UNIVERSITY OF BIRMINGHAM

Sterically Controlled Nuclease Enhanced DNA Assembly in Rapid Sepsis Diagnostics

By

Oliver James Irving (BSc MSc MRes)

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Chemistry

College of Engineering and Physical Sciences

University of Birmingham

United Kingdom

June 2022

Abstract

Sepsis is a life-threatening condition afflicting 250,000 patients in the UK alone and defined as the overstimulation of the immune system in response to a pathogen. As the condition progresses, symptoms worsen, from fever and respiration difficulties to organ failure and cardiovascular issues. The primary test for sepsis is blood cultures which can take up to 2 days to complete and can often provide a false negative result. There are currently no point of care devices available for sepsis diagnosis. While many biomarkers have been identified for sepsis, few have been incorporated into the clinic. Rapid and accurate diagnostics is the key to reducing patient mortality and improve prognosis. Nanopore sensing is a technique based on the principle of resistive pulse sensing, first described by Wallace H. Coulter. Nanopore sensing is a very sensitive technique often applied to single molecule sensing. Recent developments in biosensing demonstrate the potential for DNA modification for biomolecule capture and nanopore sensing. These techniques are often limited in their adaptability to capture multiple analytes.

This thesis discusses a newly developed technique for rapidly producing a diagnostic toolbox and testing multiple analytes simultaneously with nanopore sensing integration. It is possible to produce a DNA nanostructure containing multiple analyte binding sites that can bind IL-6 and procalcitonin proteins, detected using modified ELISA and resistive pulse sensing techniques. For a trimer structure, it is possible to capture proteins with an efficiency of 72.5%. From multiple assembly experiments it has been shown that structures of different sizes containing a variety of probes can be assembled. Most notably, this thesis shows that the limitations of the Gibson assembly method can be overcome using the developed technique. The inclusion of a biotin probe allowed for both DNA structure isolation and post assembly functionalisation.

Optimisation of nanopipette fabrication has allowed for the development of a reproducible protocol to produce nanopipettes with pore sizes at <10 nm, 10 - 30 nm, and finally <30 nm. To increase the sensing capability of the nanopipette, taper length was also optimised to consistently produce pipettes with taper lengths of ~3000 µm. Further optimisation of extrinsic electronic noise was performed to improve the signal to noise ratio for translocation experiments. It was determined that the primary contributor to noise was using equipment supplied by mains power. Therefore, optimisation explored altering the equipment to work off battery power alone. From these results, the experimental procedure used would involve using shorter, anodised electrodes, using a silver shielded data transfer cable, working on a granite slab, using a passive filter of 100 kHz or less, and working in an “off grid” approach using battery power where possible.

Standard DNA fragment translocation was performed, and the results compared against literature values. The results highlight similar translocation frequencies of ~1 per nM per second, event duration, and magnitude also. Through comparison between tetramer SCoNE DNA and bare DNA fragments, it was possible to identify significant subevents relating to probe structures at positions near 0, 0.25, 0.5, and 0.75 along the DNA backbone. Whilst it was not possible to confirm this difference for a decamer structure, subevents were observed at locations specific to probe sites. The difference was not significant between the decamer SCoNE structure and the 10 kbp fragment. The decamer structure was translocated prior to the development of the biotin probe. It is therefore possible to suggest that the gel extraction method used for isolation could have cleaved the probe structures from the backbone, limiting the potential for subevent detection.

Declaration of Originality

I hereby declare that the work presented in this thesis is my own, except where referenced. This work has not previously been submitted in any form to satisfy any degree requirement at this or any other university.

Declaration of Patents attributed to this thesis

Device - International application No. - PCT/GB2021/052870, filing date - 4 November 2021

Analyte Capture system - UK Patent Application No. 2104219.7, filing date - 25 March 2021

Copyright Declaration

Contents

Acknowledgements

There are many people that I would like to thank on both a professional and personal level for their contributions to the completion of this work.

