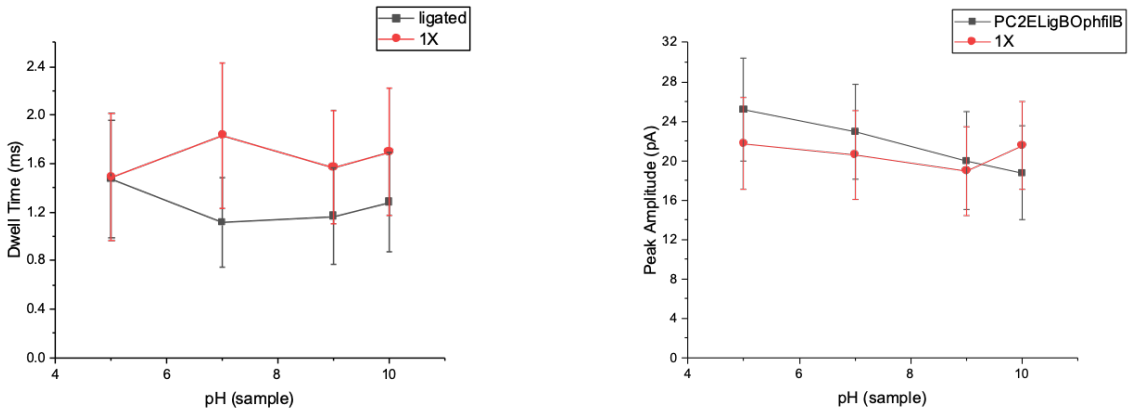


Figure 4.24*: Graphs of the average dwell time and peak amplitude of PC2ELigBOphfilB and 1X (ratio of monovalent streptavidin to PC2ELigBOphfilB is 1:1) over a variety of pHs. The buffer (100 mM KCl in TE) in the bath is kept constant at pH 8.0 while the sample pH varies from 5.0 to 10.0. The concentration of the sample is always 100 pM, the voltage -300 mV, the sampling rate 100 KHz, and the filter 30 KHz. The data comes from three different pipettes where both the carrier and the carrier with protein were recorded.*

4.13. VARIED pH

Varying the pH of both the sample and the bath was also explored. After each sample was recorded at a specific pH, the buffer in the bath was changed so that data could be recorded at a new pH. The pHs recorded were 5.0, 7.0, 8.0, 9.0, and 10.0. Similarly, to when a pH gradient was investigated, it was not expected that the changing pH would have much effect on the monovalent streptavidin-biotin bond. **Figure 4.25** shows the variation of the pHs. There again were not any clear trends, although it does seem that the greatest difference in dwell time between the bound and unbound sample occurred at pH 8.0, and unfortunately, this pH is when there is the least difference in peak amplitude. However, unlike with the gradient, it does seem that the dwell time tends to increase as the pH is increased. This is expected as, with increasing pH, there are more negative ions so there is less of a driving force for the plasmid to move through the nanopore, resulting in longer dwell times. There was no such trend for the peak amplitude. Thus, it did not seem like there was a clear ideal pH to conduct these experiments at, so pH 8.0 was continued to be used, as it is similar to physiological conditions and is well-studied.

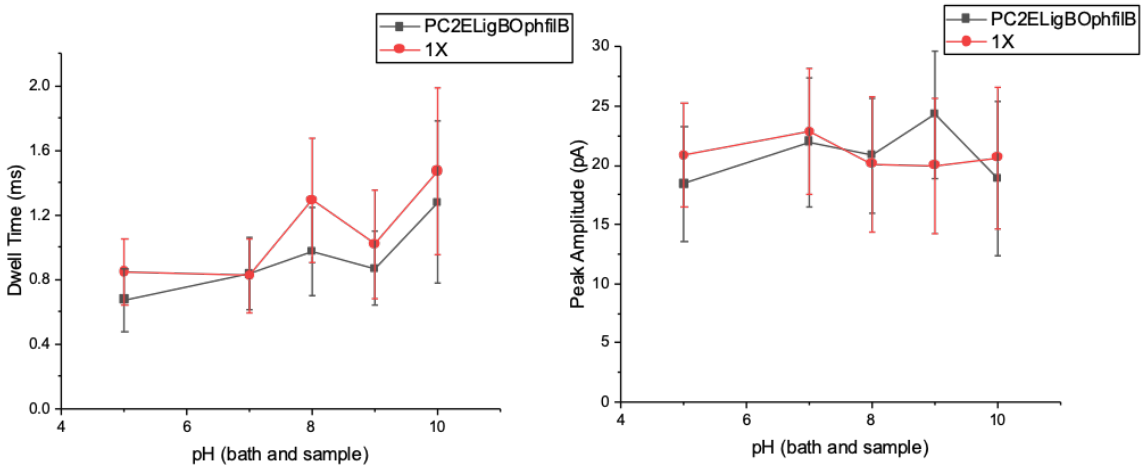


***Figure 4.25**: Graph of the change in average dwell time and peak amplitude of PC2ELigBOphfilB, and 1X, over various pHs. The buffer (100 mM KCl in TE) in the bath is changed while also varying the sample pHs, from 5.0 to 10.0. The sample concentration is always 100 pM, the voltage -300 mV, the sampling rate 100 KHz, and the filter 30 KHz. The data comes from three different pipettes where both the carrier and the carrier with protein were recorded.*

4.14 DISCUSSION OF pH CONTROLS

I-Vs were conducted for all the pipettes used and they are shown in **Figure 4.26**. For the different pHs, both bath and sample changing, the pipettes' diameter varied

significantly, from 14.2-28.4 nm, but it is possible that some of this effect has to do with the changing pH, as the same pipette was used for five different pHs; particularly, as current rectification has been observed at some buffer pHs. The gradient pipettes varied far less, only from 17.9-25.2 nm. Additionally, as the method used to calculate the diameter takes data from the I-V, and the current seems to change a bit with the changing pH, it could be that this method is not very accurate for calculating the diameter over a range of pHs. Furthermore, especially as the same pipette was used over several pHs, it is possible that AgCl could fall off the electrode and this could lead to I-V variation. This effect is likely to be pH dependent, which could explain why there is much greater variation in the Varied pH pipettes than in the pH gradient pipettes.

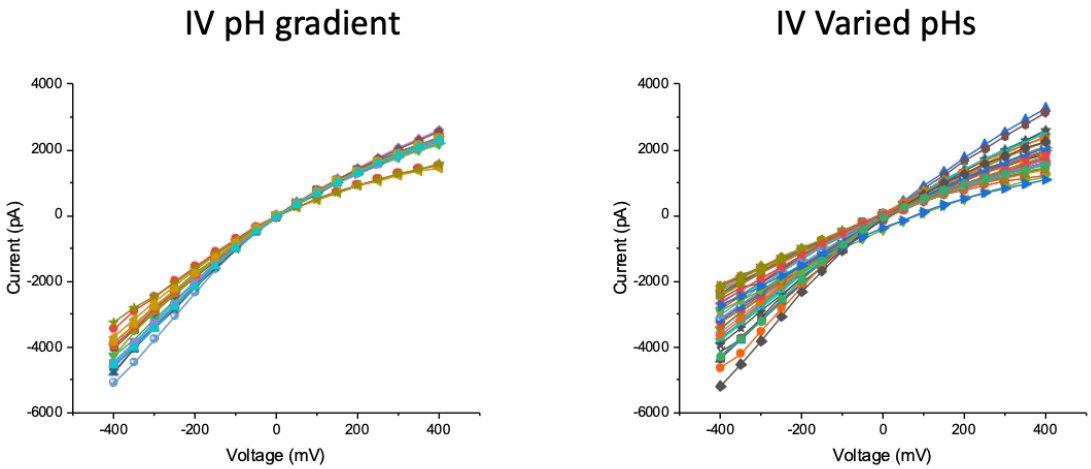


***Figure 4.26**: I-Vs for both the pH gradient experiment pipettes and the changing pH experiments.*

An additional set of controls that were performed for this experiment were current traces for the buffer at various pHs alone. With only the buffer, at all the different pHs, there were no events in the current trace, suggesting that all the events are the plasmid or plasmid with protein.

CONFIRMATION OF STREPTAVIDIN BINDING

4.15 BIOANALYSER RESULTS

In prior sections of this chapter, the inability to confirm whether streptavidin, monovalent or quadrivalent, was able to bind successfully to the plasmid carrier was discussed. As the agarose gel is not sensitive enough to determine binding, several other

methods were attempted. The first of which was using a Agilent 2100 Bioanalyser. This bioanalyser conducts electrophoresis on a chip where the sample travels through microchannels and thus is able to both use smaller sample sizes and identify smaller differences than gel electrophoresis. For these experiments, the DNA 12000 kit was used, which has a detection range of 100 bp to 12000 bp. A comparison of PC2ELigfilB, PC2ELigBOphfilB, and PC2ELigBO2filB was run, shown in **Figure 4.27**. The concentration was diluted to 50 ng/µL for each, and both the 5 Kbp DNA and 10 Kbp DNA controls and the ladder runs showed up successfully in the bioanalyser results. Unfortunately, while all three samples showed up, there was no apparent difference between them, they all showed up right around 10 Kbp DNA. Thus another method of confirmation, fluorescent binding on glass slides was performed.

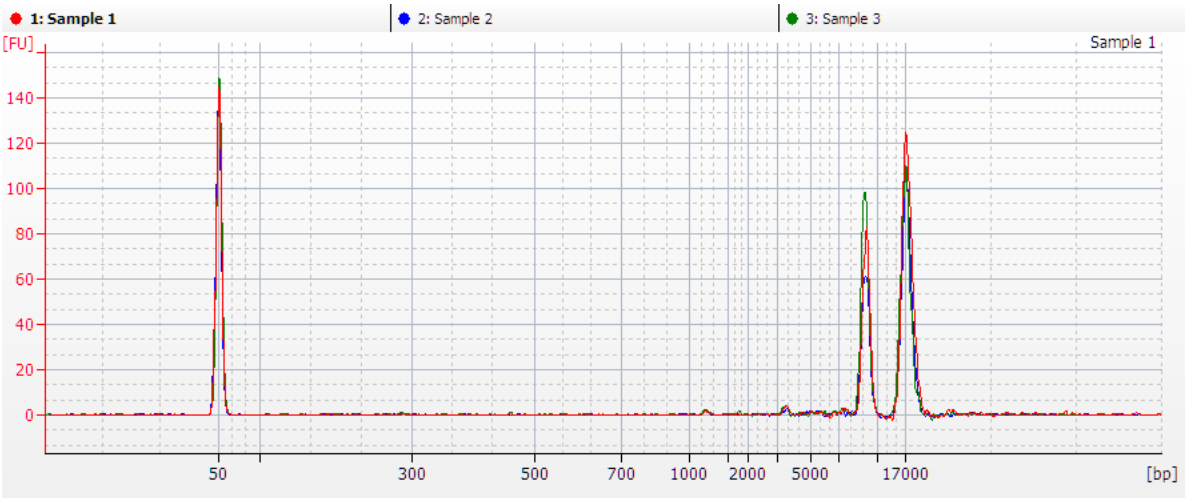


***Figure 4.27**: Bioanalyser results for a comparison of PC2ELigfilB, PC2ELigBOphfilB, and PC2ELigBO2filB. The controls at 5 Kbp, 10 Kbp, and the ladder all showed up successfully. The protocol from the Agilent DNA 12000 kit was followed.*

4.16 GLASS SLIDE PROTOCOL

The glass slide fluorescent binding experiment was adapted from Ha *et al.* (1999).[14] The general steps include binding biotinylated-BSA (Thermo Scientific) to a glass microscope slide, then incubating it with quadrivalent streptavidin (NEB). Thus the slide has quadrivalent streptavidin bound to its surface with clear binding sites for another biotin. The plasmid carriers were incubated with yoyo-1 (Invitrogen), so that they become fluorescent and are easily visible when the glass slide is imaged using a scanning confocal. As a control, half of the slide was covered with Kapton tape so that nothing was able to bind on that side of the slide.

The binding of the biotinylated-BSA is done by pipetting 100 μL of biotinylated-BSA in a 10 mM Tris, NaCl 50 mM pH 8.0 buffer (buffer A) onto the glass slide, and then letting it incubate for 10 minutes. Buffer A is then used to wash the biotinylated-BSA off and 100 μL of quadrivalent streptavidin in buffer A is added to the glass slide and left to incubate for 10 minutes. The washing with buffer A is repeated before the carrier+yoyo sample is added and incubated for 10 minutes. The slide is then washed with buffer A twice. A scanning confocal, with 60X magnification and 1 mW laser power is used to image the slides. All scanning confocal images were taken by Annie Sahota.

4.17 GLASS SLIDE RESULTS

All six samples of importance, PC2ELigfilB, PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigfilA, PC2ELigBOfilA, PC2ELigBO2filA, were imaged. Each sample was added to multiple slides and recorded. For each slide, ten sections of ten frames were measured to determine the mean fluorescence. This was done for the tape and no tape sides. As expected the tape sides had considerably less fluorescence and the lowest value from each frame was averaged to find the mean background. This value was then subtracted from the fluorescence values for the no tape sides so that the different samples could be compared. **Figure 4.28** shows the results of this analysis. The four samples that would be expected to have fluorescence, PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, and PC2ELigBO2filA, all have considerably higher levels of fluorescence. This suggests that these samples were able to bind and thus have incorporated the oligonucleotide successfully. The blank slide, the biotinylated-BSA and quadrivalent streptavidin slide, and all of the tape sides have very low fluorescence levels. The only other sample that has any significant levels of fluorescence is PC2ELigfilA. Although it has much lower fluorescence than the samples expected to bind, PC2ELigfilA's fluorescence could come from the slides not being washed sufficiently. When all of the plasmid is not washed off, areas of fluorescence can appear as the slide dries. When there is still a small layer of buffer, the unbound plasmid samples float around in the buffer, but once the slide dries, they stop moving and appear bound. However, given that this level of fluorescence is much less than all of the other samples that were expected to bind, it is reasonable to suggest that the plasmid is modified correctly and able to incorporate the biotin oligonucleotide.

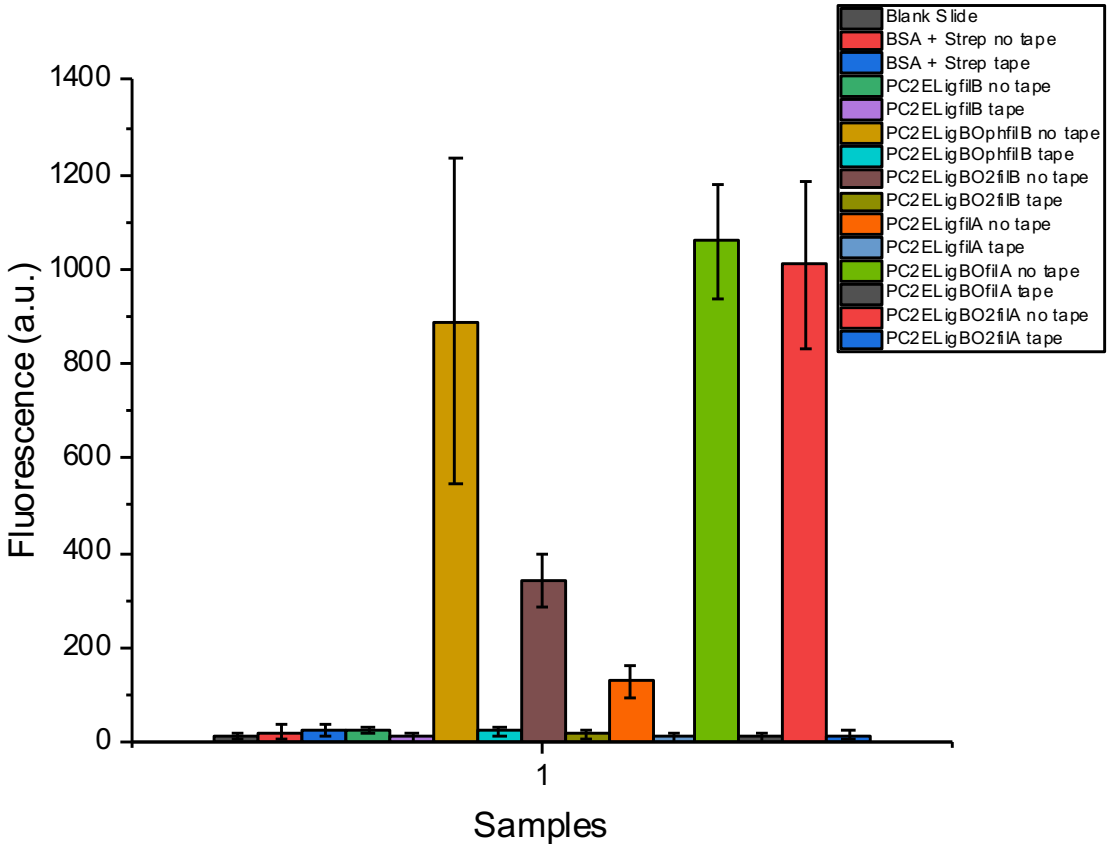


***Figure 4.28**: Bar graph comparing the levels of fluorescence for various different carrier samples and controls. The background fluorescence was subtracted for each sample and the level of fluorescence is a mean value taken from multiple slides. The controls, the blank slide, biotinylated-BSA streptavidin slide, and the tape sides of the slides, all do not have significant levels of fluorescence, suggesting that any fluorescence must come from the samples. PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, and PC2ELigBO2filA all have higher levels of fluorescence than any of the other samples.*

In addition to seeing the different levels of fluorescence, it is possible to see some difference in the binding for the samples with BO/BOph and the samples with BO2. This is shown quite clearly in **Figure 3** in **Appendix I**.

CONCLUSION

This chapter summarises the specific plasmid and modifying protocol that were chosen. It discusses the variations to the protocol for modifying the carrier as well as the final protocol. The typical nanopore behaviour of the modified plasmid was also shown and compared to the behaviour of 10 Kbp DNA. It was demonstrated that the plasmid behaves

similarly to 10 Kbp DNA after all modification steps. Additionally, the effect of pH on both the carrier and the carrier+protein was discussed. There was no definitive trend in the dwell time or peak amplitude for either a pH gradient or the varying of the pH. However, the dwell time was longer as the pH was increased, if the pH of the sample and bath were the same. Several experiments to confirm the incorporation of the oligonucleotides were then discussed. The bioanalyser was not able to show any difference between the samples, but the levels of fluorescence in the glass slide experiments were significantly different, such that it is possible to say that the oligonucleotide has been successfully incorporated into the plasmid. However, the glass slide experiment could be improved upon by using two colour fluorescence imaging. This could even allow for quantification of the oligonucleotide incorporation levels.

REFERENCES

1. Howarth, M. *et al.* A monovalent streptavidin with a single femtomolar biotin binding site. *Nat. Methods* **3**, 267–273 (2006).
2. Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, W. P. *Molecular Biology of the Cell*. (Garland Science, Taylor & Francis Group, 2014).
3. Buehler, Eugene, Cao, Qing, Kibbe, Warren A., Lurie, R. H. Oligo Calc: Oligonucleotide Properties Calculator. *Northwestern.edu* (2022). Available at: http://biotools.nubic.northwestern.edu/OligoCalc.html.
4. Dundas, C. M., Demonte, D. & Park, S. Streptavidin-biotin technology: Improvements and innovations in chemical and biological applications. *Appl. Microbiol. Biotechnol.* **97**, 9343–9353 (2013).
5. Quadrivalent Streptavidin. *Protein Data Bank in Europe*
6. NEB. Monarch DNA Gel Extraction Kit.
7. QIAquick Gel Extraction Kit. *QIAGEN* (3022).
8. Rapid DNA Ligation Kit Protocol. *Sigma Aldrich* Available at: https://www.sigmaaldrich.com/GB/en/technical-documents/protocol/genomics/cloning-and-expression/rapid-dna-ligation-kit.
9. Phosphorylation Modifications. *IDT*
10. Gong, X. *et al.* Label-free in-flow detection of single DNA molecules using glass nanopipettes. *Anal. Chem.* **86**, 835–841 (2014).
11. Wanunu, M. Nanopores: A journey towards DNA sequencing. *Phys. Life Rev.* **9**, 125–158 (2012).
12. Haywood, D. G., Saha-Shah, A., Baker, L. A. & Jacobson, S. C. Fundamental studies of nanofluidics: Nanopores, nanochannels, and nanopipets. *Anal. Chem.* **87**, 172–187 (2015).
13. Yeh, L. H., Zhang, M. & Qian, S. Ion transport in a pH-regulated nanopore. *Anal. Chem.* **85**, 7527–7534 (2013).
14. Ha, T. *et al.* Ligand-induced conformational changes observed in single RNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9077–9082 (1999).

Chapter 5 Plasmid Carrier Protein Binding

Table of Contents

Synopsis

This chapter discusses the data analysis to identify subpeaks from the plasmid events. Initially, this analysis was done manually, but then it was switched to an automated process. There is a detailed description of how the automated process works and a comparison of it to the manual results for several sample pipettes. The cis-to-trans results with monovalent streptavidin are shown, including the binding curves which show that the plasmid with one or two biotins is able to bind to monovalent streptavidin successfully. A sandwich-like assay was explored using quadrivalent streptavidin and biotinylated phosphatase. Lastly, the results for the trans-to-cis direction with the monovalent streptavidin with one and two biotins are discussed.

BACKGROUND AND EXPERIMENTAL CONDITIONS

5.1 SELECTION OF BIOTIN AND STREPTAVIDIN

Streptavidin was chosen for the proof of concept for several reasons. Firstly, the bond between biotin and streptavidin is one of the strongest noncovalent bonds, with a K_d of approximately 10^{-14} M, and is very stable over various conditions, including pH and temperature.[1] Monovalent streptavidin was used instead of wildtype quadrivalent streptavidin, as three of the binding sites are inactive, so multiple carriers cannot bind to the same monovalent streptavidin. Additionally, the size is similar to quadrivalent streptavidin, with only a minor decrease in binding strength.[2] The biotin and streptavidin bond has been used for proof of concept for other nanopore DNA carriers before with good results.[3] Thus, it was a logical choice to show the effectiveness of this carrier to bind to a specific protein.

5.2 BINDING BUFFER SELECTION AND INCUBATION TIME

Several binding buffers and incubation times were explored, but the one that yielded the best results was a phosphate buffer (50 mM phosphate, 100 mM Na^+, pH 8.0), where the monovalent streptavidin or quadrivalent streptavidin were incubated with the plasmid carrier at room temperature for an hour.

All of the experimental current recording conditions were kept the same for each binding curve and all of the controls. The plasmids' or controls' concentration was kept constant at 100 pM and the recording conditions were always the same with 100 mM KCl TE pH 8.0 for cis-to-trans and 1 M KCl TE pH 8.0 for trans-to-cis, the filter at 30 KHz, the voltage at

-300 mV for cis-to-trans and 300 mV for trans-to-cis, and the sampling rate at 100 KHz. The monovalent streptavidin or quadrivalent streptavidin concentrations for each binding curve were 50pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM, and 100 nM; these are referred to as 0.5X, 1X, 10X, 50X, 100X, 250X, and 1000X respectively. These concentrations were kept constant for both oligonucleotides, i.e. whether there was one biotin or two. Additionally, the binding curves were recorded in order of increasing protein concentration. For the later sandwich-like assay experiments, the quadrivalent streptavidin was incubated with the plasmid carrier as normal, that is at the previously mentioned ratios and conditions, and then the biophosphatase was added and incubated for another hour with the same phosphate-binding buffer. The biophosphatase's concentration was always one third of the concentration of the quadrivalent streptavidin, such that it was less likely for multiple biophosphatases to bind to the same quadrivalent streptavidin.

DATA ANALYSIS

5.3. MANUAL DATA ANALYSIS

The initial confirmation of the binding was done using a manual method. The events were identified using a custom Matlab script and then examined individually one by one for the presence of subpeaks. A subpeak was identified by looking at the shape of the event and determining if there was an additional peak. Other characteristics, such as peak height, charge, and dwell time were also considered when concluding a subpeak's presence. Ideally, the subpeaks should appear in the middle of the event, as the biotin is attached to the middle of the linearised plasmid, although this did not always seem to be the case. There was a clear difference in the event shape, with the plasmid alone events appearing more triangular. Some sample events, with and without subpeaks are shown in **Figure 5.1** for cis-to-trans events. A summary of the naming convention for the plasmid samples are shown in **Tables 4.1** and **4.2**. Event a is PC2ELigBOfilA and has a triangular shape. Events b and c are 50X (monostreptavidin incubated with PC2ELigBOfilA). These both have clear subpeaks, in addition to longer dwell times and higher peak amplitudes. It is this subpeak shape that was selected for in both the manual and automated analysis.

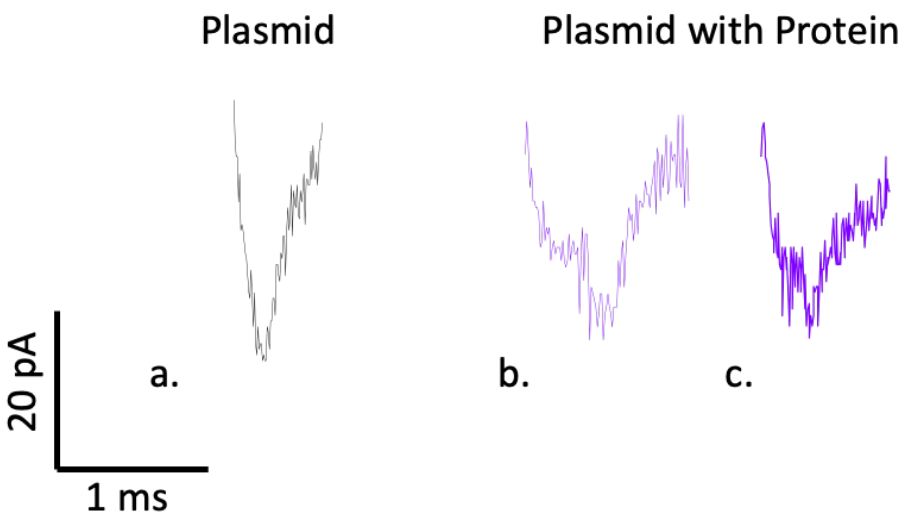


***Figure 5.1**: Sample events with and without subpeaks for PC2ELigBOfilA. Event a., in black, is without protein and events b. and c., in purple, are with protein. The latter two events have clear subpeaks, and this is the shape of event that was looked for in the manual analysis and then in the automated analysis.*

The results did seem promising, as there is a significant increase in events with subpeaks for the samples with protein compared to the samples without. However, while it is possible to identify subpeaks manually, the process is quite arduous, particularly as many pipettes have several thousand events, and allows for fluctuations in reproducibility. Therefore, an automated process using the same custom Matlab code was explored.

5.4 AUTOMATING THE DATA ANALYSIS PROCESS

In the Matlab code, there are several important parameters. These are setting the current trace (baseline) and threshold, the event parameters, including event selection and shape, and the subpeak specific parameters. A flow chart with an overview of the steps conducted for the subpeak selection process is shown in **Figure 5.2.**

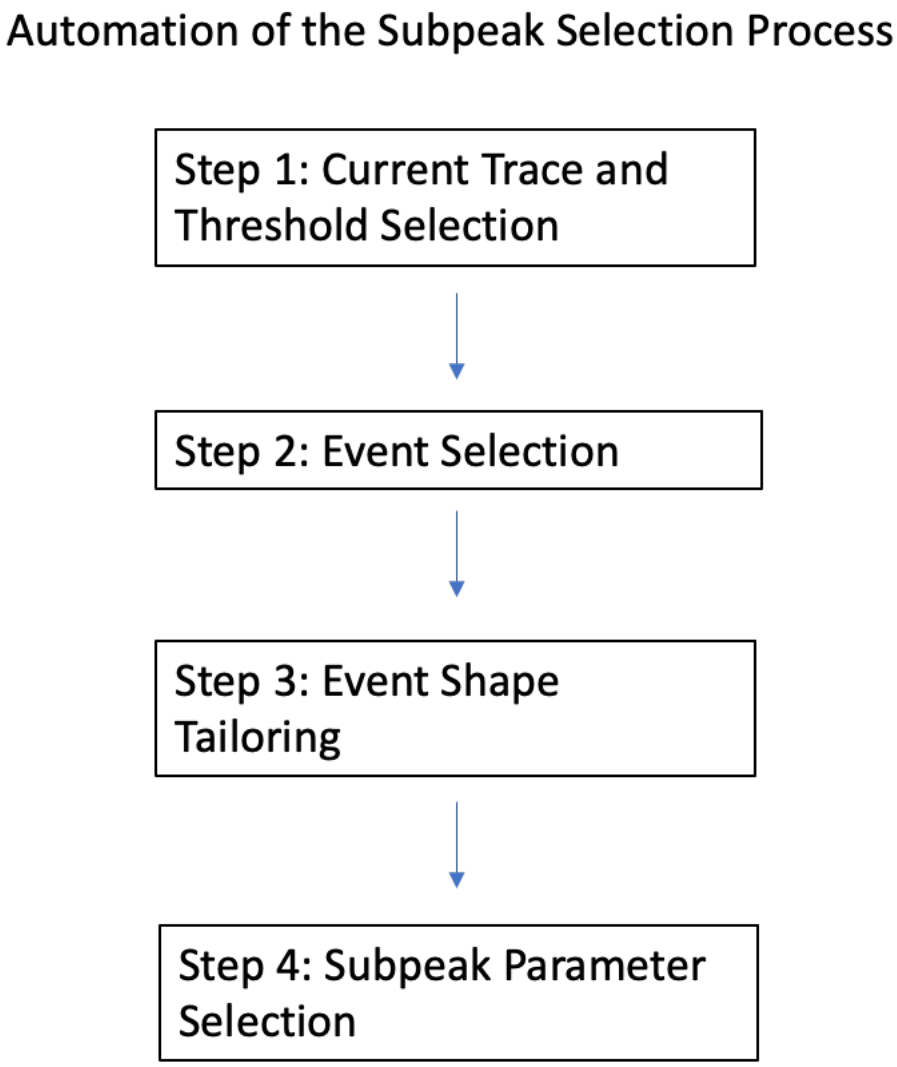


***Figure 5.2**: Flowchart showing the steps involved in the automation of the subpeak selection process.*

The current trace (baseline) and threshold, shown in **Figure 5.3**, are able to be varied from trace to trace, such that the noise was removed and the events were identified correctly. The step for the baseline was kept constant at 10, as this value allowed good fitting of the current trace, but the threshold was varied regularly to avoid selecting noise. The step offset, shown as the green line on the baseline graph and used to selected where the events start, was kept constant at 1.8. The standard deviation cut off threshold was kept constant at 7; this is shown as the black line on the baseline graph. It is possible, and indeed necessary, to vary the threshold parameters, as this allows for the selection of 'real' events, and avoids classifying noise as events. Depending on the noise levels, an override was sometimes used to set the threshold. The overall step was typically set at $1E^{-3}$ and the max

at 0.1. The override was used when the standard setting of 7 did not allow for good fitting of the threshold graph, shown with the red line. This tailoring for different samples and pipettes allowed for the best event fitting.

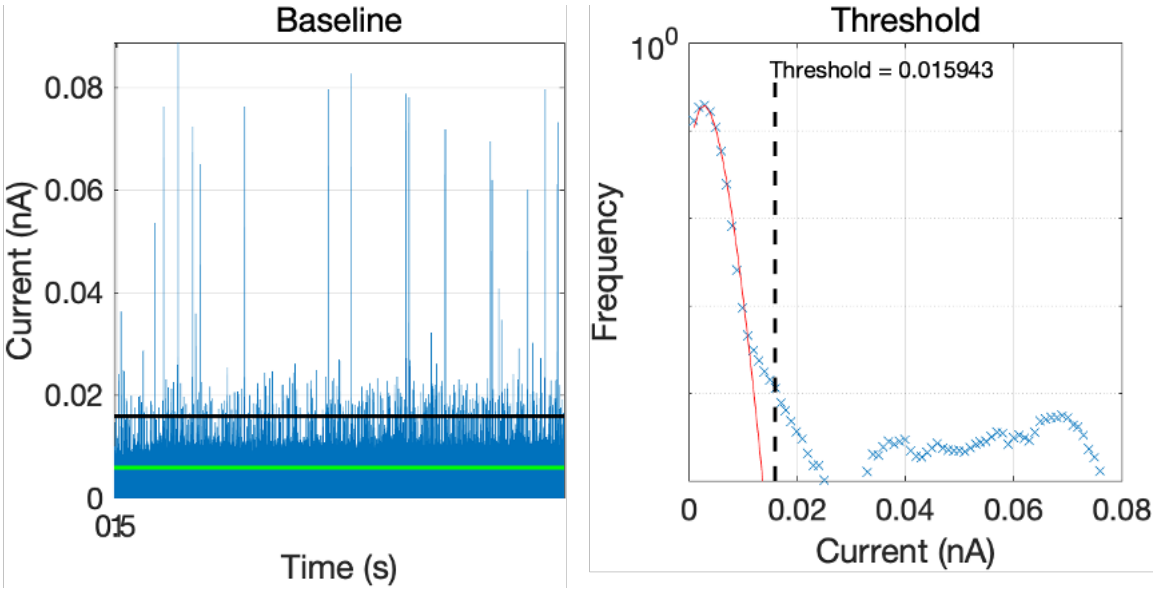


***Figure 5.3**: Custom Matlab code settings for selecting the stepsize and the threshold for the current traces. The step was kept constant at 10 as this allowed for good fitting of the current trace. The threshold varied more, in order to allow adequate fitting and to avoid classifying noise as events. The threshold fitting for all the current traces typically looked like the graph shown above. The threshold override setting was sometimes used as it allowed for better fitting.*

Figure 5.4 shows the next step in the data analysis, the selection of the events. The parameters used for this were not varied between current traces. An arbitrary maximum is imposed at 10 ms as events longer than this were assumed to be caused by blockage of the pore rather than DNA or DNA+protein moving smoothly through it. Similarly, although times down to ~0.03 ms are able to be resolved due to the 30 KHz filter and 100 KHz sampling rate, 0.1 ms was picked as the minimum, and events with times below this were assumed to be caused by DNA or DNA+protein bumping into the pore or partially entering it. There are further event parameters that allow for a more tailored selection of the events; these include approx. delta, sigma, and max levels. The approx. delta varies the step size of the levels used in the event fitting, the amplitude. Sigma was set according to the noise levels to avoid selecting events incorrectly; a lower sigma is used for traces with more noise. The max levels are the maximum number of levels that can be used to fit an event. The more levels there are the closer the fitting will follow the event trace, but the more subpeaks will be

selected incorrectly. Thus, it is necessary to find a max level value that allows for overall good fitting of the event without selecting too many subpeaks. While values of the event parameters were varied while tailoring the automated process, once values that fit the events well were found, they were kept constant. This was especially important as good fitting allowed for more accurate subpeak selection, and variation between parameters, particularly max levels, caused too many differences in subpeak selection between samples. Accurate fitting of the events also allows for a comparison of the charge. A difference in charge between samples with and without protein suggests that the protein has bound rather than the DNA simply folding and creating subpeaks.

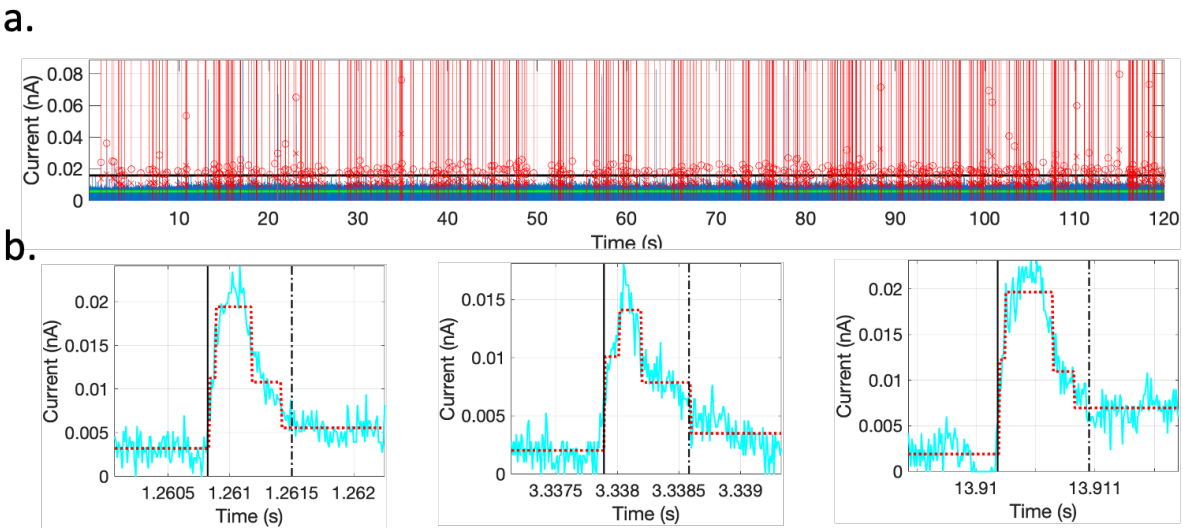


***Figure 5.4**: Selected events are shown in a. Each red line is an event from the trace with a cross showing the mean events of the peak and the max peak amplitude of the event is designated by the circle. The event parameters were kept constant for all current traces. An arbitrary minimum and maximum of 0.1 ms and 10 ms respectively were picked, as events with times lower than this were assumed to be DNA bumping the pore and longer were assumed to be DNA blocking the pore. The event parameters are used to ensure that the code traces the events accurately. Once ones that allowed for good fitting were found, it was necessary to keep them constant, as otherwise subpeak selection varied too much. Sample event fittings are shown in b.*

The last set of parameters used for the automated subpeak selection are the subpeak specific parameters. These allowed for selection based on normalised peak height, peak height, subpeak height, subpeak width, fractional width, number of subpeaks, and total event width. Normalised peak height normalises the peak height by dividing the second level by the first. This parameter was not used for any of the final data, but was

explored in the tailoring of the process. The peak height is the difference between the highest current level and the baseline. Subpeak height is the difference between the peak current when the subpeak begins and the maximum current level. Subpeak width is the dwell time of the subpeak. Fractional width is the portion of the event that corresponds to the subpeak. This allowed the removal of events that were incorrectly classified as having a subpeak, when the subpeak took up a significant fraction of the event. When tailoring the parameters, >0.5 was often used for fractional width. Number of subpeaks allows for the selection of events with one or more subpeaks. Total event width is the dwell time. **Figure 5.5** shows what these parameters look like.

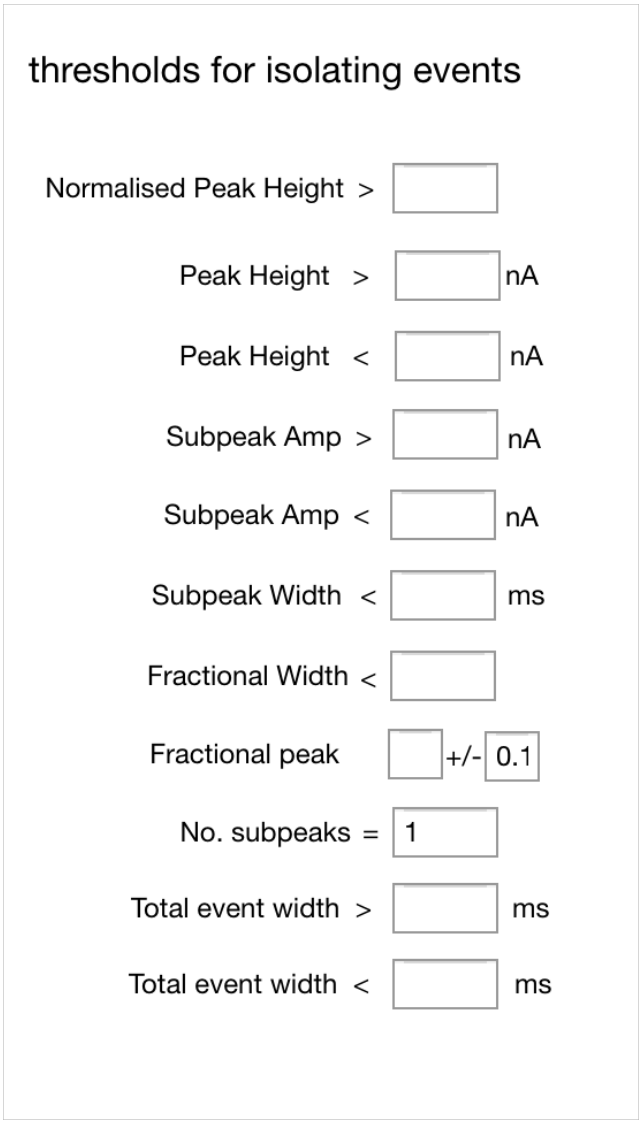


***Figure 5.5**: Subpeak parameter options. These were used to select events that had subpeaks from peaks that were deemed to be events.*

The ability to select between these characteristics of the events was necessary as differences in the averages of dwell time and peak height between the plasmid alone and plasmid with protein samples were used to select for subpeaks. Many of these subpeak parameters were explored and compared with the manual selections. **Table 5.1** shows four different ones, subpeak width <0.5 ms, frac width <0.5, subpeak amp >0.007 nA, and normalised peak height 2. The samples compared in this table are a PC2ELigBOphfilB and 50X with quadrivalent streptavidin going cis-to-trans. The agreements between the automated and the manual analysis were not particularly good, especially as one of the parameters only had 65% agreement. Additionally, for several of the parameters, there is a high rate of false negatives, sometimes up to almost 30%. Also, the percentages of the folded vs. subpeak events do not always show the subpeak percentage higher than the folded, which is the expected result. Thus, another avenue had to be explored.

	PC2ELigB Ophnew Agreement	**PC2ELigB Ophnew False Positives**	**PC2ELigB Ophnew False Negatives**	**50X Agreement**	**50X False Positives**	**50X False Negatives**	**Folded Percentage**	**Subpeak percentage**
No subpeak parameters	78.3%	21.3%	0.3%	84.0%	4.0%	12.0%	54.1%	40.5%
Subpeak width <0.5ms	78.3%	21.3%	0.3%	61.3%	1.7%	37.0%	54.1%	16.3%
Frac width < 0.5	82.3%	14.0%	3.7%	72.0%	1.7%	26.7%	44.4%	27.3%
Subpeak amp > 0.007nA	65.0%	6.3%	28.7%	71.0%	3.7%	25.3%	13.6%	28.9%
Normalised Peak Height 2	67.7%	11.0%	21.3%	63.0%	2.0%	35.0%	24.3%	16.6%

Table 5.1*: Comparison of percentage of folded events and subpeaks using various subpeak parameters.*

As the vast majority of samples with protein varied in dwell time and peak amplitude from samples of just the plasmid, these differences were explored. Typically the 50X sample was used as a comparative sample to pick the parameters, as sometimes the samples with lower amounts of protein did not always have significant enough differences. **Figure 5.6**

shows the histograms of PC2ELigBOphfilB and 50X with quadrivalent streptavidin; the same samples used in **Table 5.1**. The dwell times, charges, and peak amplitudes for these two samples vary greatly, and for the 50Xquad sample, there are clearly multiple peaks, which is what would be expected as not all of the plasmid would have incorporated the biotin oligonucleotide or bind successfully to the streptavidin. For all three, there was also significant difference when a two way T test was conducted. Scatter plots of dwell time versus current amplitude are also shown, and these again show a clear difference in the samples with and without protein.

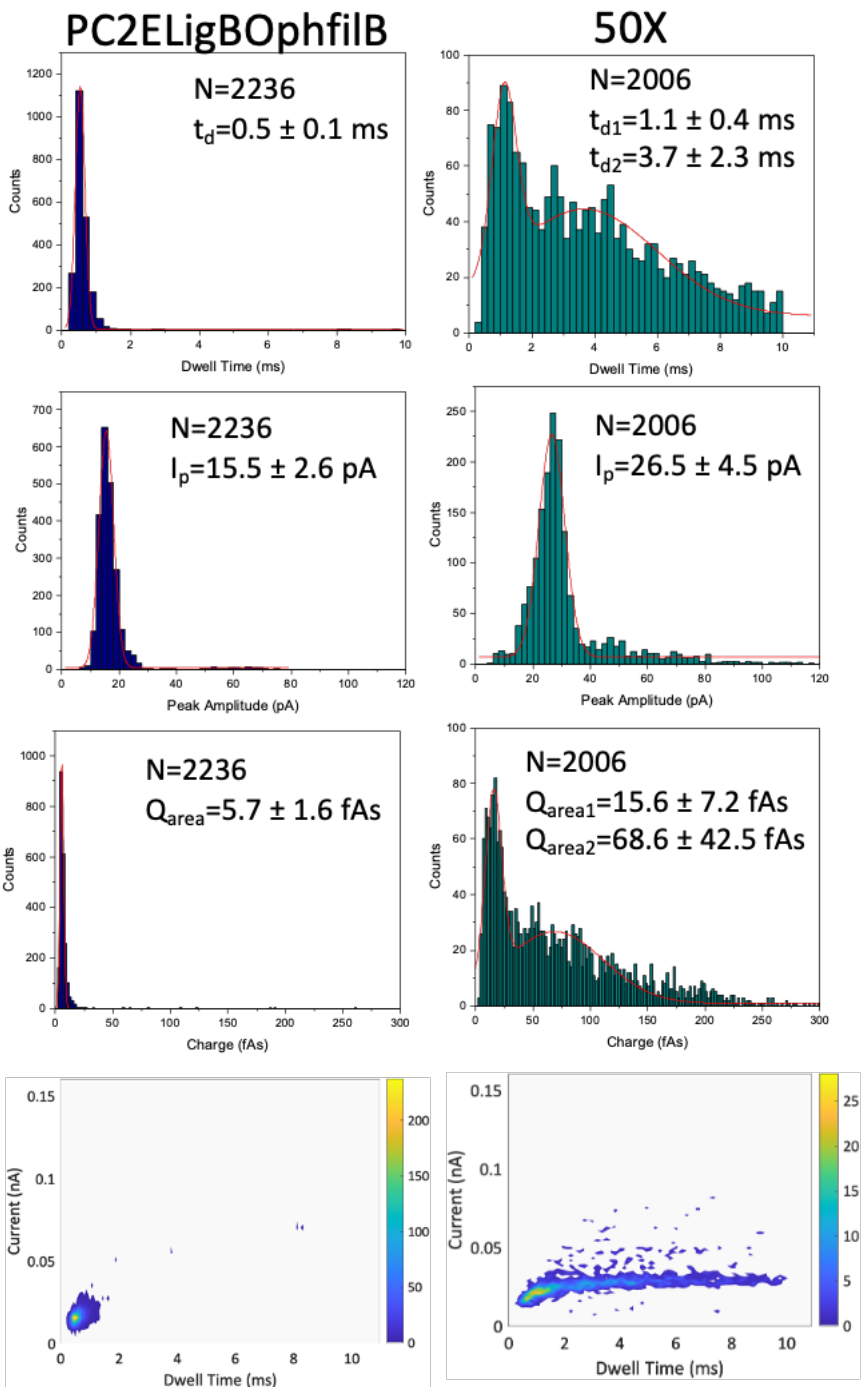


***Figure 5.6**: Histograms of the dwell time and peak amplitude of PC2ELigBOphfilB and 50Xquad quadrivalent streptavidin in cis-to-trans directions.. A Gaussian fit was used to determine the average values. A two sample T test was performed for dwell time, peak amplitude, and charge, and the means were determined to be significantly different.*

These differences in dwell time and peak amplitude are then used to select the subpeaks. As both parameters are greater in the protein sample, the subpeak parameters used are peak height > 0.015 nA and total event width > 0.5 ms. These values were chosen

from the averages of the PC2ELigBOphfilB histograms. When using these subpeak parameters, 16.7% of events are considered folded for PC2ELigBOphfilB and 43.7% of events have a subpeak for the 50X. The first 300 events were for both the manual and automated analysis were directly compared. This is shown in **Table 5.2**. Overall the agreement with the manual results was much better than the results presented earlier in this chapter, at least a 76.7% agreement between the two. Additionally, the rates of false negatives and positives were much reduced. Significantly, the percentage of subpeaks at all ratios of protein to plasmid were greater than PC2ELigBOphfilB on its own.

	Overall percentage folded/subpeak	Percentage folded/subpeak (first 300) manual	Percentage folded/subpeak (first 300) auto	Agreement with manual percentage	False negatives percentage	False positives percentage	Total Events
Ligated	16.7	21.7	25.0	76.7	10.0	13.3	2236
0.5X	38.7	33.3	25.0	82.3	13.0	4.7	3355
1X	36.1	36.0	31.3	83.3	10.7	6.0	3740
10X	43.4	39.0	44.3	76.7	9.0	14.3	3026
50X	43.7	42.0	42.7	82.0	8.7	9.3	1688
100X	45.4	44.7	47.0	85.7	6.0	8.3	1603
250X	39.1	43.3	46.0	90.7	3.3	6.0	1469
1000X	37.0	33.0	38.3	89.3	4.1	6.5	676

***Table 5.2**: A comparison of the manual and automated results, for the first 300 events, for PC2ELigBOphfilB and PC2ELigBOphfilB with various ratios of quadrivalent streptavidin.*

This comparison to the manual results was done once again, for cis-to-trans, to confirm that it is a reliable method. Additionally, it was important to confirm that it worked with the monovalent streptavidin as well. Again PC2ELigBOphfilB and 50X were compared and the histograms are shown in **Figure 5.7**. While the difference is not quite as great for the dwell time, the charge, and the peak amplitude as with the quadrivalent streptavidin sample, there is still clearly a difference between the PC2ELigBOphfilB and the 50X sample. For all three, there was also significant difference when a two way T test was conducted. Additionally, it again looks like the 50X has two populations in the dwell time and charge data. While the scatter plots do not look as different as the quadrivalent streptavidin sample, it does seem like the sample with protein is shifted both up and to the right compared to the sample without protein.

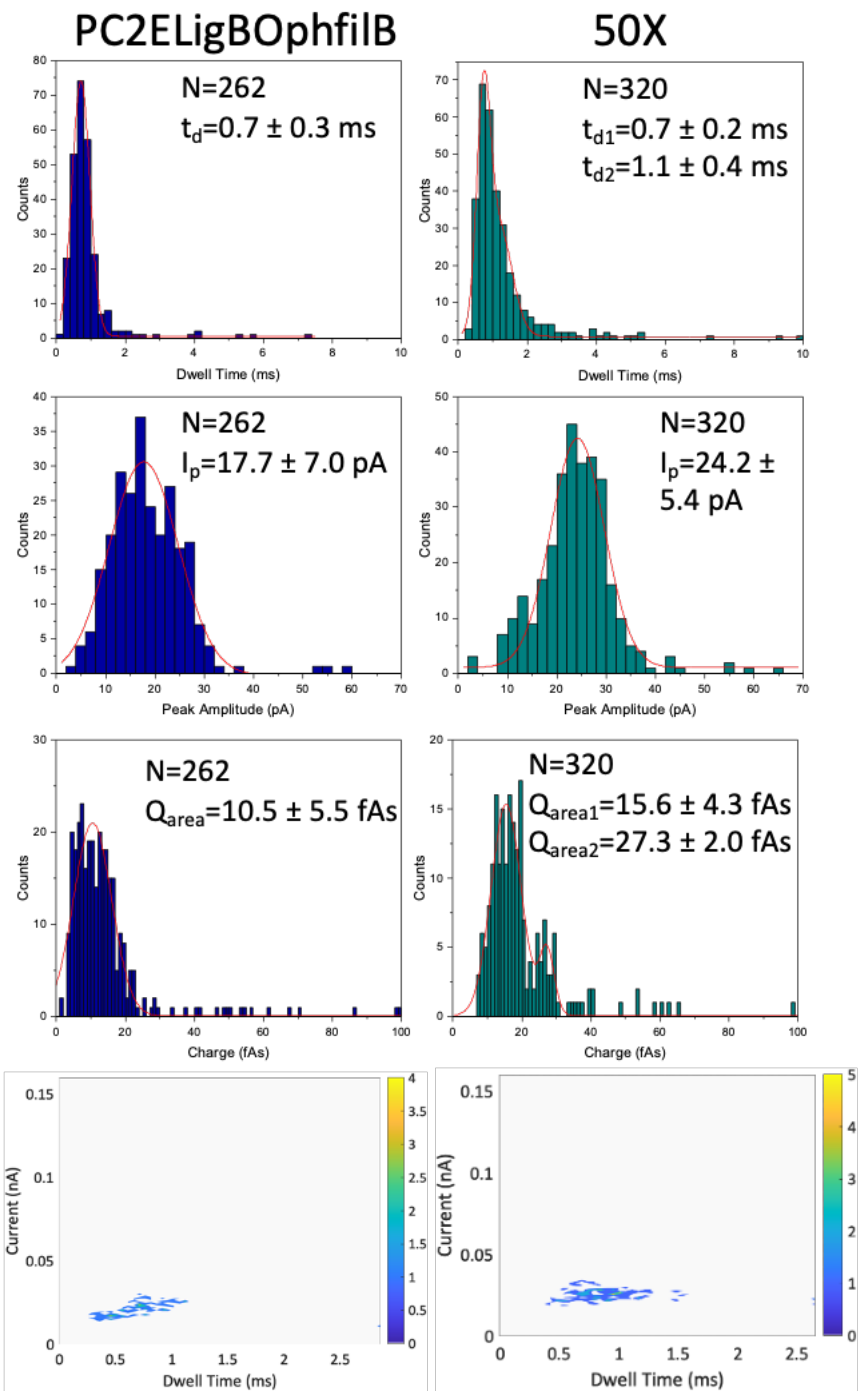


***Figure 5.7**: Histograms of the dwell time, charge, and peak amplitude of PC2ELigBOphfilB and 50X monovalent streptavidin in cis-to-trans direction A Gaussian fit was used to determine the average values. A two sample T test was performed for dwell time, peak amplitude, and charge, and the means were determined to be significantly different.*

The differences between the dwell time and peak amplitude were again used to set the subpeak parameters at peak height > 0.015 nA and total event width > 0.5 ms. The manual comparison had at least 87% agreement between manual and automated.

Additionally, the false positives and negatives rates were low, and the overall percentages of subpeaks in the samples with protein were typically larger than the percentage of folded events with PC2ELigBOphfilB. The full results are shown in **Table 5.3**. Thus, it was deemed acceptable to use this method to set the subpeak parameters for the cis-to-trans data.