Firstly, I would like to thank my supervisor Professor Tim Albrecht for your guidance and encouragement during this project. Without you, I would have never had the opportunity to explore the ideas that I had, and SCoNE would've only remained an idea on a piece of paper at the back of my notebook. I would not be the scientist I am without your help.

I would like to gratefully acknowledge the financial support of the Engineering and Physical Science Research Council (EPSRC).

I would like to thank the whole Albrecht group, past and present for their support during this work, whether helping me learn new skills, giving me ideas on how to progress with the work, or just listening to me complain after failed experiments. For assisting me in learning and developing my code writing skills I would like to thank Dr Anton Vladyaka and Christopher Weaver. Both of you patiently taught me the basics of code writing and helped me develop much of the analysis code used throughout this project. For encouragement, moral support, and reminding me "It can't be that easy" when developing the ideas behind SCoNE, I would like to thank Dr Joseph Hamill. Your support during the early design and ideas phase encouraged me to improve many of the assembly self-check steps, which have ultimately led to a functional protocol. To Nashwa Awais, I would like to thank you for always being there to pick me up or play a game after a particularly hard day of experimentation. To Dr Adrian Fortuin, I would like to thank you for helping me develop the methodology for electrode preparation, and for always being there whenever I needed you. In particular I would like to thank Lauren Mathews, who has been a great help during all the development stages of

SCoNE, translocation experiments, and us battling the noise together. It has been a great honour and privilege to work with and befriend all of you.

Additional thanks to the Neely team, for their assistance and cooperation in using their modified SAM molecules. In particular I would like to thank my second supervisor Dr Rob Neely for giving me the idea of modifying Gibson assembly to create SCoNE. I would also like to thank Steven Coulthard for providing me with the modified SAM molecule for this research.

To my friends Jack Healy and Lily Peace, I would like to thank you for always being there for me.

To my girlfriend Kristen Hawkins, I would like to thank you for all of your support, encouragement, and ideas. For being my cheerleader, always being there and helping me when the going was tough, I will be eternally grateful.

Finally, I would like to thank my parents, Karen and Andrew Irving, for everything that you do. Without you, this work would have not been possible, and so I dedicate this work to the both of you.

Heb bob un ohonoch, fyddwn I byth lle rydw I heddiw. Felly, I bob un ohonoch, diolchaf ichi o waelod fy nghalon.

Indexing

List of abbreviations

Ordered alphabetically.

Full title	Abbreviation
5-ethynyl-2'-deoxyuridine	EdU
Atomic force microscopy	AFM
avidin-biotin-peroxidase	ABC
C-reactive protein	CRP
Dibenzocyclooctyne	DBCO
Dimethyl sulfoxide	DMSO
DNA anti-gravity translocation apparatus	DATA
DNA methyl transferases	Dnmt
DNA polymerase	DNApol
Double stranded	ds
Enzyme linked immunosorbent assay	ELISA
Fas ligand	FasL
Fourier transform	FT
Interleukin	IL
Lithium chloride	LiCl
Limit of detection	LoD
Major histocompatibility complex	MHC
Methyl transferase enzyme from the extremophile species *Thermus aquaticus*	M.TaqI
Micro-RNA	miRNA
Monocyte chemoattractant protein-1/ Chemokine ligand 2	MCP1/CCL2
Neopterin	NPT
pDNA containing a biotin isolation group	iDNA
phosphate buffer solution	PBS
Point of care	PoC

Polymerase chain reaction	PCR
Power spectral density	PSD
Quantitative PCR	qPCR
Receiver operating characteristic	ROC
Root means squared	RMS
S-adenosyl methionine	SAM
S-Adenosyl-L-homocysteine	SAH
short probe strand DNA	pDNA
Silver/ silver chloride	Ag/ AgCl
Single stranded	ss
spacer strand DNA	sDNA
Sterically controlled nuclease enhanced	SCoNE
Systematic Evolution of Ligands by Exponential enrichment	SELEX
Systemic inflammatory response syndrome	SIRS
Thermus aquaticus	Taq
Tris-acetate-EDTA	TAE
Tris-HCl EDTA	TE
Tumour necrosis factor alpha	TNF-α
User interface	UI