	Overall percentage folded/subpeak	Percentage folded/subpeak (first 100) manual	Percentage folded/subpeak (first 100) auto	Agreement with manual percentage	False negatives percentage	False positives percentage	Total Events
Ligated	24.0	14.0	17.0	93.0	2.0	5.0	262
0.5X	39.3	43.0	50.0	91.0	1.0	8.0	262
1X	49.1	43.0	49.0	90.0	2.0	8.0	171
10X	37.2	33.0	35.0	88.0	5.0	7.0	266
50X	41.8	41.0	44.0	87.0	5.0	8.0	320
100X	35.9	38.0	42.0	94.0	1.0	5.0	306
250X	21.3	23.0	24.0	95.0	2.0	3.0	150
1000X	45.6	31.0	33.0	88.0	5.0	7.0	388

***Table 5.3**: A comparison of the manual and automated results, for the first 100 events, for PC2ELigBOphfilB and PC2ELigBOphfilB with various ratios of monovalent streptavidin.*

As trans-to-cis was also explored, this automated subpeak method was confirmed for this direction as well. The samples used in **Figure 5.8** are PC2ELigBO2filB and 50X, where the 50X was incubated with monovalent streptavidin. In addition to confirming it worked with trans-to-cis, these results confirm that it worked with samples with the oligonucleotide with two biotins. The dwell times and charges seem to vary between the samples, although the peak amplitudes are similar. However, even with just using the potential dwell time variation, it was possible to successfully select for subpeaks. Also, as there are fewer events, the scatter plots do not show such clear changes, but it is possible to see some shift in dwell time between the samples. Further, while the averages are not very different, there are a significant number of events that have higher charges in the sample with protein. For, dwell time and charge there was also significant difference when a two way T test was conducted.

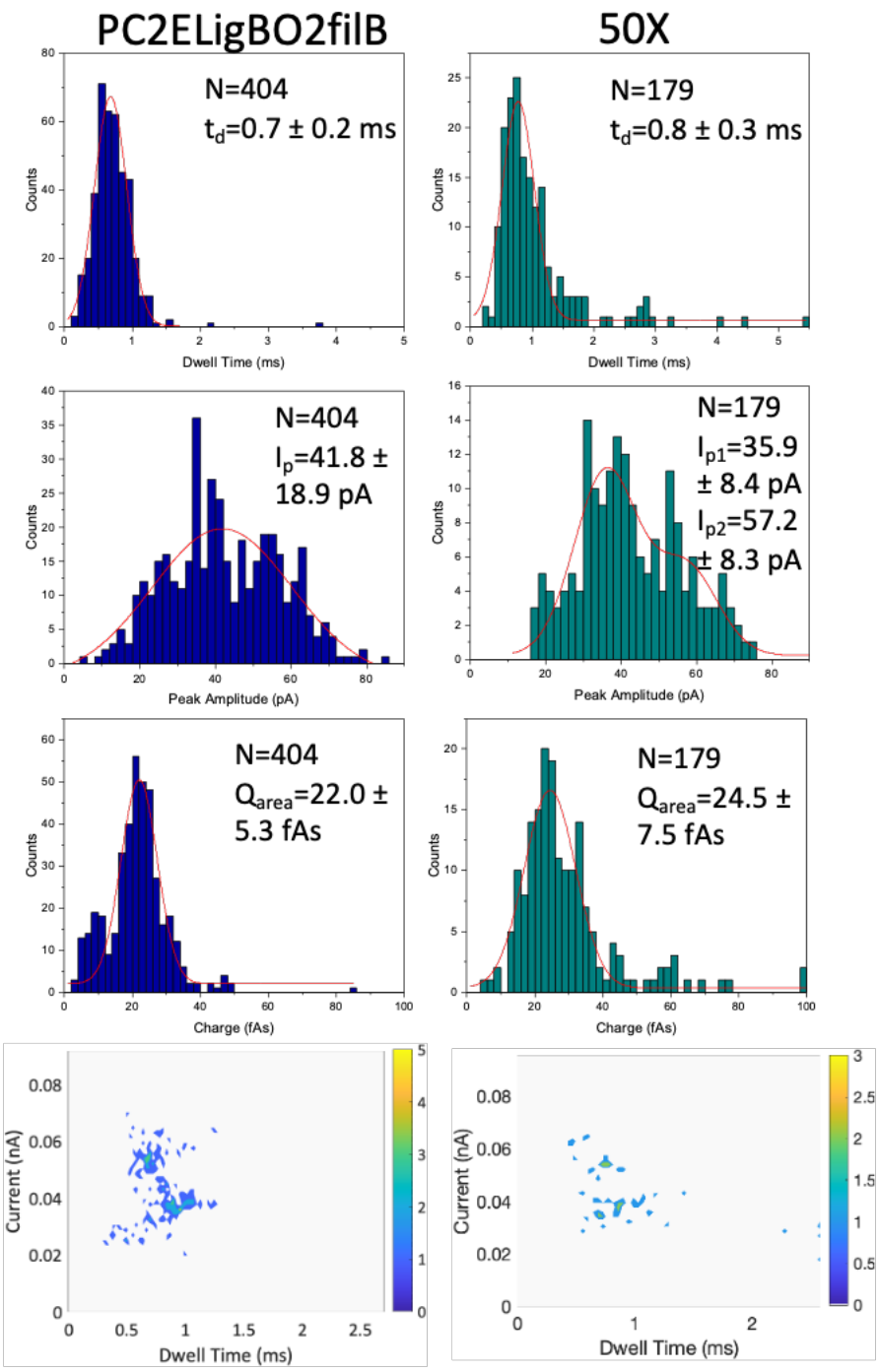


***Figure 5.8**: Histograms of PC2ELigBO2filB and 50X with monovalent streptavidin in the trans-to-cis direction. A Gaussian fit was used to determine the average values. A two sample T test was performed for dwell time, peak amplitude, and charge, and the means for dwell time and charge were determined to be significantly different, but not for peak amplitude.*

The subpeak parameter used to select for subpeak events was total event width > 0.6 ms. The first 100 events were compared to the manual results, shown in **Table 5.4**, and suggest that this method has been able to select subpeaks successfully. There is good event agreement, at least 84%, with the manual results and the false negatives and positives are

low. Additionally, while the overall percentages of subpeaks are not as drastically different as with the sample cis-to-trans examples, there is still a percentage increase for most of the protein samples.

	Overall percentage folded/subpeak	Percentage folded/subpeak (first 100) manual	Percentage folded/subpeak (first 100) auto	Agreement with manual percentage	False negatives percentage	False positives percentage	Total Events
Ligated	16.8	20.0	17.0	89.0	7.0	4.0	404
0.5X	21.3	25.0	19.0	90.0	8.0	2.0	225
1X	22.5	28.0	21.0	89.0	9.0	2.0	209
10X	18.6	25.0	20.0	89.0	8.0	3.0	220
50X	34.6	31.0	35.0	86.0	5.0	9.0	179
100X	25.1	30.0	23.0	89.0	9.0	2.0	147
250X	26.8	30.0	25.0	89.0	8.0	3.0	168
1000X	26.8	30.0	24.0	88.0	9.0	3.0	168

Table 5.4*: A comparison of the automated and manual results for trans-to-cis for the first 100 events.*

The successful use of this method for trans-to-cis was confirmed with another comparison of PC2ELigBO2filB and 50X with monovalent streptavidin. Similarly to the other samples explored earlier in this chapter, it was the dwell time differences that were used to separate for the subpeaks. **Figure 5.9** shows the histograms results. While the charge averages are not very different, there is much more variation for the sample without protein, suggesting that there are some significant differences between the charges. There were not enough events to do two way T tests to confirm for significant difference. The scatter plots also seem quite different.

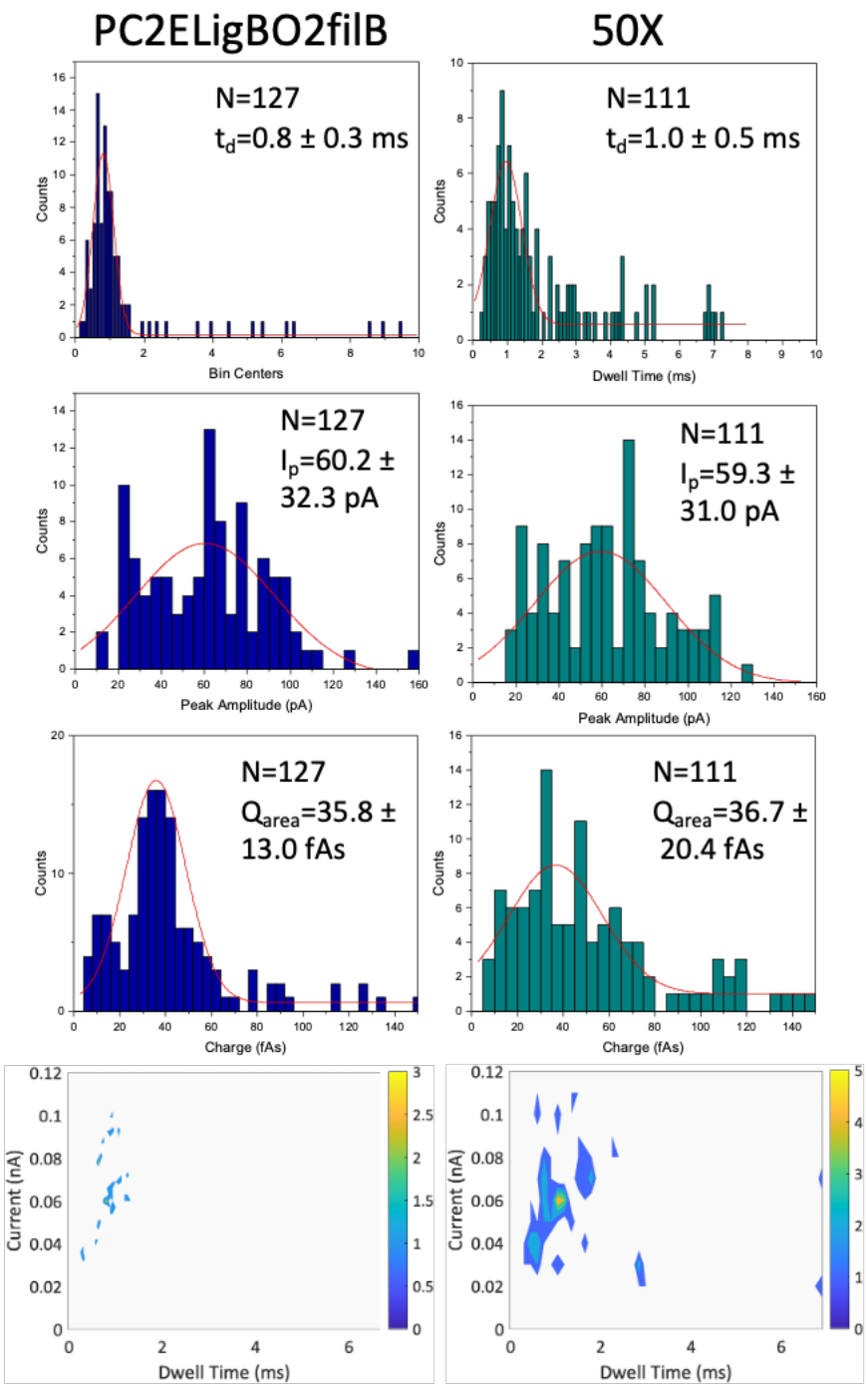


***Figure 5.9**: Histograms of PC2ELigBO2filB and 50X with monovalent streptavidin in the trans-to-cis direction. A Gaussian fit was used to determine the average values. There were not enough events for this sample to conduct a two sample T test.*

The results were compared to the manual ones, and a summary of these results is shown in **Table 5.5**. The subpeak parameter used to select for subpeak events was total event width > 0.8 ms. While the higher ratios of protein have good results, the lower ratios did not have many subpeaks. This could perhaps be due to the incubation not working

correctly or the pipette being partially blocked. The latter seems more likely as the results for 10X were not able to be analysed as they were too noisy. However, the results for 50X and greater suggest that this automated subpeak method is still valid for trans-to-cis. The overall agreement with the manual results were generally high, and the rates of false negatives and positives were quite low. Additionally, the differences between the overall percentages of the subpeaks for the protein ratios greater than 10X were higher than the folded events for PC2ELigBO2filB.

	Overall percentage folded/subpeak	Percentage folded/subpeak (first 100) manual	Percentage folded/subpeak (first 100) auto	Agreement with manual percentage	False negatives percentage	False positives percentage	Total Events
Ligated	12.6	15.0	13.0	94.0	4.0	2.0	127
0.5X	6.9	22.2	6.9	81.9	16.7	1.4	72
1X	3.5	15.8	3.5	87.7	12.3	0.0	57
10X							
50X	45.0	48.0	45.0	77.0	13.0	10.0	111
100X	44.3	45.0	44.0	87.0	6.0	7.0	106
250X	34.8	44.0	36.0	82.0	13.0	5.0	138
1000X	28.5	41.0	32.0	89.0	10.0	1.0	165

Table 5.5*: Comparison of manual and automated results for PC2ELigBO2filB and PC2ELigBO2filB with various ratios of protein. 10X was too noisy to analyse.*

Once it was established that this automated method worked for both cis-to-trans and trans-to-cis, it was important to confirm that it worked with control samples. As such, four different control samples with and without protein were analysed, two cis-to-trans and two trans-to-cis, using the same criteria for subpeak selection as in the previous samples. Although it should be noted that as the dwell time, charge, and peak amplitude did not really vary between the control sample and the control sample with protein, the dwell time and peak amplitude of the control sample were used for the subpeak parameters. The control shown in **Figure 5.10** is PCNbBVLigstepsB, so it is the plasmid that has gone through the whole protocol but only was incubated with the nicking enzyme and the ligase in the cis-to-trans direction. The protein sample is the PCNbBVLigstepsB sample incubated with monovalent streptavidin in a 1:1 ratio. The dwell times, charges, and peak amplitudes for both samples are very similar. Additionally, the scatter plots show alike results.

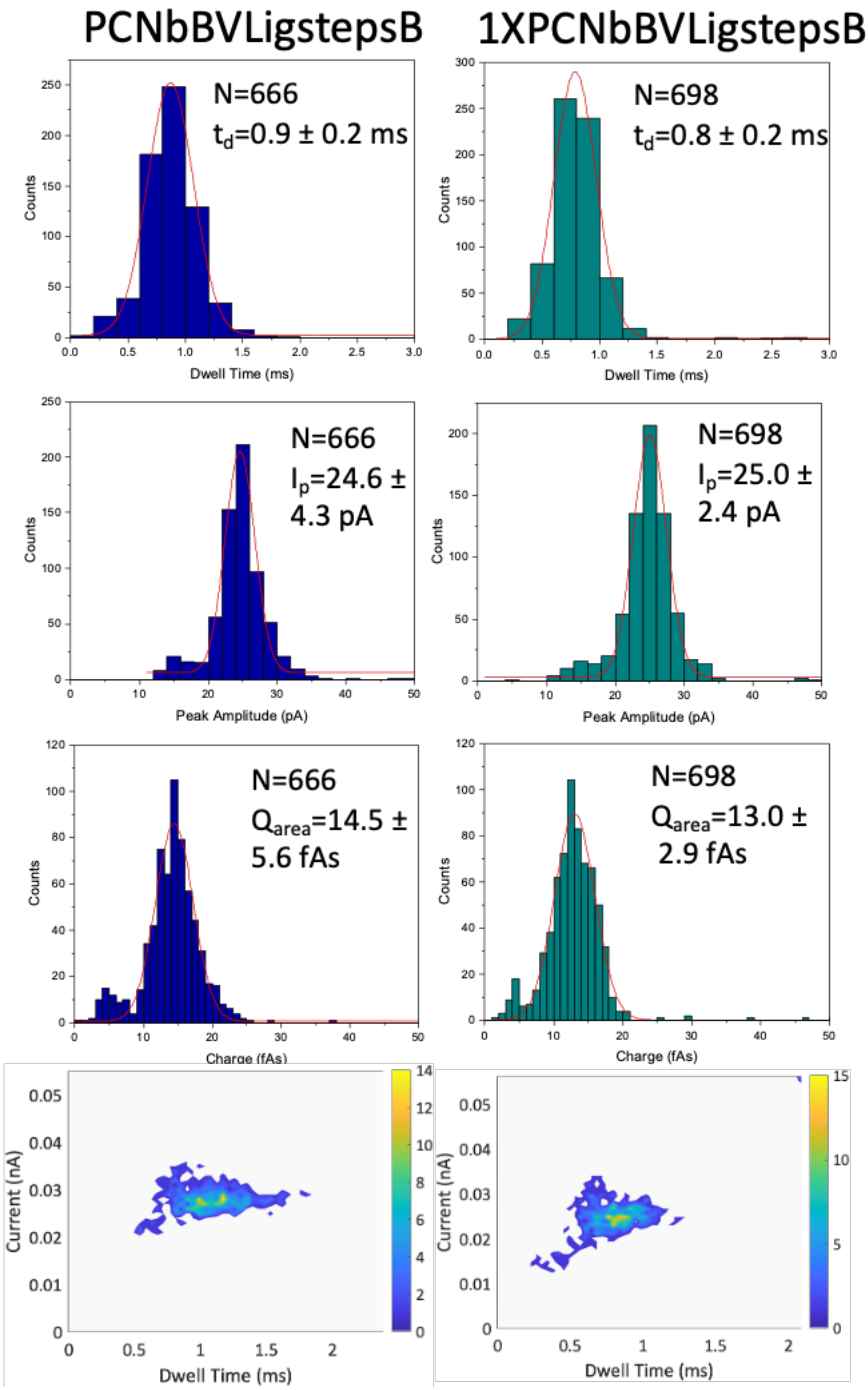


***Figure 5.10**: Histograms of the control sample PCNbBVLigstepsB, plasmid that has gone through the whole protocol, but only incubated with the nicking enzyme and the ligase, with and without protein, in the cis-to-trans direction. A Gaussian fit was used to determine the average values. A two sample T test was performed for dwell time, peak amplitude, and charge, and the mean for peak amplitude was determined to be significantly different, but not for dwell time or charge.*

This control sample was then analysed in a similar manner to the previous samples with protein and the full results are shown in **Table 5.6**. The subpeak parameters used to

select for subpeak events were peak height > 0.019 nA and total event width > 0.8 ms. It was compared to the manual results, and there was good agreement with manual and automated results, and low rates of false positives and negatives. These results, as well as the similar low overall percentages of folded/subpeak events, suggests that the protein is not able to bind to this control sample.

	Overall percentage folded/subpeak	**Percentage folded/subpeak (first 100) manual**	**Percentage folded/subpeak (first 100) auto**	**Agreement with manual percentage**	**False negatives percentage**	**False positives percentage**	**Total Events**
Ligated	12.2	15.0	10.0	91.0	7.0	2.0	666
1X	13.7	14.0	14.0	90.0	5.0	5.0	657

Table 5.6*: Comparison of the manual and automated results for PCNbBVLigstepsB.*

The second control sample that was analysed was PCstepsA going from cis-to-trans. This is the plasmid that went through all of the Protocol A modification steps, but had no enzyme or oligonucleotide added. **Figure 5.11** shows the histograms for this sample with and without monovalent streptavidin. Similar to the previous control sample, the dwell time, charges and peak amplitudes are very alike, as well as the scatter plots, suggesting that the protein does not bind.

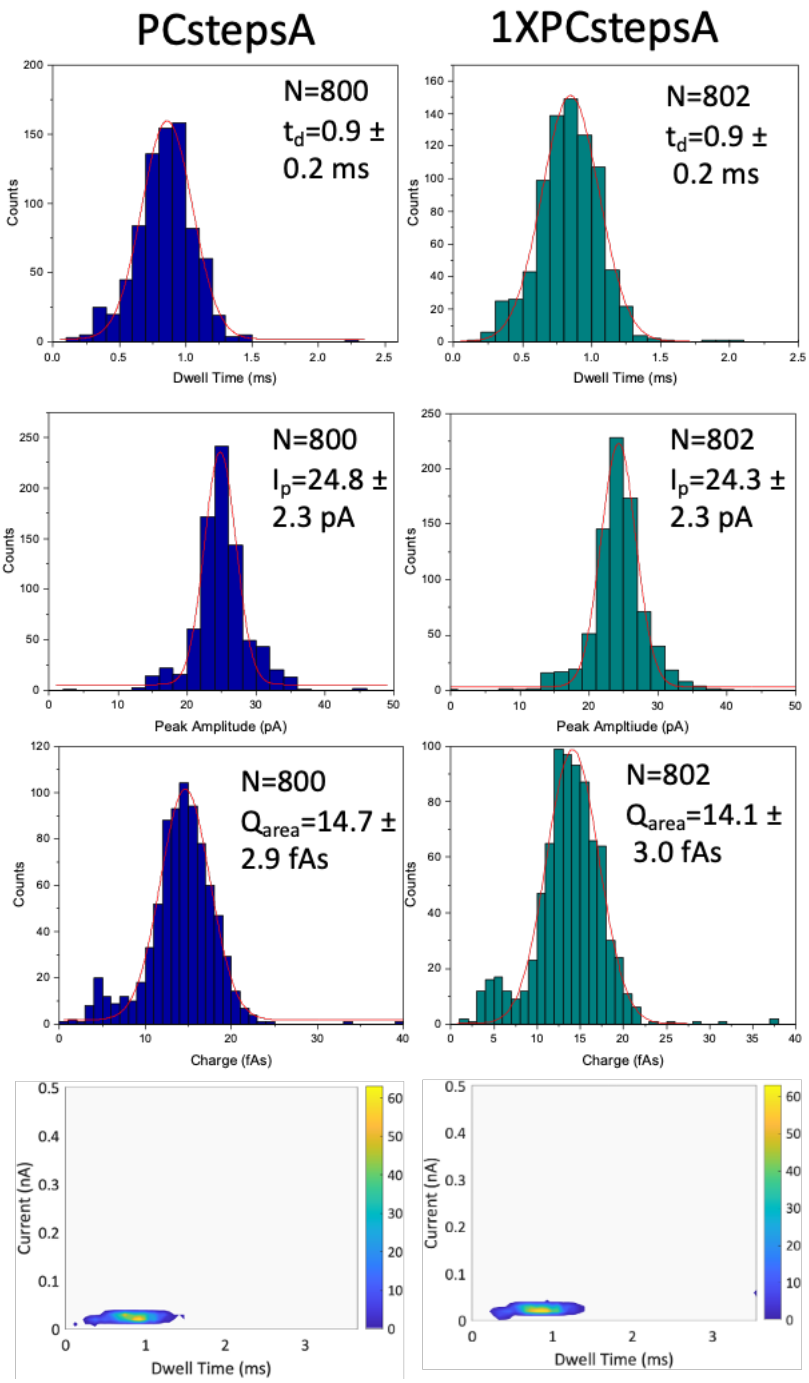


***Figure 5.11**: Histograms of PCstepsA and PCstepsA with monovalent streptavidin in the cis-to-trans direction. A Gaussian fit was used to determine the average values. A two sample T test was performed for dwell time, peak amplitude, and charge, and the mean for peak amplitude was determined to be significantly different, but not for dwell time or charge.*

The automated and manual results were compared, shown in **Table 5.7**. The subpeak parameters used to select for subpeak events were peak height > 0.019 nA and total event width > 0.8 ms. The overall subpeak percentages were nearly the same for both samples, and quite low, and there were high levels of agreement with the manual results.

Additionally, the rates of false negatives and positives were low. Thus, it is reasonable to conclude that the monovalent streptavidin is not able to bind to the plasmid on its own.

	Overall percentage folded/subpeak	Percentage folded/subpeak (first 100) manual	Percentage folded/subpeak (first 100) auto	Agreement with manual percentage	False negatives percentage	False positives percentage	Total Events
Ligated	14.9	16.0	13.0	89.0	7.0	4.0	801
1X	12.9	17.0	11.0	88.0	9.0	3.0	804

***Table 5.7**: Comparison of the manual and automated results for PCstepsA.*

The first of the trans-to-cis controls is PCstepsB. This control sample went through all of Protocol B modification steps without any added enzyme or oligonucleotide. The dwell times, charge, and peak amplitudes for this sample and when it is incubated with monovalent streptavidin are shown in **Figure 5.12**. They are quite similar, and even characteristics in the histograms that might not be expected, such as the smaller secondary peak in the lower range of peak amplitude and charge exist in both samples. The scatter plots also have a similar shape.

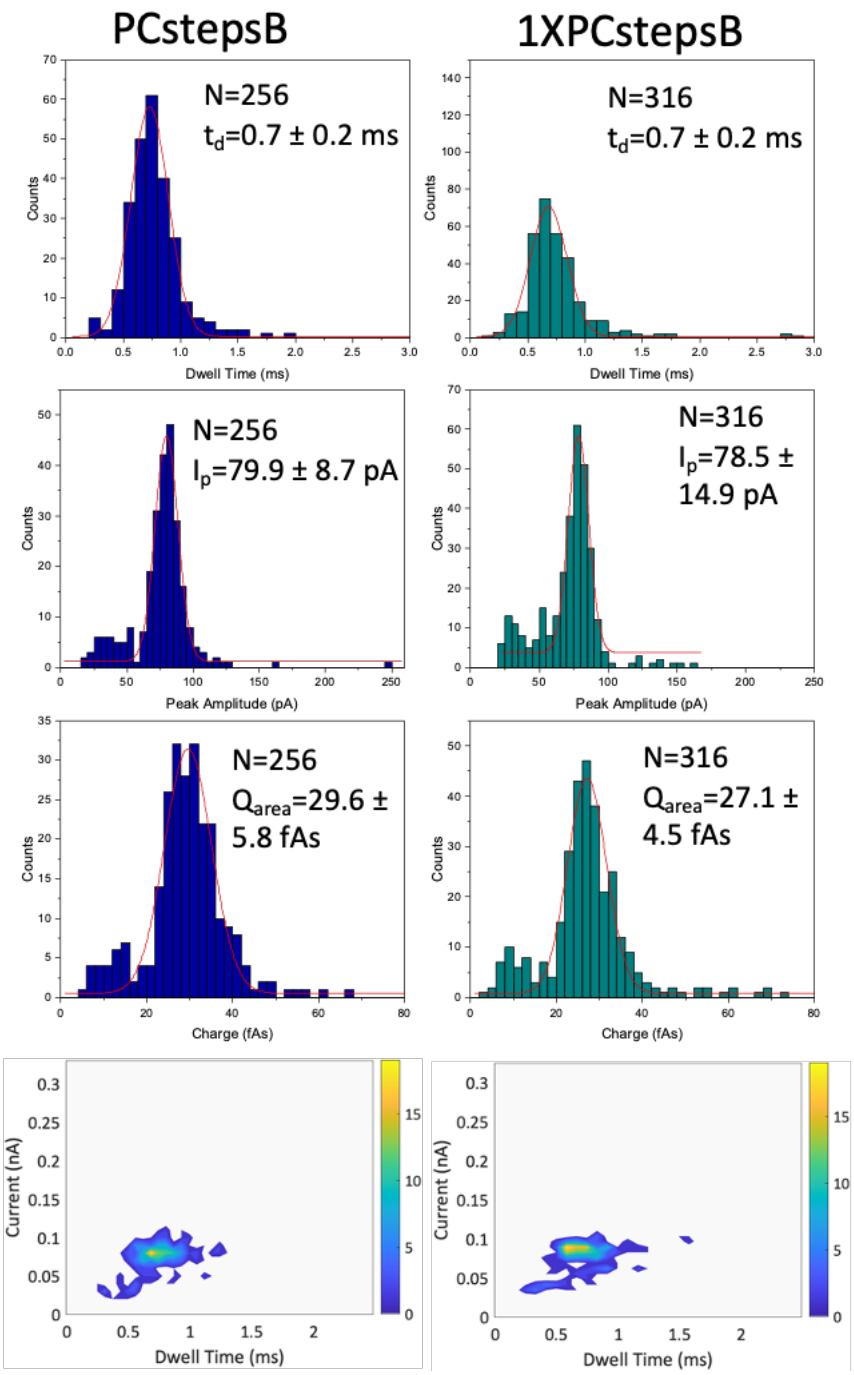


***Figure 5.12**: Histograms of PCstepsB and the control with monovalent streptavidin in the trans-to-cis direction. A Gaussian fit was used to determine the average values. A two sample T test was performed for dwell time, peak amplitude, and charge, and the mean for peak amplitude was determined to be significantly different, but not for dwell time or charge.*

Like the previous controls, the automated and manual results were compared, shown in **Table 5.8**. The subpeak parameters used to select for subpeak events were peak height > 0.07 nA and total event width > 0.6 ms. The overall percentages for PCstepsB and

1XPCstepsnew are again very similar and the agreement with the manual results is very high. These results along with the low percentage of subpeak/folded events suggests that, for this control, the protein is not able to bind.

	Overall percentage folded/subpeak	**Percentage folded/subpeak (first 100) manual**	**Percentage folded/subpeak (first 100) auto**	**Agreement with manual percentage**	**False negatives percentage**	**False positives percentage**	**Total Events**
Ligated	11.3	14.0	7.0	93.0	7.0	0.0	256
1X	9.15	8.0	5.0	95.0	4.0	1.0	317

Table 5.8*: Comparison of manual and automated results for PCstepsB and 1XPCstepsB.*

The last control that was compared to the manual results is PCNbBVLigstepsA trans-to-cis. This sample went through all of the Protocol A modification steps and was incubated with the nicking enzyme and the ligase, but no oligonucleotide. The histograms for this control with and without protein are shown in **Figure 5.13**. The dwell times, charges and peak amplitudes are again similar, and the histograms and scatter plots have a similar shape.

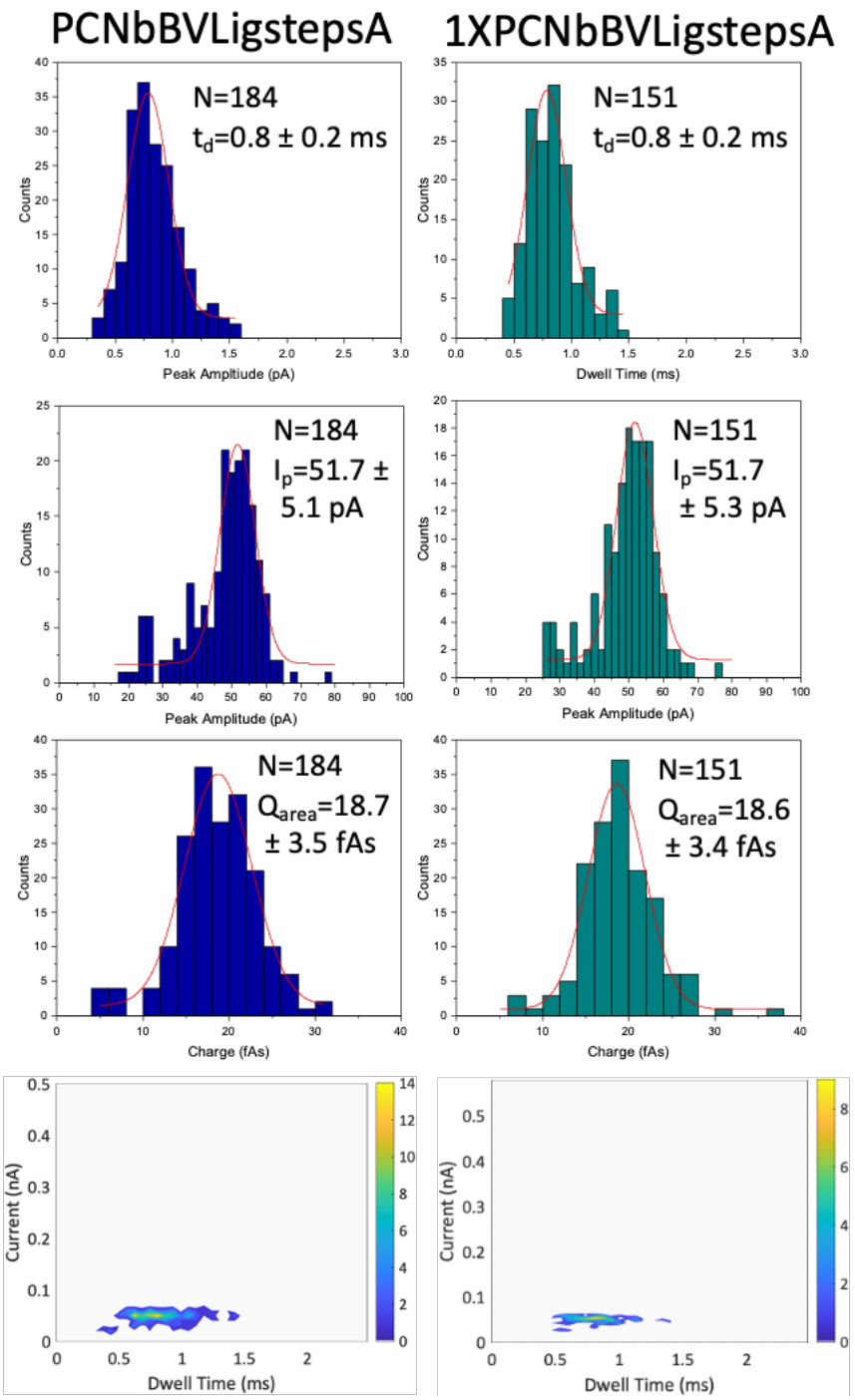


***Figure 5.13**: Histograms for PCNbBVLigstepsA and 1XPCNbBVLigstepsA in the trans-to-cis direction. A Gaussian fit was used to determine the average values. A two sample T test was performed for dwell time, peak amplitude, and charge, and none of them were determined to be significantly different.*

PCNbBVLigstepsA is the last control that had manual and automated results compared. The results of this are shown in **Table 5.9**. The subpeak parameters used to select for subpeak events were peak height > 0.045 nA and total event width > 0.7 ms. The

overall percentages for the control and the controls with protein are both low and similar in value. The agreement with the manual results is good, and the rate of false positives and negatives is low. These results again suggest that for this control the protein is not able to bind.

	Overall percentage folded/subpeak	Percentage folded/subpeak (first 100) manual	Percentage folded/subpeak (first 100) auto	Agreement with manual percentage	False negatives percentage	False positives percentage	Total Events
Ligated	9.2	10.0	7.0	95.0	4.0	1.0	184
1X	4.6	9.0	5.0	94.0	5.0	1.0	151

***Table 5.9**: Comparison of manual and automated results for PCNbBVLigstepsA with and without protein.*

From all of the above results, this automated method of identifying subpeaks is shown to be effective and accurate. The samples with protein are able to have subpeaks identified with both quadrivalent and monovalent streptavidin, and with one or two biotins, in both cis-to-trans and trans-to-cis directions. Additionally, the four controls that were analysed and compared to the manual results, show that the automated method does not select for extra subpeaks in the controls. Thus, it is possible to use this automated method to analyse the binding curve data of this plasmid carrier.

CIS-TO-TRANS SINGLE-MOLECULE BINDING CURVES

5.5 CONTROLS

To ensure that the events being seen and analysed are actually the modified plasmid bound to the protein, various controls were performed. These controls can be split into two groups, controls with protein only and controls with plasmid. The protein only controls were performed to cover the range, from the minimum and maximum concentrations, of any protein used in the binding curve experiments. For the controls with plasmid, the concentration of the plasmid was kept at 100 pM, the same as the concentration of the plasmid in any of the binding curve samples.

For the initial cis-to-trans experiments, the controls with protein only included monovalent streptavidin. Monovalent streptavidin alone at various concentrations, 10 pM, 50 pM, 300 pM, 2 nM, 10 nM, 100 nM, and 500 nM was recorded. Each protein

concentration was recorded in the same conditions as the binding curves and was performed at least three times. The results for a large amount of time, at least 1 or 2 min, are the same for all of the concentrations, no events. This suggests that monovalent streptavidin on its own cannot cause any of the events that are seen in the binding curve samples.

For the initial cis-to-trans binding curves, nine different controls with plasmid were performed. These were needed to confirm that none of the modification steps led to the events with subpeaks, and it was only with all of the modification steps and addition of monovalent streptavidin that the subpeaks were observed. The concentration of the plasmid was kept constant for all of these controls at 100 pM and the concentration of monovalent streptavidin was the same as the controls were incubated with the protein in a 1:1 ratio. The samples with and without protein were always recorded in the same pipette.

The histograms for one run of the two most significant results, PC2ELigfilA and 1XPC2ELigfilA and PC2ELigfilB and 1XPC2ELigfilB are shown below. Other sample controls for cis-to-trans are shown in **Appendix I Figures 4-10**.

These two sets of controls are highly significant. They represent a comparison between the plasmid, which has been incubated with all of the enzymes, but no oligonucleotide, and has undergone through all of the modification protocol steps, with and without protein. **Figure 5.14** compares the histograms for PC2ELigfilA and 1XPC2ELigfilA. The dwell times and peak amplitudes are quite similar and more importantly, the histograms have similar shapes. The similar distribution indicates that the monovalent streptavidin cannot bind to the plasmid and that incorporating the biotin oligonucleotide is critical for monovalent streptavidin bonding. These experiments were repeated in three different pipettes.

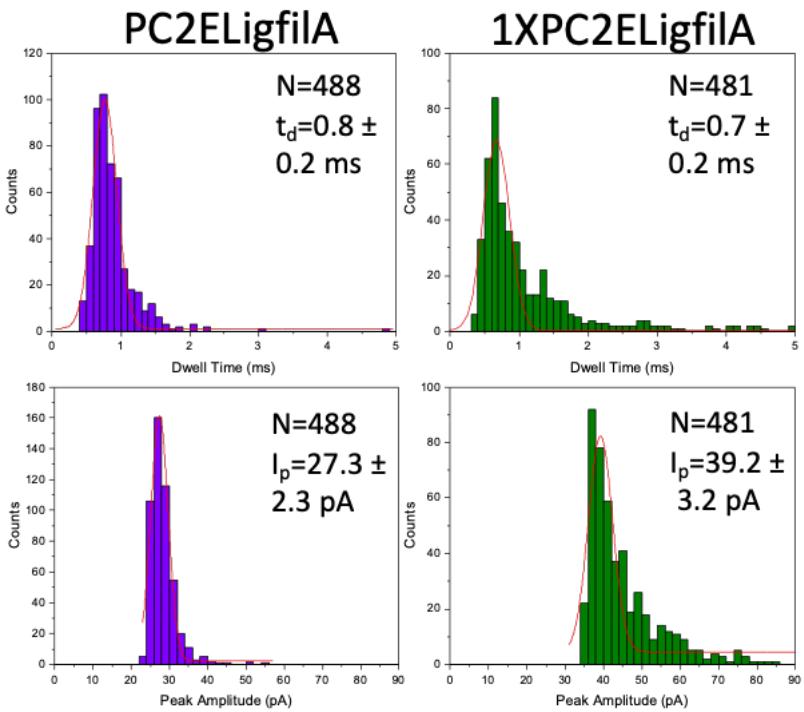


***Figure 5.14**: Histograms for PC2ELigfilA and 1XPC2ELigfilA. The dwell times and peak amplitudes are very similar. These are critical control results as they show that it is not possible for the monovalent streptavidin to bind without the presence of the biotin oligonucleotide. A Gaussian fit was used to determine the average values.*

This control is the plasmid incubated with all the enzymes and going through all of Protocol B modification steps, but without the oligonucleotide, PC2ELigfilB. Like the previously mentioned control, there is not much shift between dwell times and peak amplitudes for this control. Therefore, the monovalent streptavidin is likely not able to bind. These experiments were repeated in three different pipettes and the results from one are shown in **Figure 5.15**.

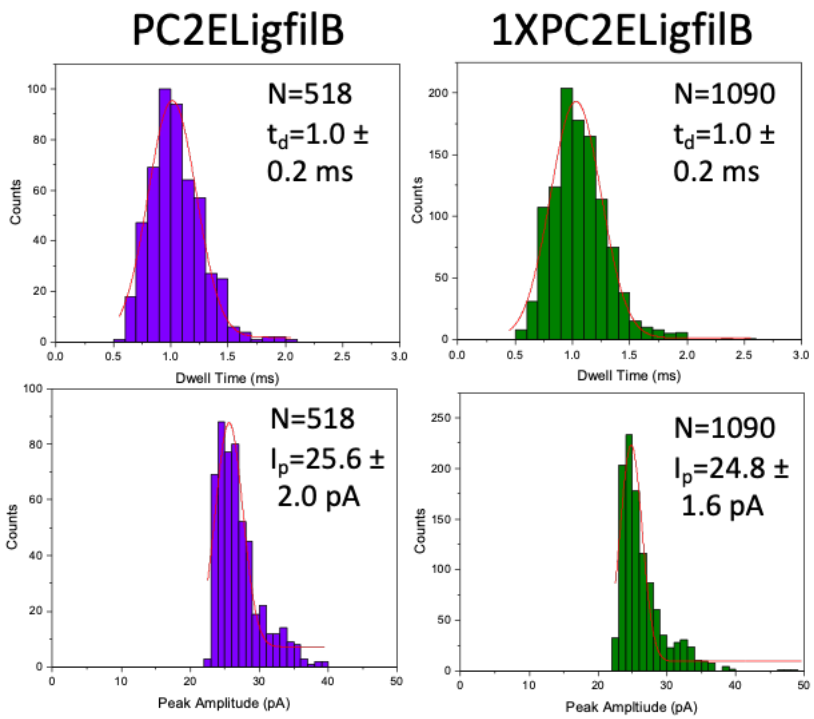


***Figure 5.15**: Comparison of PC2ELigfilB and 1XPC2ELigfilB results. The lack of dwell time shift and peak amplitude change suggest that the monovalent streptavidin was not able to bind to the control sample. A Gaussian fit was used to determine the average values.*

The controls presented in this section show that none of the plasmid modification steps, except for the addition of biotin oligonucleotides, results in large dwell time shifts or peak amplitude changes. One example for each control, for both the Protocol A and B is shown, but for each result were repeated in three different pipettes. Thus, it is reasonable to conclude that the only samples that result in high levels of events with subpeaks are the ones that have a biotin oligonucleotide and have been incubated with monovalent streptavidin.

5.6 COMPARISON BETWEEN PROTOCOLS A AND B FOR ONE BIOTIN

The first plasmid sample that was investigated with protein binding was PC2ELigBOfilA. This is the plasmid that undergoes the full Protocol A modification process with only one biotin on the oligonucleotide. The plasmid was incubated with monovalent streptavidin at various different concentrations, 50 pM, 100 pM, 1 nM, 5 nM,10 nM, 25 nM, and 100 nM, while its concentration was kept constant at 100 pM. All of these concentrations were recorded in the same pipette. There is a significant difference between

the samples without protein and with. **Figure 5.16** shows sample current traces and events for one pipette, where the PC2ELigBOfilA sample is compared to 50X. There is a difference in the general peak amplitude for the two samples, PC2ELigBOfilA is navy and 50X is turquoise, and the sample events also have differences. Event a in **Figure 5.16** has a dwell time of 1.0 ms and peak amplitude of 20.8 pA, event b a dwell time of 0.9 ms and peak amplitude of 15.9 pA, event c's dwell time is 2.1 ms and peak amplitude is 29.9 pA, and event d's dwell time is 1.6 ms and peak amplitude is 28.0 pA. Events a and b are from the PC2ELigBOfilA recording, and c and d are from 50X. Additionally the event shapes are different. The PC2ELigBOfilA ones look more triangular while the 50X ones often have an additional level, subpeak, to them. This extra level or change to the triangular shape suggests that the protein is binding to the plasmid.

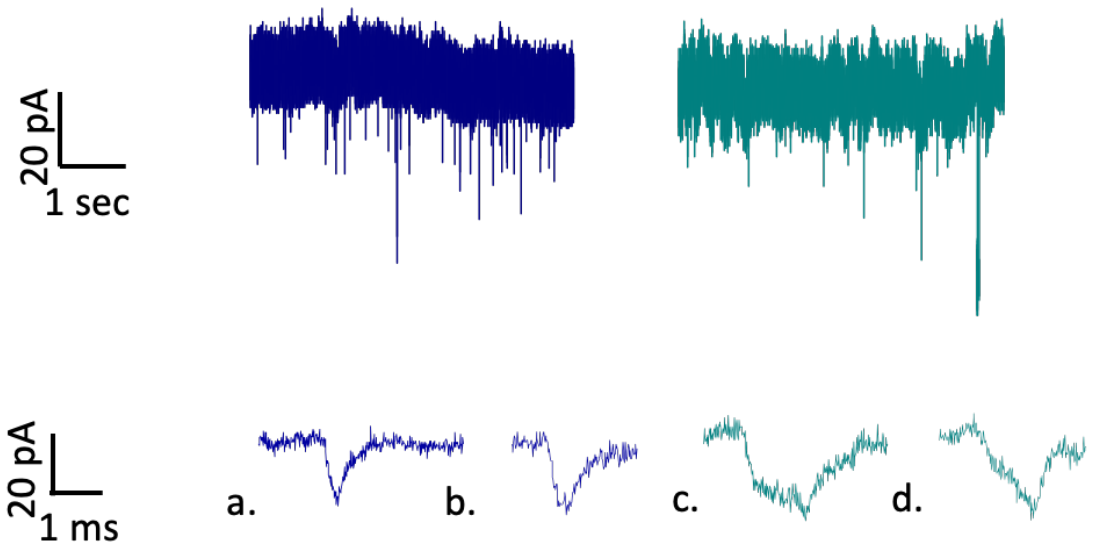


***Figure 5.16**: Sample current traces and events for PC2ELigBOfilA and 50X. PC2ELigBOfilA is shown in navy and 50X in turquoise.*

Similar binding curves for the plasmid with one biotin that uses Protocol B were also performed. The plasmid sample was again incubated with seven different monovalent streptavidin concentrations, and they were all recorded in the same pipette. **Figure 5.17** shows sample current traces and events for PC2ELigBOphfilB and 50X. Again there are

differences between the two samples, and this is particularly evident in the variation of peak amplitudes of the 50X sample. The sample events also have quite different shapes, as well as dwell times and peak amplitudes. The dwell times and peak amplitudes respectively for event a is 0.8 ms and 15.9 pA, event b is 0.9 ms and 22.0 pA, event c is 1.9 ms and 32.2 pA, and event d is 2.1 ms and 28.0 pA. Events a and b are from the PC2ELigBOphfilB recording and c and d are from 50X. The increase in dwell time and peak amplitude for the events with protein is what is expected.

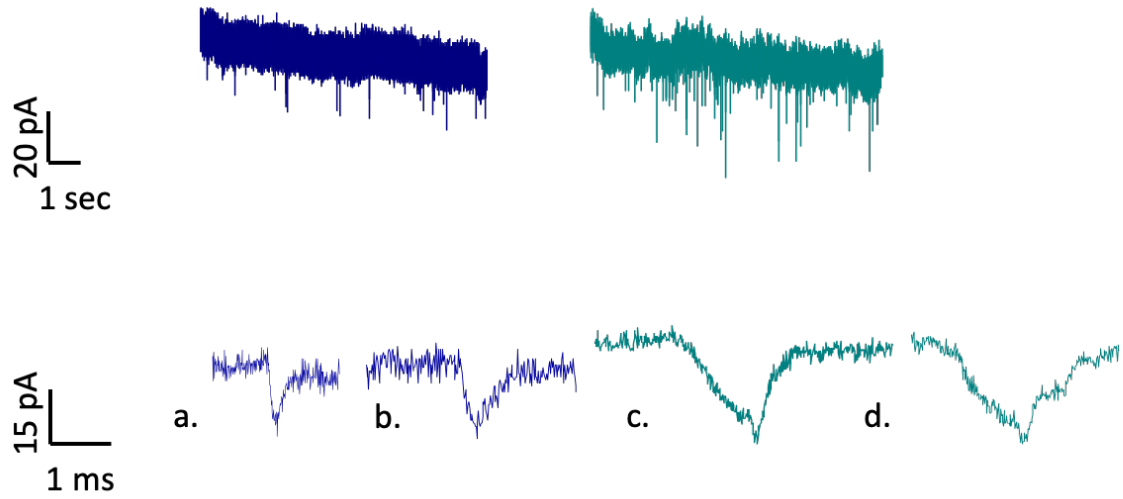


***Figure 5.17**: Sample current traces and events for PC2ELigBOphfilB and 50X. PC2ELigBOphfilB is in the navy and 50X in the turquoise.*

The results from the automated data analysis for both of these samples were then combined into binding curves. The binding curves measured how the percentage of subpeaks changed, taken from the automated results, as the concentration of monovalent streptavidin increased. Binding curves for PC2ELigBOfilA and PC2ELigBOphfilB are shown in **Figure 5.18**. For PC2ELigBOfilA, there is a very clear trend of subpeak percentage increasing as monovalent streptavidin concentration increases. For PC2ELigBOphfilB, the subpeak percentages does increase, relative to the folded events, but not with additional protein. Perhaps this has to do with the oligonucleotide being more successfully taken up when using Protocol B so that at the higher protein concentrations the biotins are already bound to monovalent streptavidin and thus saturated. Each point for the binding curve for the

Protocol A has data from at least five different pipettes and the percentage of folded events is 17.5 ± 7.9%. For Protocol B, each point has data from at least nine different pipettes and the percentage of folded events is 22.6 ± 8.5%. The diameter of the pipettes for the Protocol A varied from about 11-27 nm and for Protocol B 19-30 nm. This large variation is due to the large amounts of time between recordings, partially due to the covid-19 pandemic, and subsequent necessary tailoring of the pipette pulling protocol.

The error bars on the graphs are the standard deviation of the subpeak percentage at each concentration. The lower concentrations for both binding curves do not vary much between the samples without protein, and this result makes sense as the protein concentration is very low and one-hundred percent incorporation of the biotin oligonucleotide into the plasmid is unrealistic. For Protocol A, significant difference between the 1 nM, 5 nM, 25 nM, and 100 nM protein samples and the folded events was determined using a two sample T test. For Protocol B, significant difference between the 1 nM, 5 nM, and 100 nM protein samples and the folded events was determined using a two sample T test. Other factors that affect these results include variation between samples, particularly as the samples were very low volumes the monovalent streptavidin was difficult to incubate with the plasmid. Difficulties in protocol steps like this can explain why sometimes the protein did not seem to bind very efficiently to the plasmid. The pipette could also be susceptible to blockage and a partial blockage could greatly vary the results between samples or indeed the results within one sample. The variation in pipette diameters, and probably also shape, would again have had a significant effect on the results. Furthermore, the addition of a phosphorylated 5’ end in the biotin oligonucleotide could have a significant effect, as a higher level of uptake would be expected for this oligonucleotide; however this overall effect could be countered by the change in excess ratio for the oligonucleotide in Protocol B; that is the decrease from 100:1 to 3:1. The higher concentrations and their corresponding higher percentages show that the biotin has successfully bound to the monovalent streptavidin and that the events for these are sufficiently different to the sample without protein.

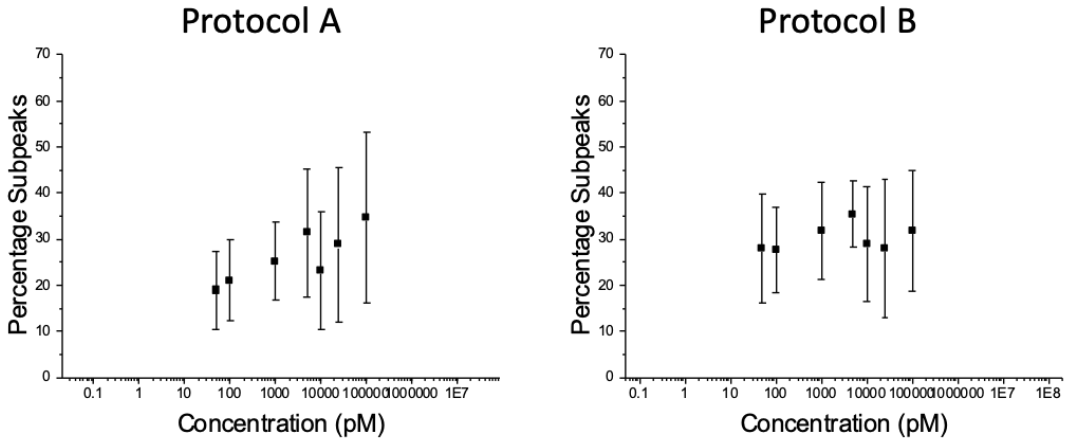


***Figure 5.18**: Binding curves of PC2ELigBOfilA and PC2ELigBOphfilB where the protein concentrations are 50 pM, 100 pM, 1 nM, 5 nM,10 nM, 25 nM, and 100 nM. The folded event percentages are 17.5 ± 7.9% and 22.6 ± 8.5% respectively*

5.7 COMPARISON BETWEEN PROTOCOL A AND B FOR TWO BIOTINS

The plasmid was also modified with an oligonucleotide that had two biotins rather than just one. As the oligonucleotide is 66 bases in length, two monovalent streptavidins should be able to bind without interference from steric effects. Again, the plasmid was incubated with monovalent streptavidin at various different concentrations, 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM, and 100 nM, while its concentration was kept constant at 100 pM. These concentrations were kept the same, although the number of biotins available for bonding had increased by a factor of two. All of these concentrations were recorded in the same pipette. Like with the plasmid with one biotin, there is a significant difference between the samples without protein and with. **Figure 5.19** shows sample current traces and events for PC2ELigBO2filA with and without protein. In the current traces, there is more variation in event peak amplitudes in the sample with protein, suggesting that the protein is binding. Also, the sample events have dwell times and peak amplitudes respectively with a having a dwell time 0.8 ms and peak amplitude of 23.2 pA, b having a dwell time of 0.9 ms and a peak amplitude of 15.0 pA, c having a dwell time of 1.3 ms and a peak amplitude of 23.3 pA, and d having a dwell time of 1.0 ms and a peak amplitude of 28.0 pA. The sample events a and b are PC2ELigBO2filA and c and d are 50X. In addition to the clear differences in dwell time and peak amplitude in the sample events, there are extra subpeaks in the events with protein. This additional peak suggests that the protein has bound successfully to the plasmid.