List of figures

List of tables

List of equations

Equation number	Equation	Description
3.1	$Ag_{(s)} + Cl^{-}_{(aq)} \leftrightharpoons AgCl_{(s)} + e^{-}$	Redox reaction showing the formation of silver chloride from a silver ion and a chloride ion producing an electron.
3.2	$d_{pore} = \frac{4Gl + \frac{\pi}{2}GD_i}{D_i \pi g(c) - \frac{\pi}{2}G}$	Pore size calculation where G is the conductance of the nanopipette, l the taper length of the nanopipette (as measured using callipers and optical microscopy), D_i is the inner diameter of the capillary (0.5 mm) and g(c) the conductivity of the electrolyte (4 M LiCl + TE determined to be 173 mS cm^{-1}).
4.1.A	$x_{RMS} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \cdots x_n^2)}$	The equation for calculating RMS noise where “x” is the mean noise value, and “n” represents the number of samples or data points to input into the equation.
4.1.B	$I_{RMS} = \sqrt{\Delta I^2}(t)$	The simplified RMS calculation for current modulation where I_{RMS} represents the RMS noise value, $\Delta I^2(t)$ represents current modulation deviating from the mean current value.
5.1	$[DNA(\mu M)] = \frac{\mu g/\mu l}{Length(bp) \times 660 g/mol}$	The calculation for determining DNA molarity.

Chapter 1

Introduction to sepsis and diagnostics

Synopsis: This chapter introduces the aims and objectives of this research project, an introduction to sepsis, a brief introduction into DNA amplification and uses in diagnostics, biomarkers and capture methods, and an overview of the field of diagnostics with relation to the work undertaken. This chapter introduces many of the concepts which are later used or adapted during experimental work.

1.1 Sepsis

Sepsis, also referred to as septicaemia or blood poisoning, is a life-threatening condition afflicting approximately 250,000 patients a year in the UK (The UK Sepsis Trust, 2019). The condition involves the overstimulation of the immune system to the presence of microorganisms, primarily bacteria, attributed to 70% of cases (Centers for Disease Control and Prevention, 2018; Lin *et al.*, 2018). There are several stages of sepsis; mild sepsis, severe sepsis and septic shock with associated mortality rates of 30, 50 and 80%, respectively (Dellinger *et al.*, 2013). Advancement of the condition leads to increased risk of vital organ failure, directly due to the effect of the immune system (Parlato and Cavaillon, 2014).

Current medical tools for diagnosis involve positive results in two of the systemic inflammatory response syndrome (SIRS) criteria as illustrated in Figure 1.1 (Deutschman and Tracey, 2014). The use of microorganism culturing has also been used for identification, however, blood cultures or swab cultures can take several days to grow and will not always produce a positive result (Kumar *et al.*, 2001; Lin *et al.*, 2018).

It has been found that rapid identification of sepsis drastically improves prognosis. As sepsis progresses, the amount of fluid entering muscular tissue (oedema) increases, causing a decrease in blood pressure. Therefore treatment includes a combination of antibiotics and saline (Taeb *et al.*, 2017).

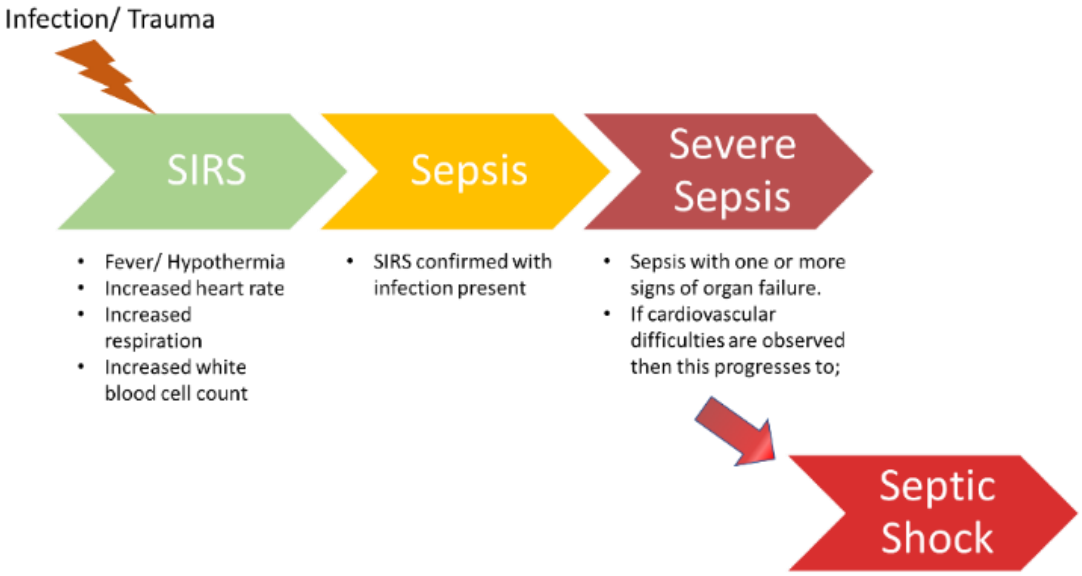

Figure 1.1. Guide to diagnosis for the progression of sepsis. Image adapted from Daniele Ramazzotti (Ramazzotti, 2014).

1.2 DNA amplification and utilisation

DNA has been used for a variety of purposes from identifying criminals, to determining the paternity of a child, and genetic disease diagnosis (Erlich, 2015). DNA is often isolated at low concentrations, reducing the usability of collected samples (Hansson *et al.*, 2009; Raymond *et al.*, 2009). To overcome this DNA samples are amplified using polymerase chain reaction (PCR), illustrated in Figure 1.2. PCR was invented in 1983 by Kary Mullis, and its first successful use highlighted in 1985, when investigating the genomic sequences of beta-globin for the diagnosis of sickle cell anaemia (Saiki *et al.*, 1985). The technique rapidly amplifies a DNA fragment using a combination of temperature control, enzymes, primers (short single stranded [ss] DNA fragments), and free nucleotides. PCR is an *in vitro* model of what occurs during cell replication, however rather than supporting the synthesis using proteins, PCR uses thermal cycling (Saiki *et al.*, 1985; Erlich, 2015). The enzyme used for PCR is a DNA polymerase (DNApol) typically from the thermophile *Thermus aquaticus* (Taq) however DNApol from other organisms have also been used (Brock and Freeze, 1969; Jia *et al.*, 2014).

Taq enzymes are used in a variety of biochemical techniques due to their heat resistance properties. Initially the hydrogen bonds between DNA strands are broken using high heat (95˚C), forming ssDNA. The heat is then lowered quickly to approximately 55˚C (varied dependant on the binding temperature) to allow the primers to bind to the ssDNA. This acts as a template to allow the DNA polymerase to bind. The temperature is then raised into the enzymes most active range (72°C for Taq DNApol). The Taq DNApol binds to the ssDNA-primer region and begins to move along the strand, incorporating base-matched nucleotides into the elongating strand. This forms a stable double strand through phosphodiester bond formation between the elongating strand and the newly incorporated nucleotide. This re-forms the double stranded (ds) DNA structure. The Taq DNApol then reaches the end of the DNA strand, finishing the final nucleotide conjugation before releasing the DNA strand and moving onto another ssDNA strand. This process is typically cycled 30 times, rapidly amplifying the starting DNA (Erlich, 2015).