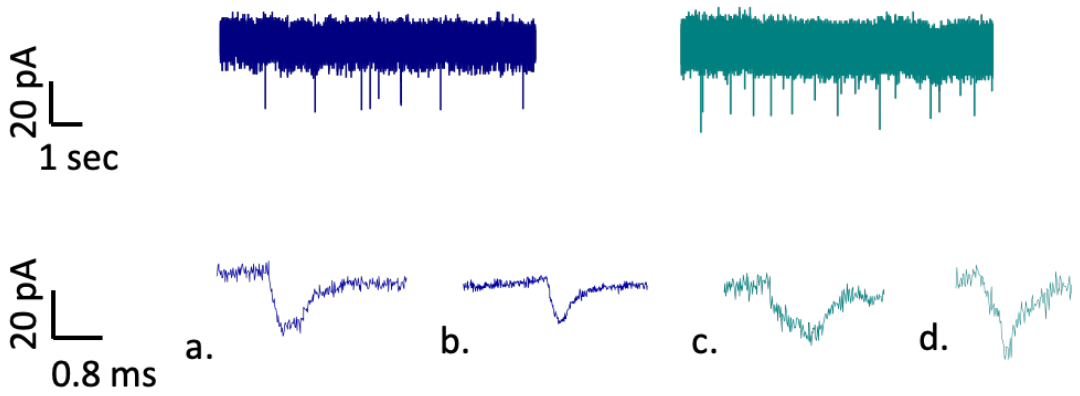


***Figure 5.19**: Sample current traces and events for PC2ELigBO2filA with, turquoise, and without, navy, protein.*

Results were also recorded for the plasmid with two biotins modified using Protocol B. This sample was similarly incubated with monovalent streptavidin at concentrations 50 pM, 100 pM, 1 nM, 5 nM,10 nM, 25 nM, and 100 nM, while the analyte concentration was kept constant at 100 pM, and again all of these were recorded in the same pipette. Like with Protocol A, there was some variation in dwell time and peak amplitude that can be seen in both current trace and sample events, shown in **Figure 5.20**. The average peak amplitude on the current trace is higher, and there is more variation in peak amplitude for the sample with protein than the sample without. The sample events have dwell times and peak amplitudes respectively of 0.9 ms and 17.1 pA, b 0.8 ms and 17.7 pA, and c 0.8 ms and 25.6 pA, and d 1.5 ms and 33.0 pA. The sample events a and b are for PC2ELigBO2filB, and c and d for the sample with protein. While the peak amplitude is much larger for these sample events, the dwell time is not that different. However, the events shape is also an extremely important characteristic, and both sample events for the 50X have clear subpeaks. It is the presence of this subpeak, as well as the change in peak amplitude that suggests that the monovalent streptavidin is able to successfully bind.

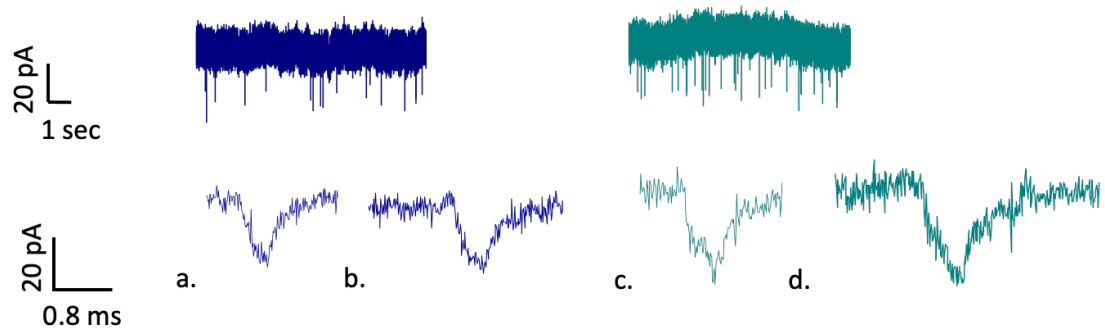


***Figure 5.20**: Sample current traces and events for PC2ELigBO2filB with, turquoise, and without, navy, protein.*

The results for both of these samples were then combined into binding curves. The binding curves measured how the percentage of subpeaks, taken from the automated results as the concentration of monovalent streptavidin increased. Binding curves for PC2ELigBO2filA and PC2ELigBO2filB are shown in **Figure 5.21**. Each point on these graphs has data from at least five different pipettes and the pipette size varied from 20-35 nm for PC2ELigBO2filA and 18-37 nm for PC2ELigBO2filB. This size variation was again probably due to the time in between recordings. There was an increase in subpeak percentage for both the binding curves for Protocols A and B compared to the sample without any protein. The folded events percentages are 20.6 ± 8.2% for PC2ELigBO2filA and 22.4 ± 8.9% for PC2ELigBO2filB. Furthermore, for Protocol A, significant difference between the 1 nM, 5 nM, 10 nM, and 25 nM protein samples and the folded events was determined using a two sample T test. For Protocol B, significant difference between the 1 nM, 5 nM, and 10 nM protein samples and the folded events was determined using a two sample T test. However, the subpeak percentage does not rise continually and for some of the higher protein there is a lack of significant difference between these samples and the samples without protein. This could stem from several things, including the high amounts of monovalent streptavidin partially blocking the pore, the larger amount of plasmid with two biotins taking longer to move through the pore and/or blocking it, or especially as the concentrations were recorded in order of increasing protein concentration, the pore deteriorating or becoming partially blocked, as more sample moved through it. There also could be issues with mixing during the protein incubation, although given that it is several of the higher protein concentration samples that have lower than expected subpeak percentages, this is unlikely.

However, particularly when looking at the percentage values with the error, the error was again calculated from the standard deviation of the results from each pipette, there is a definite increase in the percentage of subpeaks compared to the percentage of folded events.

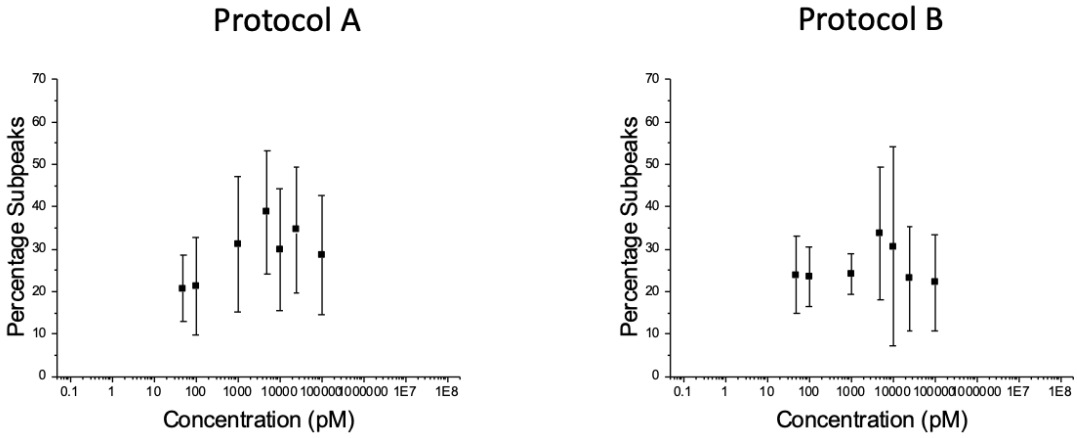


***Figure 5.21**: Binding curves of PC2ELigBO2filA and PC2ELigBO2filB where the protein concentrations are 50 pM, 100 pM, 1 nM, 5 nM,10 nM, 25 nM, and 100 nM. The folded event percentages are 20.6 ± 8.2% and 22.4 ± 8.9% respectively.*

There also does not seem to be a significant difference in event shape between the samples with one biotin and the samples with two biotins. This could be because the monovalent streptavidin are not far enough apart in the plasmid in order to be discriminated using an nanopore readout; that is, they do not have two distinct subpeaks. There do seem to be some differences in the dwell times and peak amplitudes used to separate the samples with biotin and monovalent streptavidin from the samples with only biotin, whether one or two. Typically the dwell times and peak amplitudes used for the separation in the automated process were larger for the samples with two biotins, as would be expected given the addition of another monovalent streptavidin. However, given the variation between pipettes more experiments would need to be performed, perhaps with recording both samples within the same pipette, to confirm if this is the case. Furthermore, future work within the same pipette could include a detailed investigation into this comparison that compares the average one biotin sample with protein to the average two biotin sample with protein.

5.8 POTENTIAL FOR MULTIPLEXED SENSING

The percentage of subpeaks did not increase for the two biotins sample compared to the one biotin sample as might have been expected, given that it was now theoretically twice as likely that monovalent streptavidin would be able to bind. However, it is not that unexpected when considering some other factors. The factor that probably has the most influence is the phosphorylated 5' end on the oligonucleotide with one biotin only. As it is not possible to have this phosphorylation for the two biotin oligonucleotide, there is probably lower uptake for this oligonucleotide than the single biotin oligonucleotide. In addition, the excess ratio was also changed in Protocol B for both oligonucleotides despite the two biotin oligonucleotide not having phosphorylation. However, the positive results for the two biotin oligonucleotide suggest that multiplex sensing is possible. This could be accomplished in several different ways. The first is having different aptamers attached to each end of the oligonucleotide, and as long as the proteins are not much bigger than monovalent streptavidin, steric effects should not be a problem. Additionally, it would be possible to incorporate multiple oligonucleotides into the plasmid and spread the proteins out a bit more, this could even lead to separate subpeaks for the proteins assuming they were far enough apart. Lastly, which is the direction explored in this chapter, quadrivalent streptavidin could be used to bind biotinylated proteins, using a sandwich-like assay.

CIS-TO-TRANS MULTIPLEXED SENSING

5.9 EXPERIMENTAL SETUP

The multiplexed sensing explored in this work revolves around the biotin streptavidin bond. As quadrivalent streptavidin has four binding sites, it is possible for a quadrivalent streptavidin to be bound to a plasmid carrier and then for other biotinylated proteins to be bound to another of the quadrivalent streptavidin binding sites. Thus, forming a kind of sandwich assay. This sandwich-like assay demonstrates that the plasmid can perform a different kind of multiplexed sensing than just incorporating multiple aptamers, although it will perhaps be less specific. These experiments used biotinylated phosphatase as the second protein. This protein was chosen as it was easy to acquire in the biotinylated form and also is large enough, at around 94 KDa, that theoretically it should be possible to see a change in subpeak size.[4] The quadrivalent streptavidin was incubated with the plasmid carrier as normal for one hour, and then this complex was incubated with the

biotinylated phosphatase for another hour in the same phosphate buffer. The concentration of the biotinylated phosphatase was one third that of the quadrivalent streptavidin in order to avoid as many multiple bindings of biotinylated phosphatases to a quadrivalent streptavidin as possible.

5.10 CONTROLS

As these experiments involved some proteins other than the ones described in the section before several new controls were performed. All of these experiments were only performed with PC2ELigBOphfilB so only this carrier and the version without an oligonucleotide were used in the controls. Each sample with and without protein was performed in the same pipette and the reaction conditions were identical for all of them: that is, 100 mM KCl in TE buffer pH 8, -300 mV, 100 pM plasmid concentration with a 1:1 ratio for the protein, 100 KHz sampling rate, and a 30 KHz filter. One of the most important of these controls is PC2ELigfilB with quadrivalent streptavidin. The histograms for this control are shown in **Figure 5.22**. Both the dwell times and peak amplitudes are similar with and without protein and the shape of the histograms looks very alike, suggesting that the quadrivalent streptavidin is not able to bind. Other controls included the plasmid sample being incubated with phosphatase and biophosphatase and these are shown in **Appendix I Figures 11** and **12**.

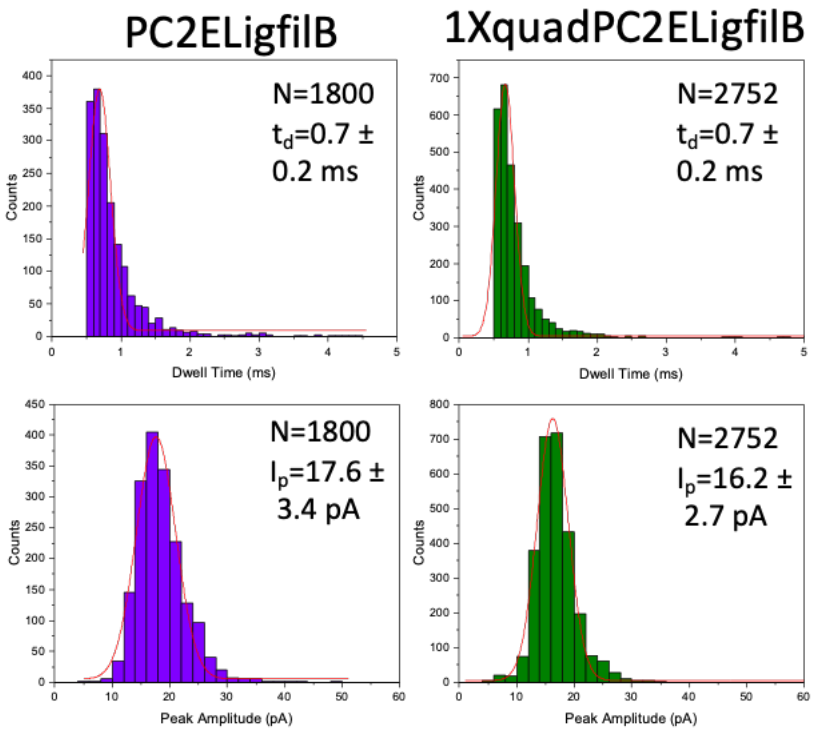


***Figure 5.22**: Histograms for the control PC2ELigfilB with and without quadrivalent streptavidin. The similar shape of the histograms, as well as close agreement between the dwell times and peak amplitudes suggests that the quadrivalent streptavidin is not able to bind. A Gaussian fit was used to determine the average values.*

The last control is PC2ELigfilB with quadrivalent streptavidin and biotinylated phosphatase. As the plasmid carrier does not have the biotin, it should not be possible for either added protein to bind. The very similar histogram shapes, particularly in the peak amplitude, in addition to the close dwell time and peak amplitudes suggest that it is indeed the case that neither is able to bind. **Figure 5.23** shows the histograms for this control.

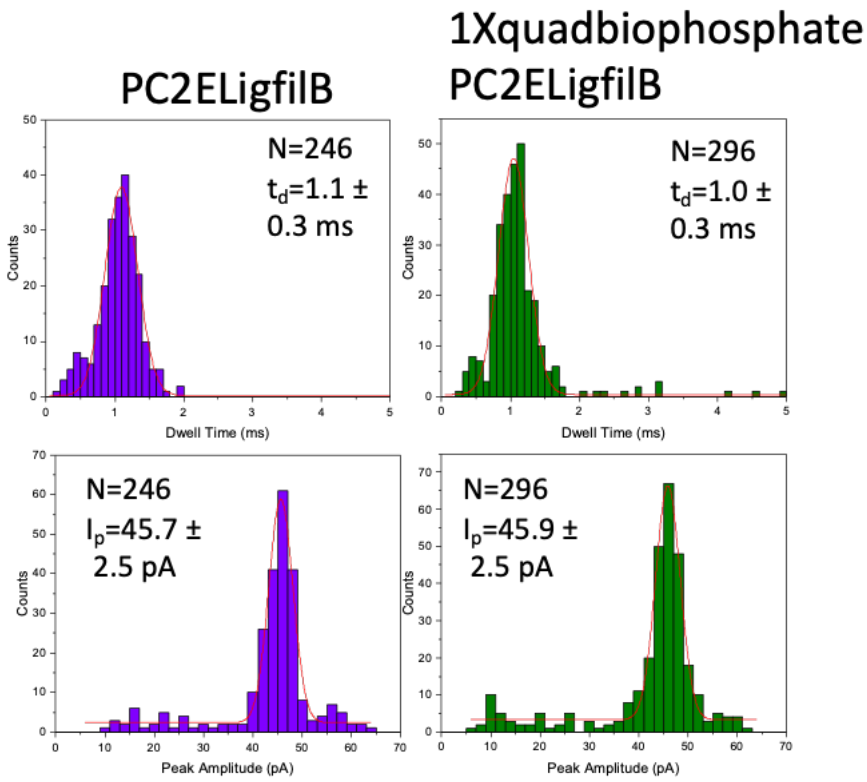


***Figure 5.23**: Histograms for PC2ELigfilB with and without quadrivalent streptavidin and biotinylated phosphatase. The similar histogram shapes, and dwell times and peak amplitudes suggest that it is not possible for either of these proteins to bind without the presence of biotin. A Gaussian fit was used to determine the average values.*

Protein alone controls were also performed for quadrivalent streptavidin, phosphatase, and biophosphatase to confirm that at the concentrations used in these experiments there were no events from the protein alone. The current traces were recorded at 10 pM and 500 nM. Each control was repeated at least three times and over long periods of recording time, 1-2 min, and there were no events.

In summary, all of these new controls support the conclusion that neither phosphatase is able to bind to the carrier. Additionally, it is not possible for the quadrivalent streptavidin to bind to the carrier without biotin even in the presence of biotinylated phosphatase.

5.11 COMPARISON OF QUADRIVALENT TO MONOVALENT STREPTAVIDIN

The binding of quadrivalent streptavidin to PC2ELigBOphfilB was investigated. The plasmid was incubated with quadrivalent streptavidin at various different concentrations, 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM, and 100 nM, while its concentration was kept

constant at 100 pM. All of these concentrations were recorded in the same pipette. However, as opposed to the samples with monovalent streptavidin, there is the possibility for multiple carriers to bind to each streptavidin. This added layer of complexity could cause some issues in the analysis of the results as, assuming steric effects do not stop the plasmids from binding, more than one can bind to each streptavidin. **Figure 5.24** shows sample current traces and events for one pipette, where the PC2ELigBOphfilB sample is compared to 50Xquad. There is a difference in the general peak amplitude for the two samples, and the sample events also have differences. Event a in **Figure 5.24** has a dwell time of 0.7 ms and peak amplitude of 15.3 pA, event b a dwell time of 0.9 ms and peak amplitude of 18.3 pA, event c's dwell time is 6.7 ms and peak amplitude is 33.0 pA, and event d's dwell time is 3.3 ms and peak amplitude is 44.6 pA. Events a and b are from the PC2ELigBOphfilB recording and c and d are from 50Xquad. In these sample events, there is a very clear difference between the events with protein and those without. Both the dwell time and peak amplitude are larger and the event shapes are different. The 50Xquad ones have a clear subpeak, and this suggests that the protein is binding to the plasmid.

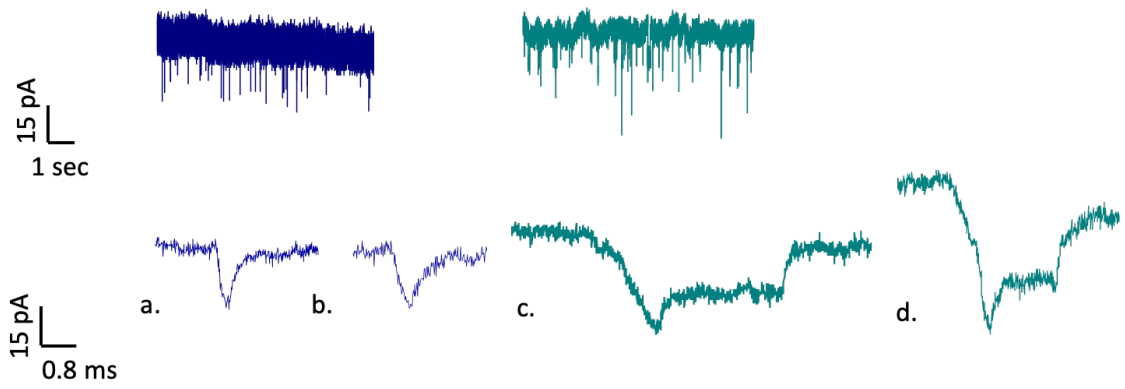


***Figure 5.24**: Sample events and current traces for PC2ELigBOphfilB and 50Xquad. The PC2ELigBOphfilB is shown in navy and the 50Xquad in turquoise.*

It was then important to compare the binding curves for the monovalent and quadrivalent streptavidin. The binding curves measure how the percentage of subpeaks, taken from the automated results, vary as the concentration of both the monovalent streptavidin and the quadrivalent streptavidin increased. Binding curves for this comparison are shown in **Figure 5.25**. The pipette sizes for these curves varied from 19-30 nm for the

monovalent streptavidin results and 17-37 nm for the quadrivalent streptavidin results. For the quadrivalent streptavidin, each point had data from at least twelve different pipettes. Additionally, the percentages for PC2ELigBOphfilB in the monovalent streptavidin and the quadrivalent streptavidin respectively are 22.6 ± 8.5% and 19.2 ± 8.4%, so quite similar. For monovalent streptavidin, significant difference between the 1 nM, 5 nM, and 100 nM protein samples and the folded events was determined using a two sample T test. For quadrivalent streptavidin, significant difference between the 50 pM, 100 pM, 1 nM, and 5 nM protein samples and the folded events was determined using a two sample T test.

Interestingly, for samples at low concentrations, the quadrivalent streptavidin that the binding is significantly higher than the folding. This could perhaps stem from the ability of the quadrivalent streptavidin to bind multiple times and that it's K_d is slightly lower than the monovalent streptavidin, so it may be able to bind better than the monovalent streptavidin and thus increase the percentage of subpeaks slightly. However, at the higher concentrations, the percentage of subpeaks for the quadrivalent streptavidin is significantly lower than for the monovalent streptavidin, particularly as none of the samples above 5 nM for the quadrivalent streptavidin are significantly different than the sample without protein. Perhaps, as the number of binding sites is so in excess compared to the plasmid, significant numbers of the quadrivalent streptavidin are able to bind multiple plasmids, and these much larger complexes either displace many carriers with just one quadrivalent streptavidin or block the pore. This hypothesis is supported by the fact that there are many longer and larger events in the quadrivalent streptavidin recordings than in the monostreptavidin ones. The pipettes also got blocked more during the quadrivalent streptavidin recordings. Given that some of the monovalent streptavidin pipettes were recorded around the same time as the quadrivalent streptavidin pipettes and that the average size is similar, it is unlikely to entirely be an effect of pore shape or length. Indeed at 1000Xquad, there are barely more subpeaks than folded events which suggests that there is some other effect causing such a drastic change.

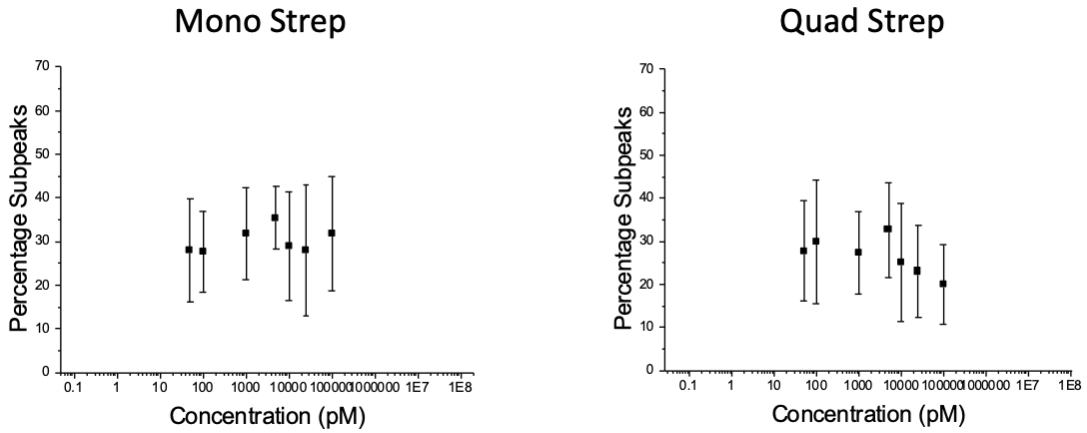


***Figure 5.25**: A comparison of the binding curves for monovalent and quadrivalent streptavidin with PC2ELigBOphfilB. The protein concentrations are 50 pM, 100 pM, 1 nM, 5 nM,10 nM, 25 nM, and 100 nM. The folded event percentages are 22.6 ± 8.5% and 19.2 ± 8.4% respectively.*

5.12 QUADRIVALENT STREPTAVIDING + BIOTINYLATED PHOSPHATASE BINDING CURVES

To explore the sandwich-like assay binding, biotinylated phosphatase binding was performed. The plasmid was incubated with quadrivalent streptavidin at various different concentrations, 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM, and 100 nM, while its concentration was kept constant at 100 pM. It was then incubated again with biotinylated phosphatase at a concentration one third of that of the quadrivalent streptavidin. All of these concentrations were recorded in the same pipette. Sample current traces and events for PC2ELigBOphfilB and 50Xquadrivalent streptavidin biotinylated phosphatase are shown in **Figure 5.26**. There is a difference in the general peak amplitude and again the sample with protein is shown in turquoise and without in navy. The sample events have much greater difference than the general recording. Event a in **Figure 5.26** has a dwell time of 0.6 ms and peak amplitude of 15.2 pA, event b a dwell time of 0.7 ms and peak amplitude of 15.0 pA, event c's dwell time is 1.1 ms and peak amplitude is 22.9 pA, and event d's dwell time is 0.9 ms and peak amplitude is 17.1 pA. Events a and b are from the PC2ELigBOphfilB recording and c and d are from 50Xquadrivalent streptavidin biotinylated phosphatase. The peak differences, particularly the shape with the subpeak present for the events with protein, but also the increase in dwell time suggest that the protein is binding.

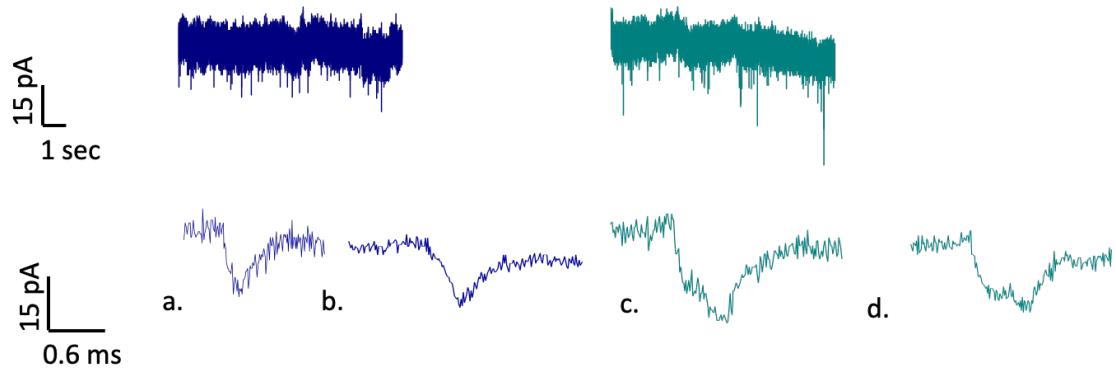


***Figure 5.26**: Sample events and current traces for PC2ELigBOphfilB and 50X quadrivalent streptavidin biotinylated phosphatase. The PC2ELigBOphfilB is shown in navy and the 50Xquadrivalent streptavidin biotinylated phosphatase in turquoise.*

The binding curves of quadrivalent streptavidin and quadrivalent streptavidin with biotinylated phosphatase were compared. They have the same protein concentrations for the quadrivalent streptavidin, but the biotinylated phosphatase is one third of the concentration of the quadrivalent streptavidin at any given point. Binding curves for this comparison are shown in **Figure 5.27**. The pipette sizes for these curves varied from 17-37 nm for the quadrivalent streptavidin results and 18-33 nm for the quadrivalent streptavidin biotinylated phosphatase so it can be assumed that the pipette shape and size were similar. Additionally, the percentages for PC2ELigBOphfilB in the quadrivalent streptavidin and the quadrivalent streptavidin biotinylated phosphatase respectively are 19.2 ± 8.4% and 15.3 ± 7.6%. For the quadrivalent streptavidin biotinylated phosphatase binding curve, each point had data from at least eight different pipettes. For Quad Strep, significant difference between the 50 pM, 100 pM, 1 nM, and 5 nM protein samples and the folded events was determined using a two sample T test. For Quad + Biophos, significant difference between the 5 nM, 10 nM, 25 nM, and 100 nM protein samples and the folded events was determined using a two sample T test.

The data again for these binding curves are quite different. The quadrivalent streptavidin biotinylated phosphatase binding curve increases as protein concentration increases, while the quadrivalent streptavidin one increases at the lower protein

concentrations, but not at the higher ones. Perhaps with the presence of biophosphatase, even if it is not possible to control for the addition of one, two, or three biophosphatases, stops or greatly reduces the binding of multiple plasmid carriers to one quadrivalent streptavidin. As biophosphate is considerably smaller than a 10 Kbp plasmid, it is not able to block the pore in the same way and thus can continue having a large number of subpeaks at large concentrations of protein. This is reflected in that all the protein concentrations higher than 1 nM have significant difference compared to the samples without protein. However, a downside of this sandwich-like assay is that it is not possible to tell if the biophosphate has bound to the quadrivalent streptavidin. While subpeaks exist for many events, there is no clear difference between them and the subpeaks that exist for quadrivalent streptavidin, although some biophosphatase must be binding to get such a different effect in the overall percentages of subpeaks.

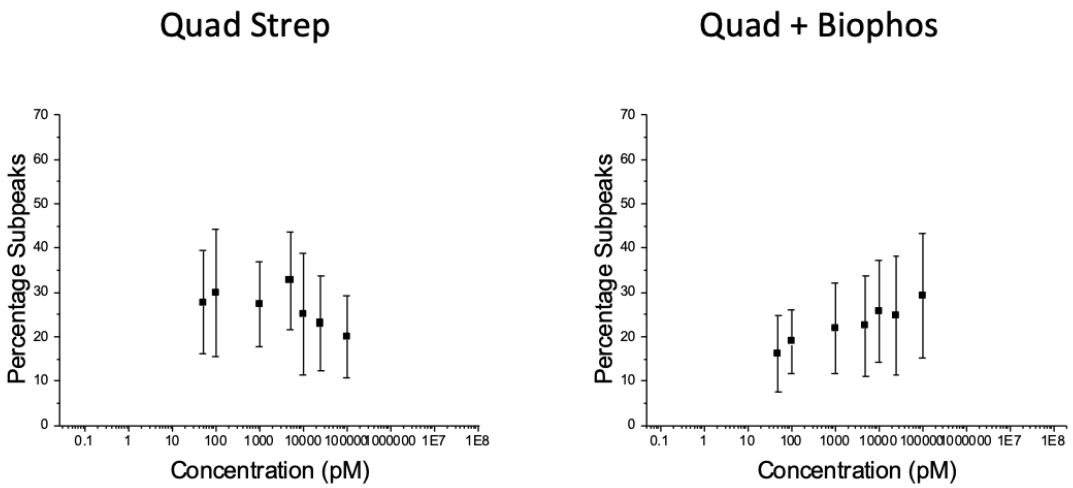


Figure 5.27*: A comparison of the binding curves for quadrivalent streptavidin and quadrivalent streptavidin biotinylated phosphatase with PC2ELigBOphfilB. The protein concentrations are 50 pM, 100 pM, 1 nM, 5 nM,10 nM, 25 nM, and 100 nM for quadrivalent streptavidin, and one third of those values of the biophosphatase. The folded event percentages are 19.2 ± 8.4% and 15. ± 7.6% respectively.*

Overall for multiplexed sensing, the biophosphatase definitely seems to bind, as shown by the different results for the quadrivalent streptavidin binding curve and the quadrivalent streptavidin biotinylated phosphatase binding curve, but it is not as clear cut as expected. That is, it is not so easy to tell from individual events that the biophosphatase is binding to the quadrivalent streptavidin that is bound to the plasmid carrier. Future work could compare average events with quadrivalent streptavidin and average events with

quadrivalent streptavidin and biophosphatase. This would illuminate differences between the different binding scenarios. Additionally, it could be interesting to compare these to another biotinylated protein that is much larger than biophosphatase.

TRANS-TO-CIS SINGLE-MOLECULE BINDING CURVES

5.13 CONTROLS

Similar controls were performed for trans-to-cis as for cis-to-trans, although only with monovalent streptavidin, and again can be split into two groups, controls with protein only and controls with plasmid. The protein only controls were recorded at the maximum and minimum protein concentrations for monovalent streptavidin, and at no concentration were any events shown over long periods of time, such as one or two minutes. These were repeated at least three times and in the same conditions as the binding curves, suggesting that the monovalent streptavidin still does not cause any events on its own.

For the controls with plasmid, the concentration of the plasmid was kept at 100 pM, like the concentration of the plasmid in any of the binding curve samples. The monovalent streptavidin concentration was always the same as they were performed at a 1:1 ratio. Additionally, they were conducted in the same conditions as all of the binding curves. The controls performed included a control for each step of the plasmid modification process to make sure that no step allowed the monovalent streptavidin to bind without the presence of the oligonucleotide. The controls are PCSAL1 and 1XPCSAl1, PC and 1XPC, PCstepsB and 1XPCstepsnew, PCstepsA and 1XPcstepsold, PCNbBVstepsB and 1XPCNbBVstepsB, PCNbBVstepsA and 1XPCNbBVstepsA, PCNbBVLigstepsA and 1XPCNbbBVLigstepsA, PC2ELigfilA and 1XPC2ELigfilA, and PC2ELigfilB and 1XPC2ELigfilB. All of these controls had similar dwell times, peak amplitudes, and histogram shapes with and without protein suggesting that in each case the monovalent streptavidin was not able to bind to the carrier.

5.14 COMPARISON BETWEEN PROCTOCOLS A AND B

For trans-to-cis, the Protocol A and B plasmid carrier binding curves were only performed with the one biotin oligonucleotide. The plasmid modified by either protocol was incubated with monovalent streptavidin at various different concentrations, 50 pM, 100 pM, 1 nM, 5 nM,10 nM, 25 nM, and 100 nM, while its concentration was kept constant at

100pM. All of these concentrations were recorded in the same pipette. Sample current traces and events for PC2ELigBOfilA and 50X are shown in **Figure 5.28**. There is a difference in the general peak amplitude, however this time the trace with protein has the lower peak amplitude, and again the sample with protein is shown in turquoise and without in navy. There is also a difference between the event charges, with the charges of the samples without protein being larger than the charges of those with protein. However, the event charges are not shown in the sample events as this analysis was not used to distinguish events, but is an area that should be explored further. The current trace also shows much fewer events than the cis-to-trans data, and that is because at the same concentration trans-to-cis will have a much lower capture rate compared to cis-to-trans. The sample events have much greater difference than the general recording. Event a in **Figure 5.28** has a dwell time of 0.7 ms and peak amplitude of 65.0 pA, event b a dwell time of 0.6 ms and peak amplitude of 60.4 pA, event c a dwell time of 0.9 ms and peak amplitude of 36.2 pA, and event d a dwell time of 0.8 ms and peak amplitude of 30.5 pA. Events a and b are from the PC2ELigBOphfilB recording and c and d are from 50X. The lower peak amplitude for events with protein, shown in both the current trace and the sample events, makes sense as the pI of streptavidin is 5, so it will be negative at pH 8.[5] A lower peak amplitude means that the current does not change as much as the complex moves through it. This can be explained by the streptavidin likely being less negative than the plasmid, so being able to bring fewer positive counterions with it. Also, as streptavidin is more compact than the plasmid generally, it will displace fewer of the negative ions in the solution. The sample events, despite having the lower peak amplitude have clear subpeaks that do not exist in the carrier only samples.

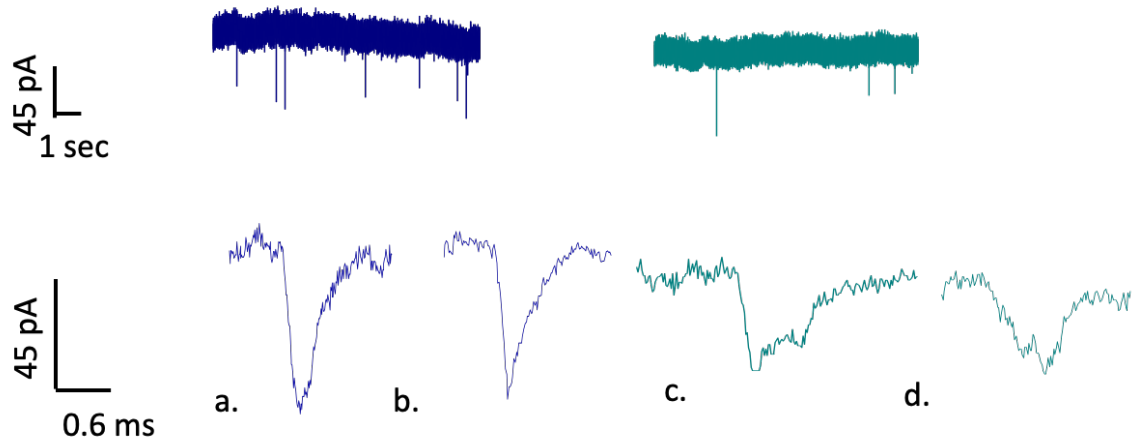


***Figure 5.28**: Sample events and current traces for PC2ELigBOfilA and 50X. The PC2ELigBOfilA is shown in navy and the 50X in turquoise.*

Sample current and traces are also shown for PC2ELigBOphfilB for trans-to-cis in **Figure 5.29**. They were performed in similar manner to the results presented before and in the same conditions. These again agree with Protocol A results, where the peak amplitude, particularly overall, is higher for the carrier without protein. Event a in **Figure 5.29** has a dwell time of 0.7 ms and peak amplitude of 35.8 pA, event b a dwell time of 0.8 ms and peak amplitude of 54.3 pA, event c a dwell time of 0.9 ms and peak amplitude of 33.6 pA, and event d a dwell time of 1.0 ms and peak amplitude of 42.6 pA. Events a and b are from the PC2ELigBOphfilB recording and c and d are from 50X. While the events are a bit different in terms of dwell time and peak amplitude, the most important difference is in shape. The events with protein have a clear subpeak and are not triangular as opposed to the ones without.

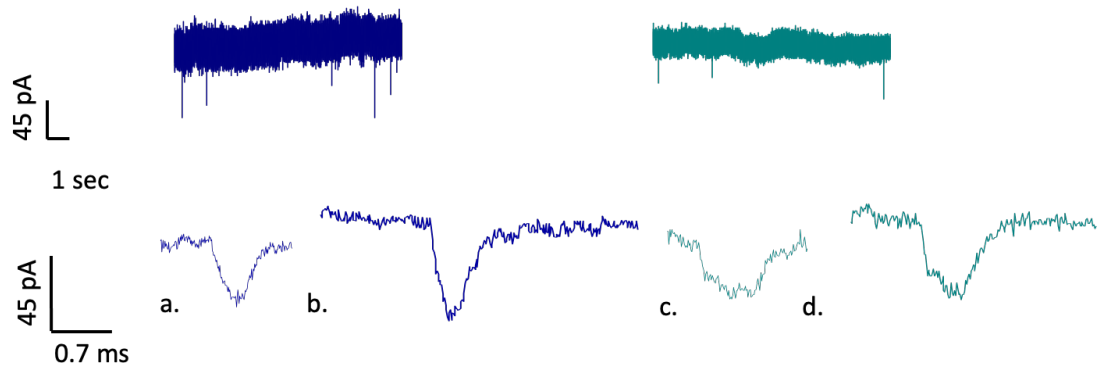


***Figure 5.29**: Sample events and current traces for PC2ELigBOphfilB and 50X. The PC2ELigBOphfilB is shown in navy and the 50X in turquoise.*

Then the binding curve results for Protocols A and B were compared. The binding curves measured how the percentage of subpeaks, taken from the automated results, as the concentration of monovalent streptavidin increased. Binding curves for PC2ELigBOfilA and PC2ELigBOphfilB are shown in **Figure 5.30**. The error bars are again calculated from the standard deviation of all of the results at each concentration. Each point has data from at least four pipettes for both binding curves. The pipette size varied from 17-23 nm PC2ELigBOfilA and 15-22 nm for PC2ELigBOphfilB. This is much less variation than in the cis-to-trans results. Perhaps this lack of variation explains why the data is much more similar for each pipette for both of these binding curves. Additionally, the percentages overall continuously rise as the protein concentration is increased. The plasmid carrier alone folded percentages are 16.1 ± 8.2% for PC2ELigBOfilA and 22.4 ± 3.3% for PC2ELigBOphfilB. For Protocol A, significant difference between the 100 pM, 1 nM, 5 nM, and 100 nM protein samples and the folded events was determined using a two sample T test. For Protocol B, significant difference between the 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM and 100 nM protein samples and the folded events was determined using a two sample T test. This means that for every protein concentration in Protocol B, there is a significant difference for the samples with protein. Protocol B also has much higher subpeak percentages, and even crosses the 50% threshold. This suggests that Protocol B, given that it uses an oligonucleotide with phosphorylation on the 5’ end, is able to incorporate the oligonucleotide more than Protocol A.

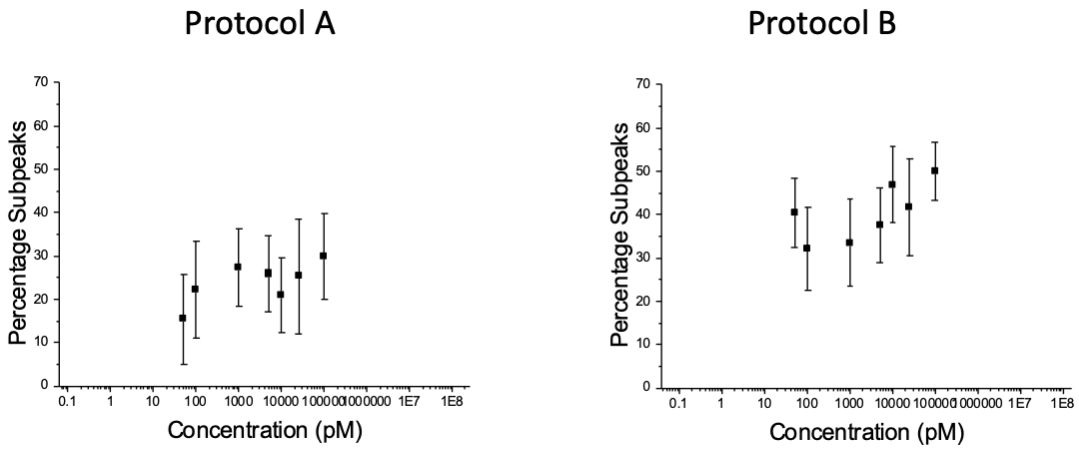


***Figure 5.30**: A comparison of the binding curves for Protocols A and B of the plasmid with one biotin. The protein concentrations are 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM, and 100 nM for monovalent streptavidin. The folded event percentages are 16.08 ± 8.24% and 22.39 ± 3.34% respectively.*

5.15 COMPARISON BETWEEN ONE AND TWO BIOTINS

As the two biotin oligonucleotide carrier was only recorded using Protocol B, this section will compare one and two biotin carriers for Protocol B only. Like with previous samples, the plasmid with two biotins was incubated with monovalent streptavidin at various different concentrations, 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM, and 100 nM, while its concentration was kept constant at 100 pM. All of these concentrations were recorded in the same pipette. Sample current traces and events are shown in **Figure 5.31**. Like the other trans-to-cis current traces, there are fewer events and the sample with protein generally has lower peak amplitudes than the one without. Additionally, there is a difference between the event charges, with the charges of the samples without protein being larger than the charges of those with protein. However, the event charges are not shown in the sample events as this analysis was not used to distinguish events, but like with the one biotin samples, it is an area that should be explored further. Event a in **Figure 5.31** has a dwell time of 0.7 ms and peak amplitude of 57.8 pA, event b a dwell time of 0.7 ms and peak amplitude of 50.0 pA, event c a dwell time of 1.0 ms and peak amplitude of 44.6 pA, and event d's dwell time is 1.1 ms and peak amplitude is 47.0 pA. Events a and b are from the PC2ELigBO2filB recording and c and d are from 50X. The events without protein have very different shapes than the ones with. They are sharp and triangular, while the ones

with protein are more square and often have a clear change in levels that suggests a subpeak.

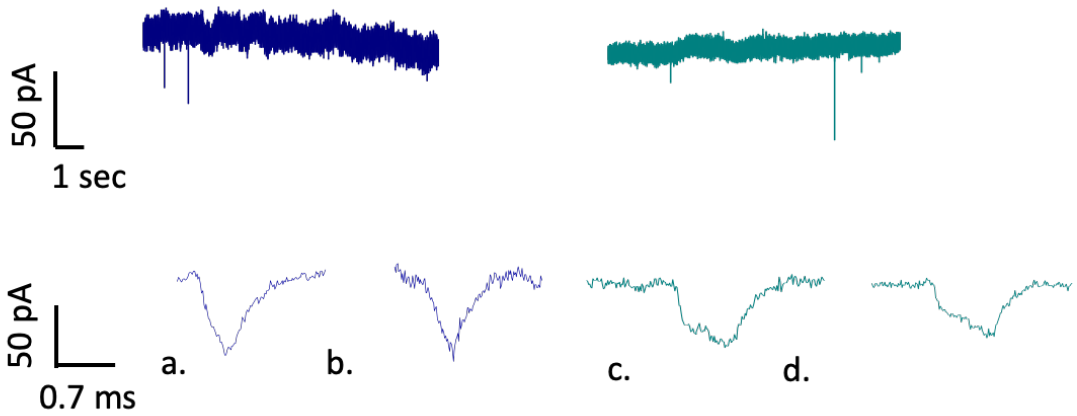


***Figure 5.31**: Sample events and current traces for PC2ELigBO2filB and 50X. The PC2ELigBO2newl is shown in navy and the 50X is in turquoise.*

The binding curves for the samples with one and two biotins were compared. They were calculated in the exact same manner as described previously in this chapter and are shown in **Figure 5.32**. Each point has data from at least four pipettes for the one biotin binding curve and five for the two biotin binding curve. The pipette size varied from for 15-22 nm PC2ELigBOphfilB and 13-21 nm PC2ELigBOphfilB. Again, there is much less variation in size than with the cis-to-trans events. The plasmid carrier alone folded percentages are 22.4 ± 3.3% for PC2ELigBOphfilB and 22.4 ± 8.9% for PC2ELigBO2filB. For one biotin, significant difference between the 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM and 100 nM protein samples and the folded events was determined using a two sample T test. For Protocol B, significant difference between the, 5 nM, 10 nM, and 100 nM protein samples and the folded events was determined using a two sample T test. Interestingly, the two biotins did not have the sample clear trend as the one biotin sample. Perhaps, this is because there are not two clear subpeaks for each plasmid that binds two biotins and therefore this reduces the overall amount of subpeaks that are detected. However, this data suggest that it is possible to detect monovalent streptavidin using a plasmid with one or two biotins in a nanopore.

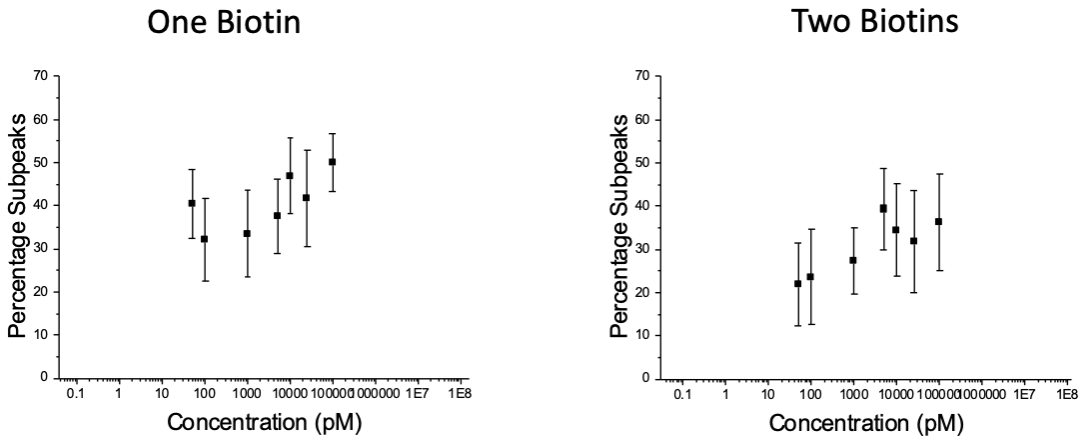


***Figure 5.32**: A comparison of the binding curves for the plasmid with one and two biotins The protein concentrations are 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 25 nM, and 100 nM for monovalent streptavidin. The folded event percentages are 22.4 ± 3.3% and 22.4 ± 8.9% respectively.*

CONCLUSIONS

In conclusion, it is clear that the plasmid is able to bind to proteins and that these events are clearly visible in the nanopore. There was a clear increase in the percentage of subpeak events for both the quadrivalent and monovalent streptavidin binding curves compared to the plasmid carrier alone in the cis-to-trans direction. However, for the multiplexed sensing, the sandwich-like assay had mixed results, where it is possible to see an effect of the biophosphatase binding, but it is not possible to see this very clearly in individual events. This lack of resolution could stem from the size of the biophosphatase in relation to the 10 Kbp DNA. As such, the increase in subpeak dwell time and amplitude is minimal compared the overall event and therefore would not be easily apparent. It was also shown that the monovalent streptavidin is able to bind to the plasmid carrier and have clear events showing that binding in the trans-to-cis direction. Further work could include adding a machine learning component to the subpeak selection process, as this would allow for more consistent selection, as well as better selection of subpeaks. This better selection of subpeaks could allow for binding curves to be recorded at lower protein concentrations. The lack of the sigmoidal shape for these binding curves could be due to the relatively high protein concentrations that were required to obtain significantly different results to the samples without protein.

REFERENCES

1. Dundas, C. M., Demonte, D. & Park, S. Streptavidin-biotin technology: Improvements and innovations in chemical and biological applications. *Appl. Microbiol. Biotechnol.* **97**, 9343–9353 (2013).
2. Howarth, M. *et al.* A monovalent streptavidin with a single femtomolar biotin binding site. *Nat. Methods* **3**, 267–273 (2006).
3. Bell, N. A. W. & Keyser, U. F. Specific protein detection using designed DNA carriers and nanopores. *J. Am. Chem. Soc.* **137**, 2035–2041 (2015).
4. Murphy, J.E., Tibbitts, T.T., Kantrowitz, E. R. Alkaline Phosphatase. *RCSB PDB*
5. Streptavidin. *Invitrogen* Available at: https://www.fishersci.com/shop/products/pierce-streptavidin-protein-conjugates-3/pi21323.

Chapter 6 Conclusions

Table of Contents

Synopsis: This chapter offers a summary of the conclusions presented in this thesis for both DNA carriers. Additionally, it discusses various future work outcomes and wider applications for the plasmid carrier.

PROJECT AIMS AND OUTCOME

The aim of this thesis was to explore two potential novel DNA carrier structures for use in nanopore sensing. This included establishing their successful construction and modification, and their distinct differentiable signals both with and without bound protein. Exploring the use of these carriers in multiplexed sensing offered another avenue of investigation.

The necessity of DNA carriers, and the challenges associated with protein signalling particularly, are discussed in Chapter 1. Chapter 2 summarised the general materials and methods used throughout this research, including the fabrication of the nanopipettes.

Chapter 3 discusses the feasibility of using DNA dendrimers as a carrier. It was established that dendrimer structures with thirty base overhangs were constructed, while structures with lengths of four and thirteen were also explored. The shorter base overhangs were not long enough to form stable dendrimers. Gel electrophoresis was used to confirm the formation of these structures, and it was quite clear that they were fabricated successfully. However, when it came to running the dendrimer through the nanopore, there were significant issues with dendrimer stability, particularly due to the salt conditions. A salt study was performed, and even at 100 mM KCl, the dendrimers appeared to regularly fall apart back into their Y structures. Thus, these traces did not contain many clear translocation events. Thrombin binding, with an incorporated aptamer, was also investigated. However, again, there were not many translocation events, due to the instability of the dendrimers. Thus, it was established that the dendrimers were not acceptable DNA carriers.

After the DNA dendrimers were deemed unacceptable carriers, a 10 Kbp plasmid was explored in Chapter 4. The plasmid was modified to incorporate an oligonucleotide with a biotin molecule and to linearise it. This protocol was tailored and then, further changed and optimised due to Covid-19 pandemic restrictions. The successful modification of the plasmid was confirmed using a glass binding experiment with fluorescence. The behaviour of the plasmid in the nanopore was then examined. Voltage studies confirming that it

behaved with expected trends and displayed similar behaviour to 10 Kbp DNA were performed. A pH study was also performed, and it was established that pH 8.0 presented the optimal experimental conditions.

Once the modification of the plasmid carrier and its general nanopore behaviour were determined, extensive binding curves with biotin and streptavidin were performed for the plasmid, in Chapter 5. Biotin and streptavidin were selected as they are extremely well studied. This chapter also includes a detailed discussion of the automated subpeak selection process. Automation ensured that the criteria used to select subpeaks were rigorous and the high volume of translocation events eliminated the option of manual selection as a feasible process. Cis-to-trans binding curves for plasmids that had one or two biotins incorporated were performed. Protocol A had a clearer upward trend for the binding curve, however both showed significant numbers of subpeaks. Contrary to the previous samples, binding curves for the plasmid carriers with two biotins indicated that despite a general upward trend, the two highest concentrations deviated from this and had fewer subpeaks. This potentially could stem from large amounts of monovalent streptavidin blocking the pore or larger amounts of the plasmid successfully binding the monovalent streptavidin and then blocking the pore more easily. A comparison for monovalent streptavidin and quadrivalent streptavidin was also performed. The quadrivalent streptavidin had significantly different subpeak percentages at lower protein concentrations, but not at higher ones. This perhaps could be due to the quadrivalent streptavidin binding multiple plasmids. A sandwich-like assay was also performed with quadrivalent streptavidin and biotinylated biophosphatase. The binding curve looks significantly different for the biophosphatase compared to the quadrivalent streptavidin, suggesting that there is definitely some effect, but on an individual level the subpeaks themselves did not have a drastically different shape. Trans-to-cis binding curves were also performed for monovalent streptavidin. These had significantly higher percentages of subpeaks compared to cis-to-trans, and should be explored more. Thus, the effectiveness of the plasmid as a carrier was shown.

CONCULSIONS, OUTLOOKS, AND FUTURE WORK

The plasmid, while being established as a successful DNA carrier, still requires more investigation and optimisation. There are some challenges of using nanopipettes in biomedical applications, particularly for the sensing of proteins, as proteins are not uniformly charged and nanopipettes typically cannot have diameters < 10 nm. The conical shape formed in nanopipettes also leads to uneven forces that can affect expected results. Furthermore, these can be amplified if there are differences in pipette diameter and shape. All of these differences in pipette shape can have profound effects on the results of the binding curves. As such, it could be interesting to performe experiments with the plasmid as a DNA carrier in another type of solid state pipette.