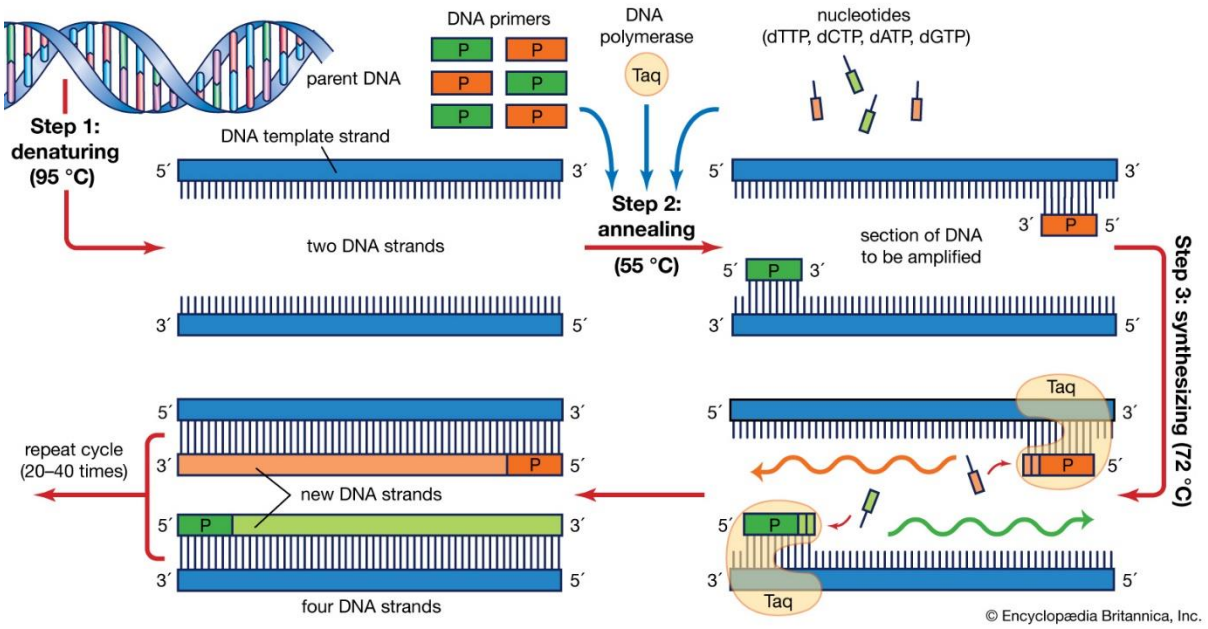

Figure 1.2. Diagram illustrating the PCR process. Image taken from Britannica (Britannica, 2020).

Most recently, PCR has been employed in the diagnosis of COVID-19 (van Kasteren *et al.*, 2020). Most viruses use RNA as their genetic material, using other organism's cellular machinery to produce more of the viral proteins and genetic material. Therefore, to use PCR for diagnosis, additional steps are added to the process. A reverse transcriptase enzyme is added to the reaction mixture which converts the RNA into ssDNA. An intercalating dye is also incorporated to perform quantitative PCR (qPCR). This allows for the real time detection of newly formed dsDNA (Stein *et al.*, 2020). With regards to COVID-19 PCR tests, samples are loaded along with the qPCR mix and a positive result determined if the DNA concentration reaches a certain threshold by a certain number of cycles (typically 24, however this is dependent on the test kit used) (van Kasteren *et al.*, 2020).