As trans-to-cis was not a major focus on this thesis, but the early results look promising, a more detailed exploration of this experimental setup should be performed. Furthermore, with the increased investigation into the trans-to-cis direction, it would also be valuable to see how the DEP affects the plasmid in both directions. The concentrating effect of the DEP, when it traps DNA, could explain some of the discrepancies in the binding curve results, particularly as the variation in pipette size meant that the DEP could vary greatly between pipettes. Perhaps a more stringent selection of pipettes with both similar diameters and taper lengths could be used to see how the DEP affects the results. Additionally, improvement into the modification process itself could greatly increase the sensitivity of this DNA carrier. The successful modification could be quantified using similar glass slide binding experiments, which would allow for optimisation of every step of the modification process. This would also help improve the accuracy of the subpeak percentages, as the amount of carrier that was not able to bind to protein would be minimised. Additionally, improving the automated process, such that subpeaks could be selected with greater accuracy would be an area that ought to be expanded upon. This automation would also have a diverse set of applications as many other DNA carriers exhibit subpeaks. Machine learning could be applied to this, particularly as it would be relatively simple to acquire the large amounts of data required for training. Further work could also include a deeper investigation of the sandwich-like assay with biotin and streptavidin to see if there are distinctive subpeaks. Additionally, adding an aptamer to the oligonucleotide instead of biotin, perhaps a thrombin one, would allow for a more practical application of this carrier. This carrier's effectiveness in selecting specific proteins could also be measured using a serum sample. Lastly, the plasmid's design could be modified such that it would be

possible to insert two oligonucleotides with at least several thousand bases in between them, to see if two distinct subpeaks appear when two proteins are bound to the oligonucleotides, to confirm its effectiveness at multiplexed sensing.

In short, this plasmid has many potential applications as a carrier. It is completely customisable and quite reliable, as shown in this thesis. As such, its application into a nanopore sensor should be explored further. It has the potential to be a widely applicable, label-free, rapid, low-cost sensing tool for biological samples. Finally, it even offers the possibility of the identification and quantification of various biological analytes of medical interest.

APPENDIX I: SUPPLEMENTARY FIGURES

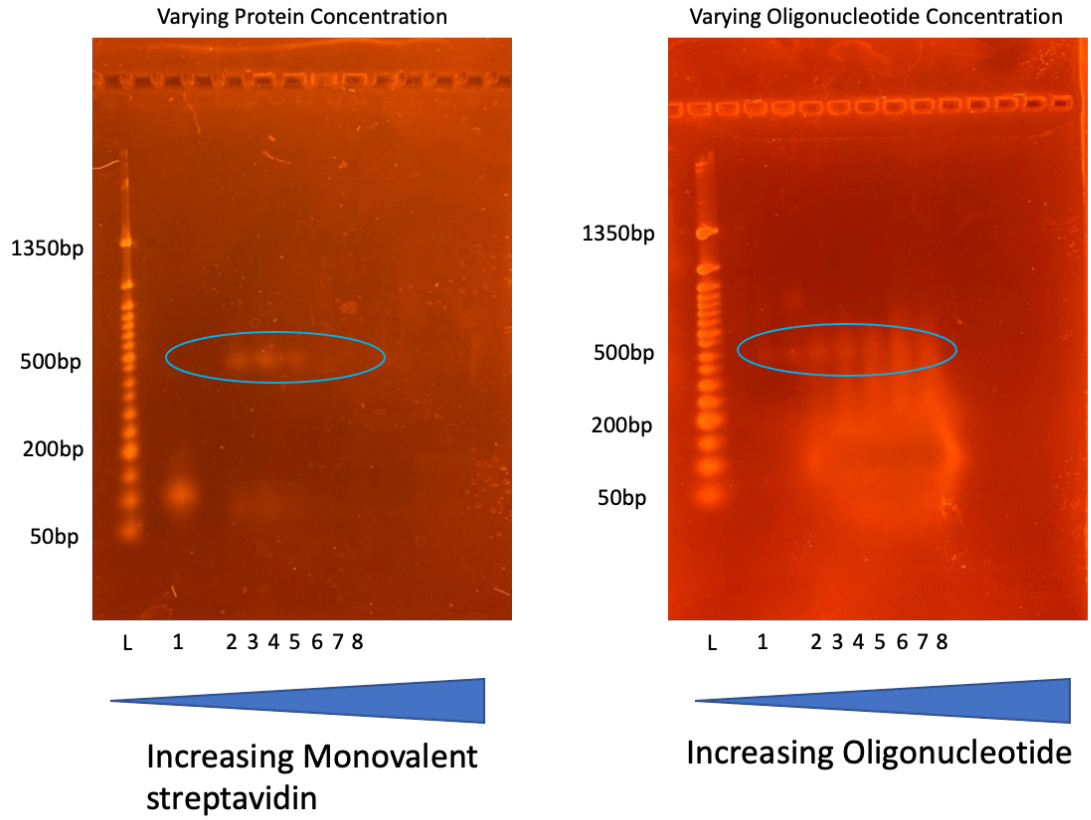


Figure 1: *These 2% agarose gels show that the phosphorylated oligonucleotide is able to bind to monovalent streptavidin. The ladder ranges from 50 bp to 1350 bp. For varying protein concentration the wells are 1. BOph, 2. 0.5X (monovalent streptavidin to oligonucleotide), 3. 1X, 4. 10X, 5. 50X, 6. 100X, 7. 250X, and 8. 1000X. The wells for the varying oligonucleotide concentration are 1. Monovalent streptavidin, 2. 0.5X (oligonucleotide to monovalent streptavidin), 3. 1X, 4. 10X, 5. 50X, 6. 100X, 7. 250X, 8. 1000X, 9. BOph only. The bands where binding occurred are marked with a circle.*

Figure 1 shows that this new oligonucleotide with phosphorylation is also able to successfully bind to monovalent streptavidin The agarose gel concentration for both of these gels is 2%, so that the shift that occurs between the unbound and bound oligonucleotide is shown clearly. The Varying Protein Concentration gel keeps a constant amount of oligonucleotide while gradually increasing the amount of protein. Unlike with BO, the highest ratio clearly visible is only 10X, but there seems to be only one band at this ratio, which would suggest that most of the oligonucleotide has bound the monovalent streptavidin. Similarly, the necessary dilution of the oligonucleotide to get the required protein ratios is probably the reason why the higher protein bands cannot be seen. The

Varying Oligonucleotides Concentration gel shows that as there is more oligonucleotide a smaller percentage of it is bound. On this gel, unlike for the previous oligonucleotides, it is possible to see that a similar amount of it is bound to the protein even as the oligonucleotide concentration increases, and the unbound oligonucleotide simply moves further down the gel.

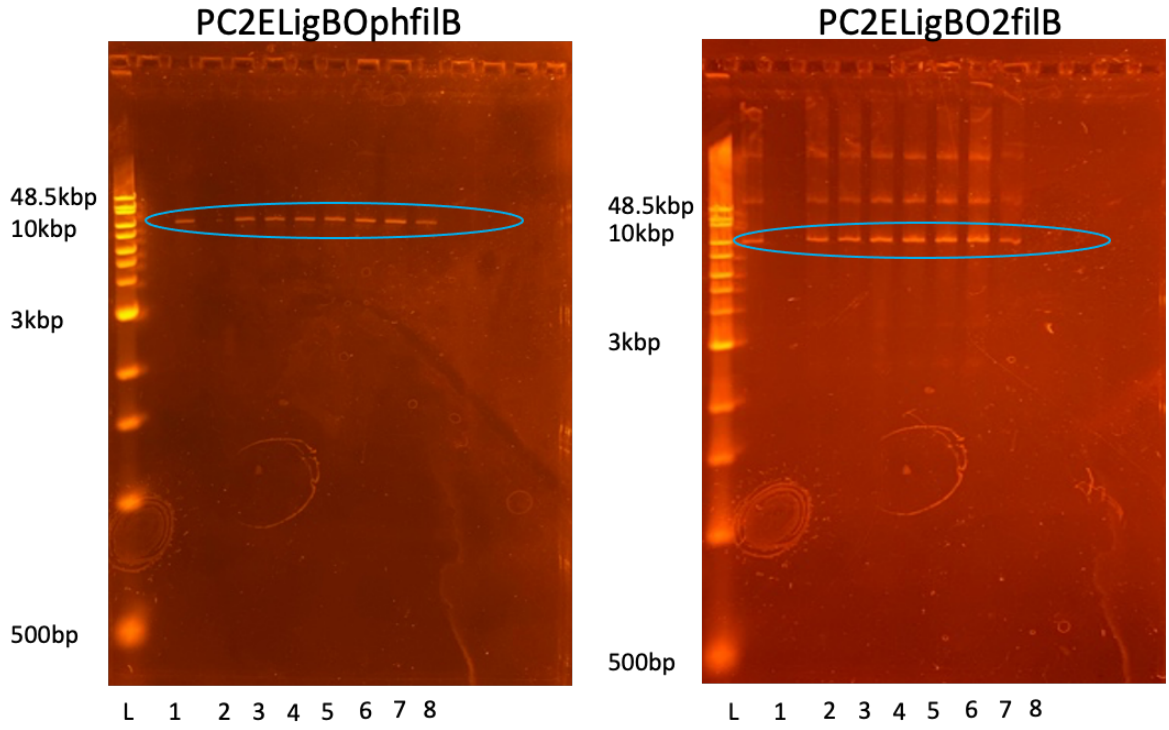


Figure 2: *1% agarose gels showing the binding curve for PC2ELigBOphfilB and PC2ELigBO2filB with monovalent streptavidin. Well 1 is PC2ELigBOphfilB or PC2ELigBO2filB, well 2 is 0.5X (ratio of monovalent streptavidin to plasmid), well 3 is 1X, well 4 is 10X, well 5 is 50X, well 6 is 100X, well 7 is 250X, and well 8 is 1000X. The ladder for this gel is 500 bp to 48.5 Kbp. The 10 Kbp bands are marked with a circle.*

Binding curves for the new plasmid protocol with one and two biotins. For both gels, well 1 is PC2ELigBOphfilB or PC2ELigBO2filB, well 2 is 0.5X (ratio of monovalent streptavidin to plasmid), well 3 is 1X, well 4 is 10X, well 5 is 50X, well 6 is 100X, well 7 is 250X, and well 8 is 1000X. Unfortunately, it was not possible to see any shift with the bound protein, probably because the shift is not very big. Even in the PC2ELigBO2filB gel, there are multiple bands as the SAL1 did not cut properly rather than the binding of the protein.

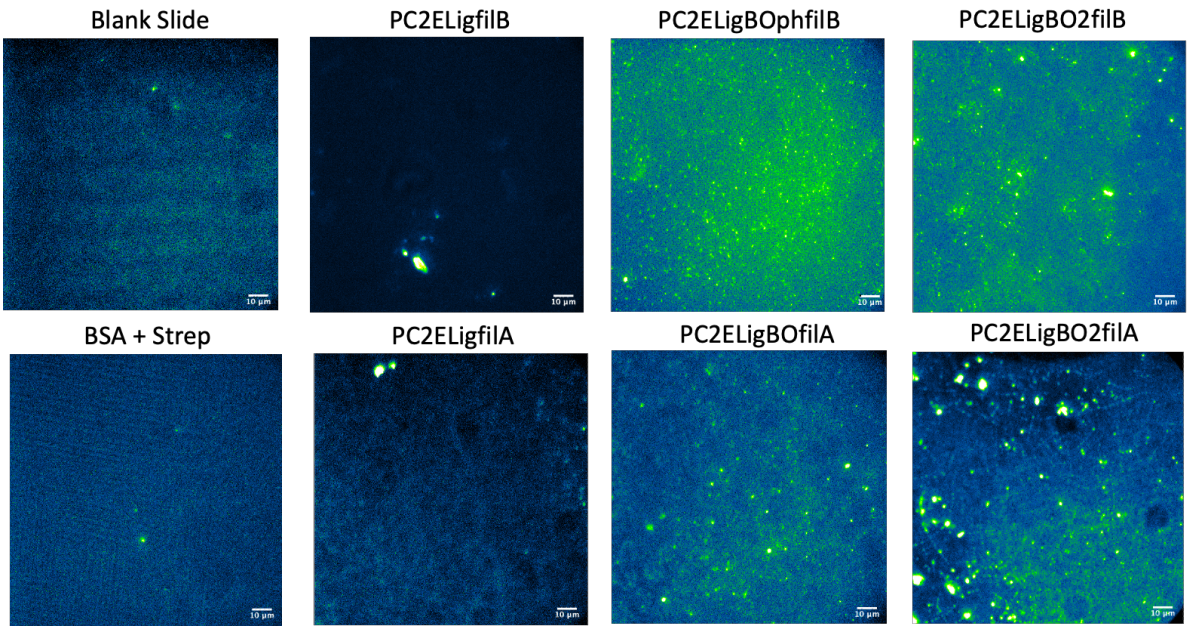


***Figure 3**: Frame images for each sample from the glass slide control experiments. The samples that should bind to the slide, PC2ELigBOphfilB, PC2ELigBO2filB, PC2ELigBOfilA, and PC2ELigBO2filA, have much more fluorescence than any of the control slides. Additionally, individual dots are seen quite clearly. Some of these, particularly for PC2ELigBO2filB and PC2ELigBO2filA, are a bit larger and brighter, and this may be due to the higher probability of closer binding for the samples with two biotins. These results show that the plasmid has been successfully modified. Images taken by Annie Sahota.*

Sample slides for the binding of different plasmid samples. PC2ELigBO2filB and PC2ELigBO2filA seem to have large spots of very high fluorescence. These spots are not large enough to suggest that dirt or other contaminants, especially since the slides were plasma cleaned before the experiments. Perhaps when they have two biotin binding sites, the probability of several plasmid carriers binding close together is higher, causing these brighter spots. This figure also confirms what is shown in **Figure 4.30**, that the samples that have the biotin oligonucleotide added are able to bind to the slide and fluoresce. The level of fluorescence for the control slides is much lower.

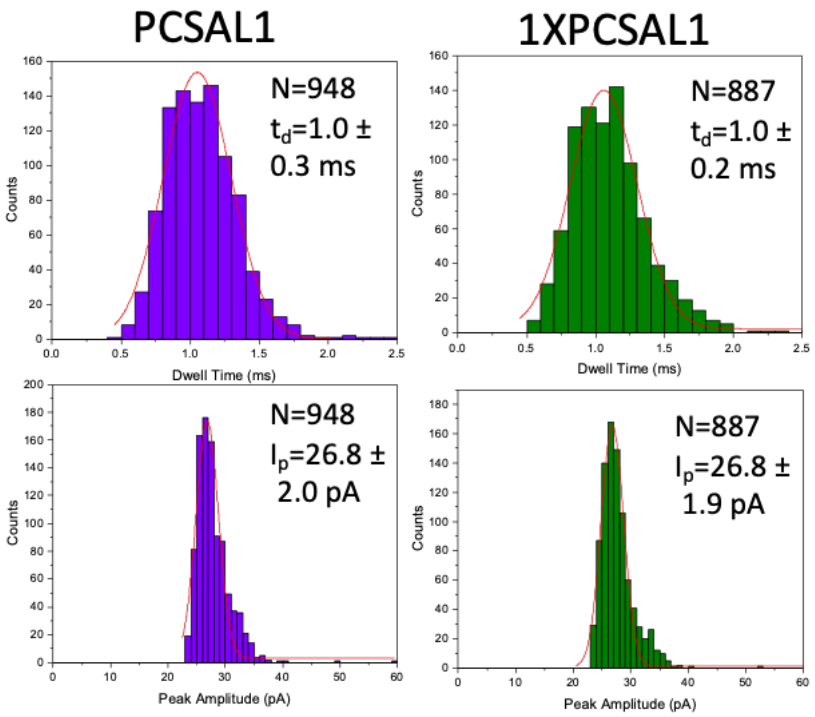


***Figure 4**: Example histograms of the linearised plasmid control. The similar dwell times and peak amplitudes for the plasmid with and without protein suggest that it is not able to bind to the monovalent streptavidin. A Gaussian fit was used to determine the average values.*

Figure 4 is PCSAL1, which is the plasmid where it does not go through all of the modification steps, but it only is incubated with the linearising enzyme, SAL1. It was filtered afterwards. The dwell times and peak amplitudes for PCSAL1 and 1XPCSAL1 are similar, suggesting that the monovalent streptavidin is not able to bind to the linearised plasmid. This control was performed in three different pipettes with very similar results.

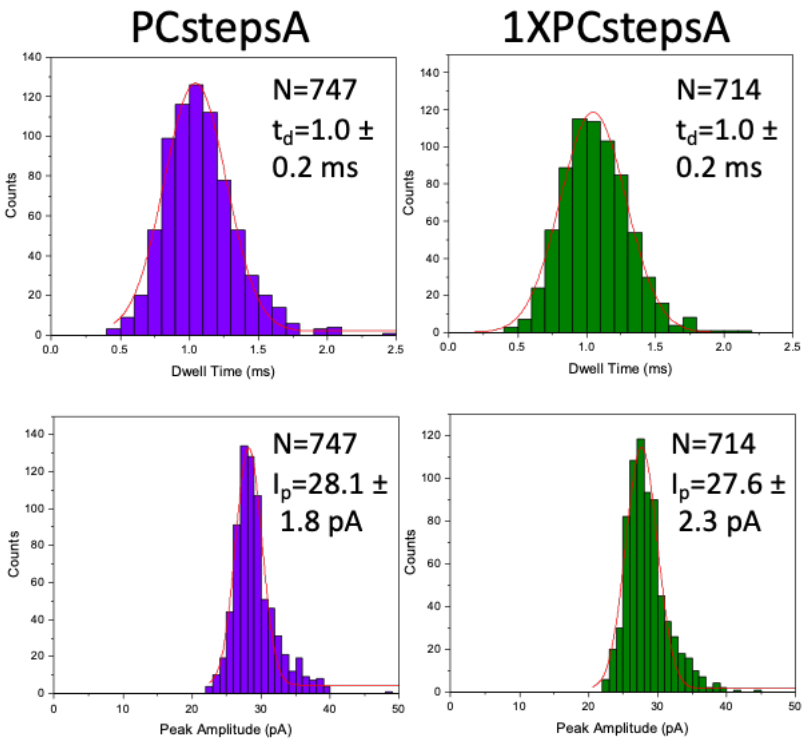


***Figure 5**: Histograms of PCstepsA control with and without protein. The similar dwell times and peak amplitudes confirm that the monovalent streptavidin is not able to bind to the plasmid alone even after the modification steps. A Gaussian fit was used to determine the average values.*

The next control is the plasmid that has gone through all of the Protocol A modification steps, but did not have any oligonucleotide or enzymes added. **Figure 5** shows sample histograms for PCstepsA and 1XPCstepsA. The dwell time and peak amplitude are again near in value, and the histograms have a similar shape for both the sample with protein and without. This control was repeated in three different pipettes with similar outcomes. These results therefore suggest that the monovalent streptavidin is not able to bind the plasmid alone even when it has gone through all of the modification conditions.

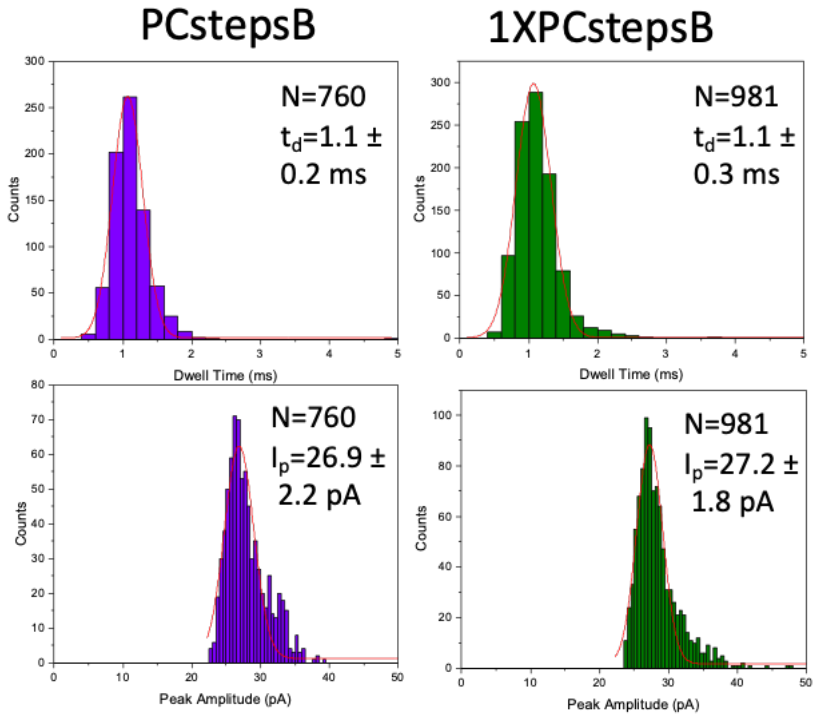


***Figure 6**: Histograms of PCstepsB control with and without protein. The similar dwell times and peak amplitudes confirm that the monovalent streptavidin is not able to bind to the plasmid alone even after the modification steps. A Gaussian fit was used to determine the average values.*

Similarly, another control is the plasmid that went through all of Protocol B modification steps. This control is important as particularly steps, like the ligation step, had drastically different conditions in both protocols. ThePCstepsB control results are shown in **Figure 6**, and have similar results to PCstepsA. The dwell times and peak amplitudes are similar with and without monovalent streptavidin. This control was also repeated in three different pipettes.

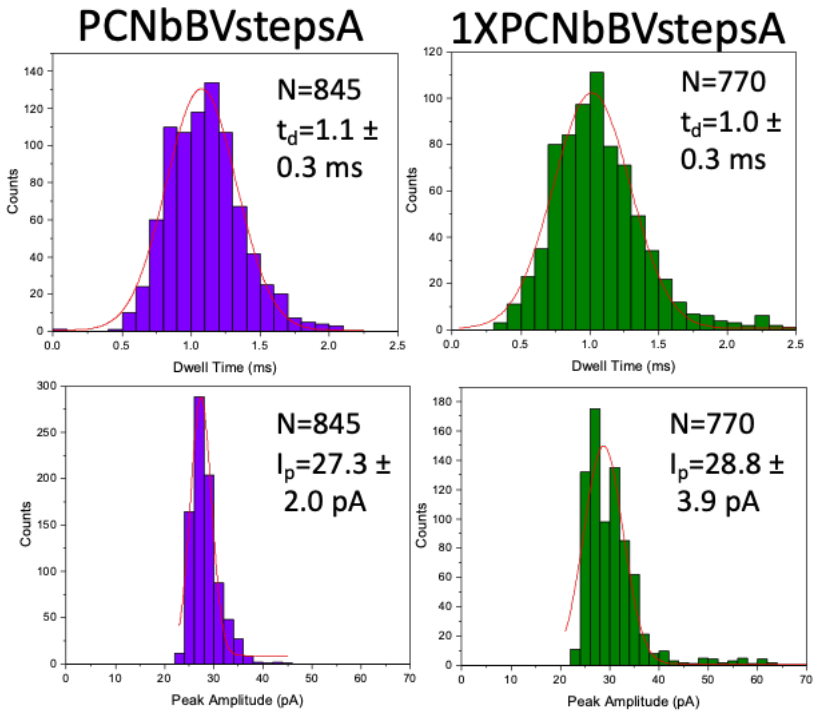


***Figure 7**: Histogram results for PCNbBVstepsA with and without protein. The dwell times and peak amplitudes are quite similar, and therefore the monovalent streptavidin is unlikely to be binding in the sample that was incubated with protein. A Gaussian fit was used to determine the average values.*

The next control is the nicked plasmid that went through all of the modification steps, PCNbBVsteps, without any added SAL1, ligase, or oligonucleotide. **Figure 7** shows the results for PCNbBVstepsA with and without protein. These samples went through Protocol A. Again, the dwell times and peak amplitudes are similar, and thus it is reasonable to conclude that it is not the nicking enzyme that allows monovalent streptavidin to bind to the plasmid. This control, like the previous ones was repeated in three different pipettes.

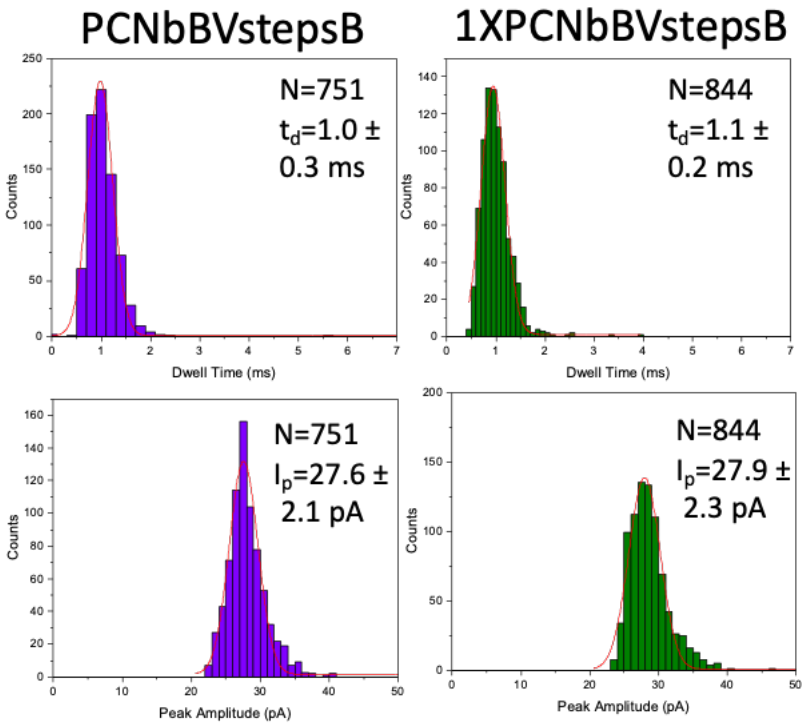


***Figure 8**: Histogram results for PCNbBVstepsB control with and without protein. These demonstrate that the monovalent streptavidin is not able to bind to the plasmid when it has only been nicked even after going through all of the modification steps. A Gaussian fit was used to determine the average values.*

This control was repeated with Protocol B, that is for PCNbBVstepsB, and shown in **Figure 8**. These results also demonstrated that the monovalent streptavidin is unlikely to be binding to the plasmid sample, as the dwell times and peak amplitudes are similar. Additionally, the results were repeated in three different pipettes.

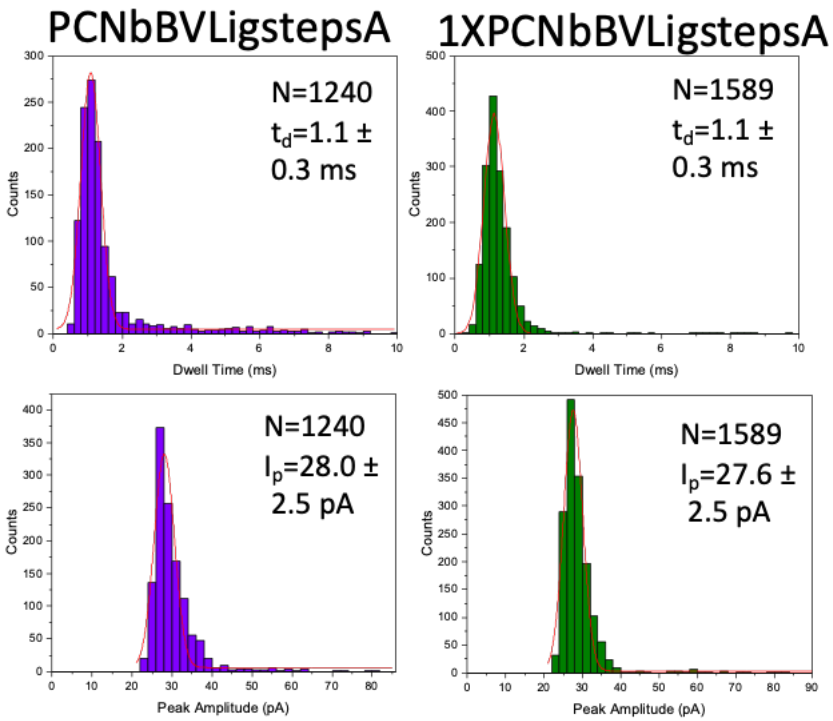


***Figure 9**: Histograms showing that the dwell times and peak amplitudes do not vary between PCNbBVLigstepsA with and without monovalent streptavidin. This shows that the protein is not able to bind. A Gaussian fit was used to determine the average values.*

The control PCNbBVLigstepsA was also run. The plasmid went through all of the modification steps, but only was incubated with the nicking enzyme and the ligase. **Figure 9** shows the results of these experiments with and without monovalent streptavidin. The dwell times and peak amplitudes are similar and the results were repeated three times, suggesting that the monovalent streptavidin is not able to bond to the plasmid.