1.3 Biomarkers

Biomarkers are defined as biologically active molecules by which a pathological or physiological process can be identified (Strimbu and Tavel, 2010; Hirsch and Watkins, 2020). Biomarkers are classed in two distinct ways (Strimbu and Tavel, 2010). Firstly, they are classified on the type of information that they provide. These are; diagnostic, prognostic, pharmacological, and surrogate biomarkers (Haschek *et al.*, 2013; FDA-NIH Biomarker Working Group, 2016; Khan, 2016; Faqi, 2017). The focus in primary research is to discover and utilise diagnostic and prognostic biomarkers (Srinivas *et al.*, 2001; Vliet *et al.*, 2017). A diagnostic biomarker provides an identifier, with a particular sensitivity and specificity, to determine whether a patient has the disease under investigation (Hsu *et al.*, 2013; Gardner *et al.*, 2020). A prognostic biomarker provides information on the likelihood of mortality, and/ or resistance to a particular type of treatment (Zinellu *et al.*, 2021). Secondly, biomarkers are classified into

physiologic characteristics, radiographic, histologic, and molecular (U.S. Food & Drug Administration, 2019).

Physiologic characteristic biomarkers are defined as a physiological change occurring from the presence of a condition or disease. These include physiological changes such as blood pressure and oxygen saturation. The use of this type of biomarker is well established and routinely performed both diagnostically and prognostically. Many conditions alter these basic physiological features due to the stress induced. Whilst these are useful for general understanding of a patients' condition, they lack diagnostic specificity, and therefore further biomarkers are required for an accurate diagnosis (Medicine, 2004).

Radiographic biomarkers are classed as biomarkers or biological features which can be observed using an imaging technique. These include examples such as lesions or tumour imaging using techniques like X-ray or MRI. Imaging techniques such as these are used in both diagnosis and prognosis of various conditions such as tumour growth or multiple sclerosis lesion monitoring. These techniques contribute to a more in-depth diagnostic and prognostic understanding of various conditions; however, these rely on significant changes to physiology to image. Small changes in physiology are often missed by imaging techniques, and more invasive procedures are employed to confirm diagnosis (Smith *et al.*, 2003).

Histologic biomarkers are defined as an alteration in normal cellular behaviour resulting in a change in cell structure and behaviour. This type of biomarker is typically used in the grading and staging of cancer from biopsy samples. This very invasive technique is employed often only once imaging studies have been performed, and a diagnosis

derived (Kilty *et al.*, 2007). These types of biomarker are primarily used to confirm a diagnosis and provide a more accurate prognosis (Gazquez *et al.*, 2012).

Molecular biomarkers are defined as non-imaging based biomarkers which have bio-physical properties and are further divided based on the species of molecule they belong to (Laterza *et al.*, 2007). This type of biomarker is extracted from bio-fluids such as; blood, urine, cerebral spinal fluid, saliva, mucus, semen and vaginal secretions (Broza *et al.*, 2019). Molecular biomarkers include, but are not limited to, proteomic, lipidomic, metabolomic, and genomic (Nordström and Lewensohn, 2009; Sinclair and Dudley, 2019). This is both classed as an invasive, and non-invasive technique, dependant on the type of bio-fluid used, and the condition of the patient at the time of extraction. This diagnostic technique is employed routinely in blood tests for molecules such as glucose, creatinine, and ion concentration. This type of biomarker analysis provides both diagnostic and prognostic information and is often used to monitor patient condition (NHS, 2020). Whilst molecular biomarkers are often one of the most studied for diagnostic and prognostic purposes, they are often the least utilised (Selleck *et al.*, 2017).

1.3.1 Proteomic Biomarkers in Sepsis

There are many proteomic biomarkers for sepsis, however the most significant interest for study includes interleukins (IL) 6, 8, and 10 proteins, C reactive protein (CRP), and procalcitonin (Parlato and Cavaillon, 2014). IL proteins are responsible for the modulation of the immune system, either through promotion or suppression (Akdis et al., 2011). IL-6 is a pro-inflammatory mediator released from macrophages in the tissue in response to an infection and to induce oedema. IL-6 released from these cells also stimulates the activation of B cells (Aziz et al., 2013). Clinically, normal blood IL-6