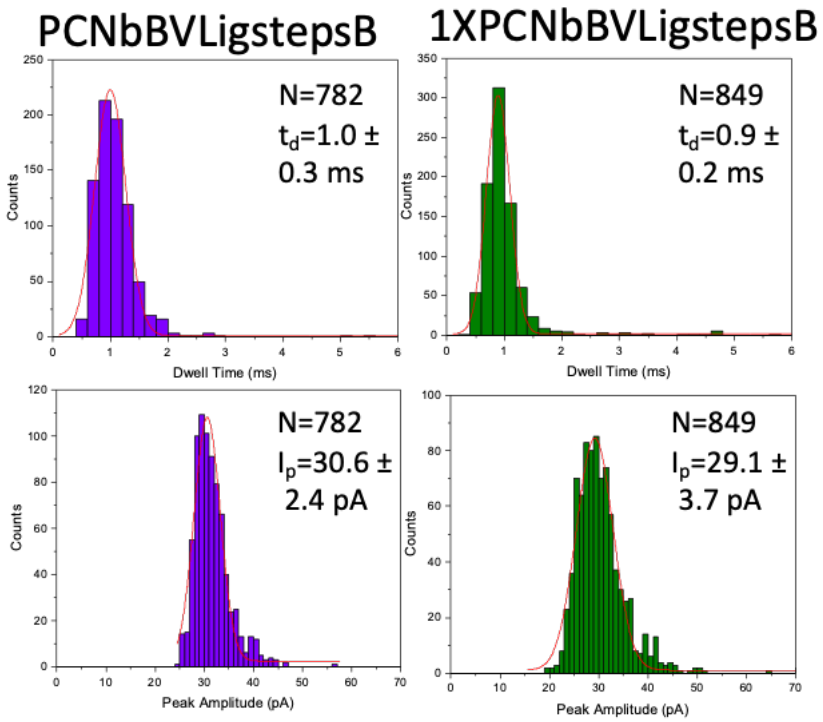


***Figure 10**: Comparison of controls results for PCNbBVLigstepsB and 1XPCNbBVLigstepsB. The similarity between dwell times and peak amplitudes demonstrates that the monovalent streptavidin is not able to bind. A Gaussian fit was used to determine the average values.*

This control was also repeated with Protocol B, PCNbBVLigstepsB. The histograms for these experiments are shown in **Figure 10**. The lack of any shift in the dwell time and peak amplitude shows that the monovalent streptavidin is not able to bind. These experiments were also repeated three times.

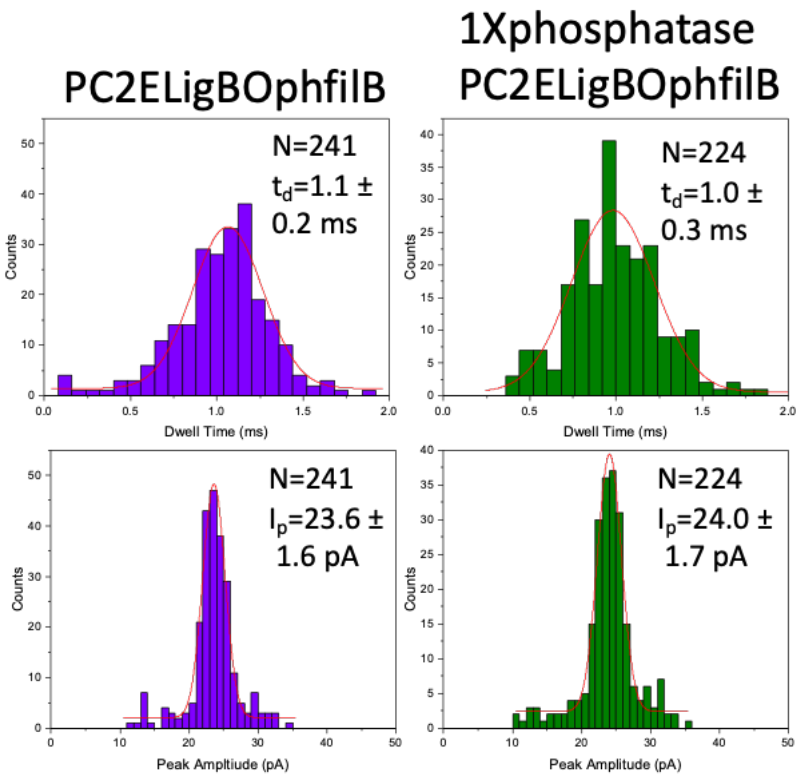


***Figure 11**: Histograms for PC2ELigBOphfilB with and without phosphatase. As this phosphatase has not been biotinylated, and there is no quadrivalent streptavidin it should not be possible for it to bind to the plasmid carrier. The similar values for dwell time and peak amplitudes and the histogram shapes suggest that this is the case. A Gaussian fit was used to determine the average values.*

The is control is PC2ELigBOphfilB incubated with phosphatase. This is a phosphatase that has not been biotinylated and should not be able to bind to the carrier even in the presence of quadrivalent streptavidin. The histograms for this control are shown in **Figure 11**. Again the dwell times and peak amplitudes are similar and so are the histogram shapes, suggesting that the phosphatase does not bind.

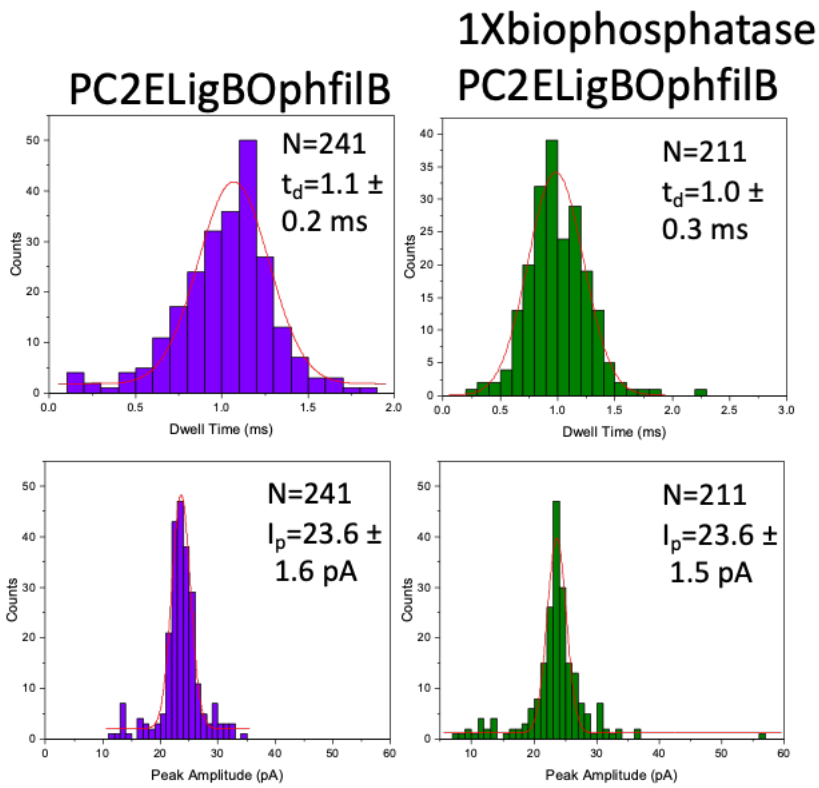


***Figure 12**: Histograms for PC2ELigBOphfilB with and without biotinylated phosphatase. The similar histogram shapes, as well as peak amplitudes and dwell times suggest that this phosphatase is not able to bind either. A Gaussian fit was used to determine the average values.*

This control is very similar to the previous one in that it is the biotinylated version of the phosphatase. However, as there is no quadrivalent streptavidin present in this control, it should not be able to bind to the carrier either. The similar shape of the histograms for the carrier with and without protein suggest that the biotinylated phosphatase is not binding. Additionally the dwell time and peak amplitudes are very similar. **Figure 5.12** shows these histograms.

APPENDIX II: COPYRIGHT PERMISSION

Figure Number	Source	Copyright Holder	Permission	Licence Number	Approval Number
1.1	*Phys. Life Rev.* 2012, **9**, 125–158	Elsevier	Granted	5292760992887	1
1.3	*Phys. A Stat. Mech. its Appl.* 2021, **582**, 126252	Elsevier	Granted	5292991161337	2
1.4	*Anal. Chem.* 1997, **69**, 4627–4633	ACS Publications	Granted	N/A	3
1.5	*Proc. Natl. Acad. Sci.* 2008, **105**, 417–421	PNAS	PNAS exclusive License to Publish	N/A	
1.5	Elsevier, 2013	Elsevier	Granted	1213808-1	4
1.6	*Science (80-.).* 1996 **274**, 1859–1866	Science	Open Access (under CCBY)	N/A	
1.7	*Nature* 2001 **412**, 166–169	Nature Publishing Group	Granted	5293000314956	5
1.7	*Nat. Mater.* 2003 **2**, 537–540	Nature Publishing Group	Granted	5293000444454	6
1.8	*Anal. Chem.* 2014 **86**, 835–841	ACS Publications	Granted	N/A	7
1.9	*Nat. Commun.* 2017 **8**, 1–8	Nature Publishing Group	Open Access (under CCBY)	N/A	
1.10	*Nat. Nanotechnol.* 2013 **11**, 1093–1097	Nature Publishing Group	Granted	5293001134964	8
1.11	*IEEE Sens. J.* 2018 **1748**, 1–1	IEEE	Granted	N/A	9
1.11	*Angew. Chemie - Int.*	John Wiley and Sons	Granted	5293001416608	10

	Ed. 2012 **51**, 4864–4867				
1.13	*Nat. Commun.* (2017).	Nature Publishing Group	Open Access (under CCBY)	N/A	
1.13	*J. Am. Chem. Soc.* 2015 **137**, 2035–2041	ACS Publications	Open Access (under CCBY)	N/A	
4.3	*Nat. Methods* 2006 **3**, 267–273	Nature Publishing Group	Granted	5293010594518	11

Approval 1

CCC RightsLink®

My Orders My Library My Profile Welcome ad4216@ic.ac.uk Log out | Help | FAQ

My Orders > Orders > All Orders

License Details

This Agreement between Imperial College London -- Alexandra Dias-Lalcaca ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Print Copy

License Number	5292760992887
License date	Apr 19, 2022
Licensed Content Publisher	Elsevier
Licensed Content Publication	Physics of Life Reviews
Licensed Content Title	Nanopores: A journey towards DNA sequencing
Licensed Content Author	Meni Wanunu
Licensed Content Date	Jun 1, 2012
Licensed Content Volume	9
Licensed Content Issue	2
Licensed Content Pages	34
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System
Institution name	Imperial College London
Expected presentation date	Jun 2022
Portions	Figure 2
Requestor Location	Imperial College London Molecular Sciences Research 82 Wood Lane London, W12 7SL United Kingdom Attn: Imperial College London
Publisher Tax ID	GB 494 6272 12
Total	**0.00 USD**

BACK

Approval 2

RightsLink®

My Orders My Library My Profile Welcome ad4216@ic.ac.uk Log out | Help | FAQ

My Orders > Orders > All Orders

License Details

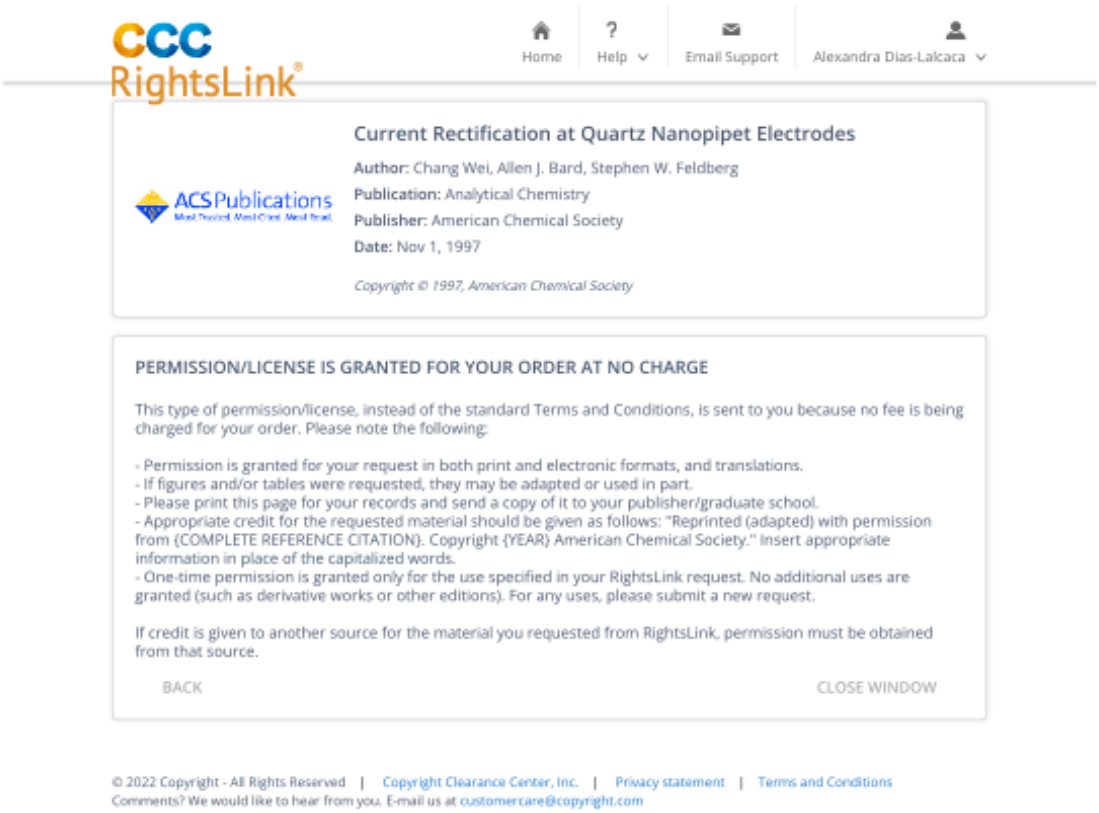
This Agreement between Imperial College London -- Alexandra Dias-Lalcaca ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Print Copy

License Number	5292991161337
License date	Apr 20, 2022
Licensed Content Publisher	Elsevier
Licensed Content Publication	Physica A: Statistical Mechanics and its Applications
Licensed Content Title	On the Gouy–Chapman–Stern model of the electrical double-layer structure with a generalized Boltzmann factor
Licensed Content Author	Anis Allagui,Hachemi Benaoum,Oleg Olendski
Licensed Content Date	Nov 15, 2021
Licensed Content Volume	582
Licensed Content Issue	n/a
Licensed Content Pages	1
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System
Institution name	Imperial College London
Expected presentation date	Jun 2022
Portions	Figure 1
Requestor Location	Imperial College London Molecular Sciences Research 82 Wood Lane London, W12 7SL United Kingdom Attn: Imperial College London
Publisher Tax ID	GB 494 6272 12
Total	**0.00 USD**

BACK

Approval 3


CCC RightsLink
Home
Help
Email Support
Alexandra Dias-Lalcaca
ACS Publications
Current Rectification at Quartz Nanopipet Electrodes
Author: Chang Wei, Allen J. Bard, Stephen W. Feldberg
Publication: Analytical Chemistry
Publisher: American Chemical Society
Date: Nov 1, 1997
Copyright © 1997, American Chemical Society
PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE
This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:
- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.
If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.
BACK
CLOSE WINDOW
© 2022 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Terms and Conditions
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Approval 4

CCC | Marketplace™

Order Number: 1213808
Order Date: 22 Apr 2022

Payment Information

Alexandra Dias-Lalcaca
ad4216@ic.ac.uk
Payment method: Invoice

Billing Address:
Alexandra Dias-Lalcaca
Imperial College London
Molecular Sciences Resear
ch 82 Wood Lane
London, W12 7SL
United Kingdom

+44 7957747156
ad4216@ic.ac.uk

Customer Location:
Alexandra Dias-Lalcaca
Imperial College London
Molecular Sciences Resear
ch 82 Wood Lane
London, W12 7SL
United Kingdom

Order Details

1. Engineered Nanopores for Bioanalytical Applications

Billing Status: Open

Order License ID	1213808-1	**Type of use**	Republish in a thesis/dissertation
Order detail status	Completed	**Publisher**	Elsevier
ISBN-13	9781437734737	**Portion**	Chart/graph/table/figure

0.00 USD
Republication Permission

LICENSED CONTENT

Publication Title	Engineered Nanopores for Bioanalytical Applications	**Country**	United States of America
Author/Editor	Edel, Joshua B.,Albrecht, Tim	**Rightsholder**	Elsevier Science & Technology Journals
Date	04/12/2013	**Publication Type**	Book
Language	English		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	**Distribution**	Worldwide
Number of charts / graphs / tables / figures requested	1	**Translation**	Original language of publication
Format (select all that apply)	Print,Electronic	**Copies for the disabled?**	No
Who will republish the content?	Academic institution	**Minor editing privileges?**	No
Duration of Use	Life of current edition	**Incidental promotional use?**	No
		Currency	USD

Lifetime Unit Quantity	Up to 499
Rights Requested	Main product

NEW WORK DETAILS

Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System	**Institution name**	Imperial College London
		Expected presentation date	2022-06-21
Instructor name	Joshua Edel		

ADDITIONAL DETAILS

The requesting person / organization to appear on the license	Imperial College London

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure 3.7	**Title of the article/chapter the portion is from**	CHAPTER 3 Instrumentation for Low-Noise High-Bandwidth
Editor of portion(s)	N/A		
Volume of serial or monograph	N/A	**Author of portion(s)**	Edel, Joshua B.; Albrecht, Tim
Page or page range of portion	80	**Publication date of portion**	2013-04-12

Elsevier Science & Technology Journals Terms and Conditions

Elsevier publishes Open Access articles in both its Open Access journals and via its Open Access articles option in subscription journals, for which an author selects a user license permitting certain types of reuse without permission. Before proceeding please check if the article is Open Access on http://www.sciencedirect.com and refer to the user license for the individual article. Any reuse not included in the user license terms will require permission. You must always fully and appropriately credit the author and source. If any part of the material to be used (for example, figures) has appeared in the Elsevier publication for which you are seeking permission, with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder. Please contact permissions@elsevier.com with any queries.

Total Items: 1

Subtotal: 0.00 USD

Order Total: 0.00 USD

Approval 5

RightsLink®

My Orders My Library My Profile Welcome ad4216@ic.ac.uk Log out | Help | FAQ

My Orders > Orders > All Orders

License Details

This Agreement between Imperial College London -- Alexandra Dias-Lalcaca ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Print Copy

License Number	5293000314956
License date	Apr 20, 2022
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature
Licensed Content Title	Ion-beam sculpting at nanometre length scales
Licensed Content Author	Jiali Li et al
Licensed Content Date	Jul 12, 2001
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System
Institution name	Imperial College London
Expected presentation date	Jun 2022
Portions	1,2
Requestor Location	Imperial College London Molecular Sciences Research 82 Wood Lane London, W12 7SL United Kingdom Attn: Imperial College London
Total	**0.00 USD**

BACK

Approval 6

RightsLink®

My Orders My Library My Profile Welcome ad4216@ic.ac.uk Log out | Help | FAQ

My Orders > Orders > All Orders

License Details

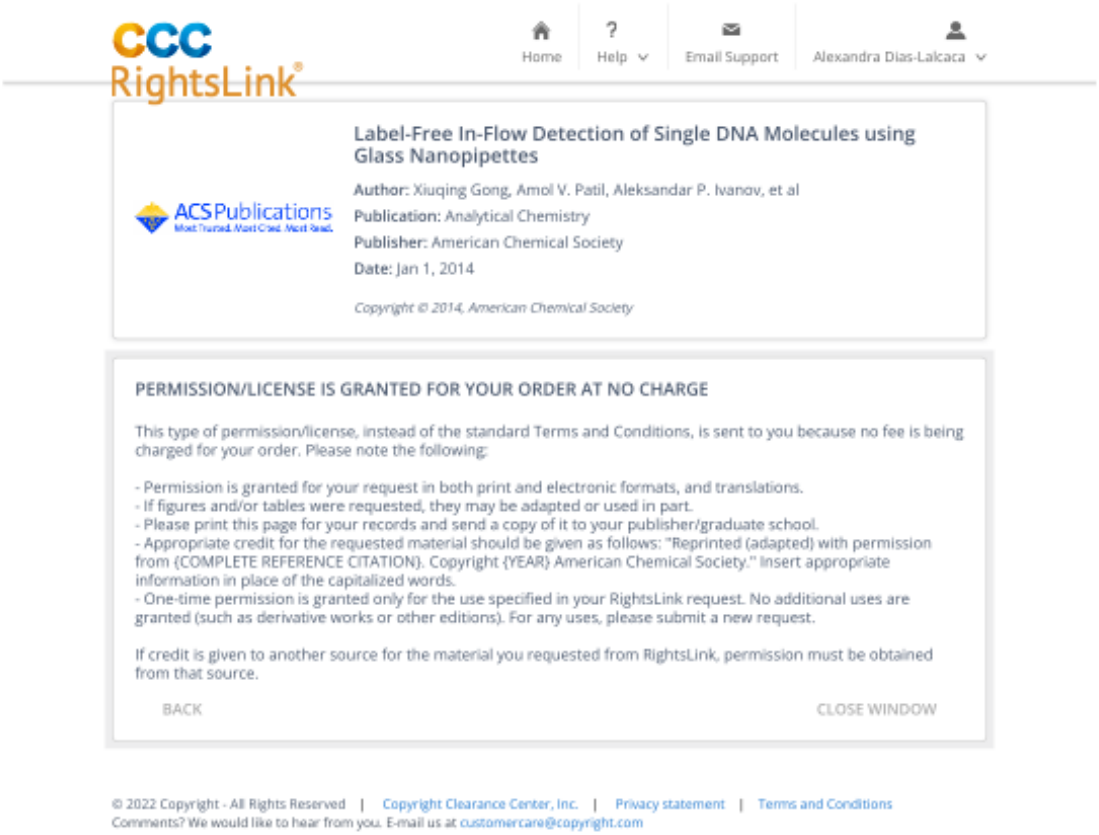
This Agreement between Imperial College London -- Alexandra Dias-Lalcaca ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Print Copy

License Number	5293000444454
License date	Apr 20, 2022
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Materials
Licensed Content Title	Fabrication of solid-state nanopores with single-nanometre precision
Licensed Content Author	A. J. Storm et al
Licensed Content Date	Jul 13, 2003
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System
Institution name	Imperial College London
Expected presentation date	Jun 2022
Portions	1
Requestor Location	Imperial College London Molecular Sciences Research 82 Wood Lane London, W12 7SL United Kingdom Attn: Imperial College London
Total	**0.00 USD**

BACK

Approval 7


CCC RightsLink
Home
Help
Email Support
Alexandra Dias-Lalcaca
ACS Publications
Most Trusted. Most Cited. Most Read.
Label-Free In-Flow Detection of Single DNA Molecules using Glass Nanopipettes
Author: Xiuqing Gong, Amol V. Patil, Aleksandar P. Ivanov, et al
Publication: Analytical Chemistry
Publisher: American Chemical Society
Date: Jan 1, 2014
Copyright © 2014, American Chemical Society
PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE
This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:
- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION). Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.
If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.
BACK
CLOSE WINDOW
© 2022 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Terms and Conditions
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Approval 8

RightsLink®

My Orders My Library My Profile Welcome ad4216@ic.ac.uk Log out | Help | FAQ

My Orders > Orders > All Orders

License Details

This Agreement between Imperial College London -- Alexandra Dias-Lalcaca ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Print Copy

License Number	5293001134964
License date	Apr 20, 2022
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Nanotechnology
Licensed Content Title	Direct observation of DNA knots using a solid-state nanopore
Licensed Content Author	Calin Plesa et al
Licensed Content Date	Aug 15, 2016
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System
Institution name	Imperial College London
Expected presentation date	Jun 2022
Portions	Figure 2
Requestor Location	Imperial College London Molecular Sciences Research 82 Wood Lane London, W12 7SL United Kingdom Attn: Imperial College London
Total	**0.00 USD**

BACK

Approval 9

Home | Help | Email Support | Alexandra Dias-Lalcaca

Intelligent Quantification of Picomolar Protein Concentration in Serum by Functionalized Nanopores

Author: Naren Das
Publication: IEEE Sensors Journal
Publisher: IEEE
Date: 15 Dec.15, 2018

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK CLOSE WINDOW

Approval 10

RightsLink®

My Orders My Library My Profile Welcome ad4216@ic.ac.uk Log out | Help | FAQ

My Orders > Orders > All Orders

License Details

This Agreement between Imperial College London -- Alexandra Dias-Lalcaca ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Print Copy

License Number	5293001416608
License date	Apr 20, 2022
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Angewandte Chemie
Licensed Content Title	DNA Origami Gatekeepers for Solid-State Nanopores
Licensed Content Author	Hendrik Dietz, Ulrich Rant, Thomas G. Martin, et al
Licensed Content Date	Apr 4, 2012
Licensed Content Volume	124
Licensed Content Issue	20
Licensed Content Pages	4
Type of Use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No
Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System
Institution name	Imperial College London
Expected presentation date	Jun 2022
Portions	Figure 2
Requestor Location	Imperial College London Molecular Sciences Research 82 Wood Lane London, W12 7SL United Kingdom Attn: Imperial College London
Publisher Tax ID	EU826007151
Total	**0.00 USD**

BACK

Approval 11

RightsLink

My Orders My Library My Profile Welcome ad4216@ic.ac.uk Log out | Help | FAQ

My Orders > Orders > All Orders

License Details

This Agreement between Imperial College London -- Alexandra Dias-Lalcaca ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Print Copy

License Number	5293010594518
License date	Apr 20, 2022
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Methods
Licensed Content Title	A monovalent streptavidin with a single femtomolar biotin binding site
Licensed Content Author	Mark Howarth et al
Licensed Content Date	Mar 22, 2006
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Novel DNA Carrier Structures for Protein Detection and Analysis in a Nanopore Sensing System
Institution name	Imperial College London
Expected presentation date	Jun 2022
Portions	Figure 1
Requestor Location	Imperial College London Molecular Sciences Research 82 Wood Lane London, W12 7SL United Kingdom Attn: Imperial College London
Total	**0.00 USD**

BACK