concentration should be <7 pg/mL. In SIRS patients, IL-6 in the blood has a median concentration of 62 pg/mL, 131 pg/mL for sepsis, 346 pg/mL for severe sepsis, and 659 pg/mL for septic shock (NHS, 2021a). During infection, CRP is synthesised by hepatocytes, stimulated by IL-6 (Póvoa et al., 1998). The protein binds to the surface of cells and activates the complement system (Thompson et al., 1999). This often activates T-cells and stimulates the adaptive immune response, further potentially initiating an autoimmune reaction, leading to increased organ damage (Pepys and Hirschfield, 2003). It has been found that increased blood CRP levels in a cohort of 306 patients had a specificity and sensitivity, for sepsis, of 75% and 98.5% respectively (Póvoa et al., 1998). Clinical reference values state that CRP should have a concentration of ~0-5 mg/L in the blood. Elevated levels above 5 mg/L are considered to reflect a worsening condition (NHS, 2021b). IL-8 is a pro-inflammatory mediator responsible for the recruitment and activation of other immune cells at the site of infection (Modi et al., 1990). Increased release of IL-8 has been strongly associated with the progression of sepsis from severe to septic shock in a clinical study of 47 patients, however it is not currently used in routine medical tests (Berner et al., 1998; Livaditi et al., 2006). Previous research has suggested that the concentration of IL-8 in healthy volunteers is <10 pg/mL, and levels above this are considered to be clinically significant (Hack *et al.*, 1992; Yoshio Hirao *et al.*, 2000). IL-10 is an anti-inflammatory mediator produced by macrophages, T and B cells (Park et al., 2018). Increased synthesis of IL-10 has been linked (along with IL-33 synthesis) to a decreased mortality rate in mouse models with bacterial sepsis (Nascimento et al., 2017). The proposed hypothesis is that increased IL-10 synthesis is vital for preventing severe sepsis progressing to septic shock (Steinhauser et al., 1999). However, if the protein becomes

too abundant in the blood, the immune system can be supressed leading to immunoparalysis (Moore *et al.*, 2001; Abe *et al.*, 2008). IL-10 is not used in clinical blood tests, however previous research suggests that healthy adults have a mean blood concentration of 5.7 pg/mL, concentrations above 10 pg/mL are clinically relevant, and that concentrations above 23 ng/mL were linked to increased mortality (Alsaif *et al.*, 2012; Li *et al.*, 2017; Dunne *et al.*, 2017; Gao *et al.*, 2017). Procalcitonin is another protein released during microbial infection and is the precursor of calcitonin, an endocrine calcium concentration modulator. During an inflammatory response to bacterial infection, synthesis of procalcitonin in adipocytes is greatly increased and secreted in the blood, however, the protein is not cleaved to its active form. The specificity of procalcitonin allows for a more accurate diagnosis. It has also been shown that an increased concentration of procalcitonin in the blood directly links to the antibiotic concentration required for treatment, reducing hospital stay time (Jin and Khan, 2010). Clinically, circulating PCT should have a concentration below 70 pg/ml. Concentrations of <0.5 ng/mL are considered to be caused by a local infection, between 0.5 – 2 ng/mL, sepsis is considered likely but further tests are required to confirm, between 2 – 10 ng/mL sepsis becomes the most likely candidate, and above 10 ng/mL is considered to be a severe inflammatory response to infection (NHS, 2021c).

1.3.2 Genetic Biomarkers in Sepsis

As well as protein biomarkers, there are two microRNA (miRNA) markers that have been associated with sepsis. miRNA-146a is a blood circulating miRNA which is responsible for the modulation of the immune system, and regulation of inflammation (Sonkoly et al., 2008). In comparison with other inflammatory diseases, it was found

that miRNA-146a was significantly reduced in septic patients with an accuracy of 85.8% as determined by area under a receiver operating characteristic (ROC) curve (Wang et al., 2010; Wang et al., 2013). miRNA-150 has also been identified as a prognostic biomarker for sepsis (Roderburg et al., 2013). miRNA-150 is also an important regulator of the inflammatory pathway; although in recent research it was shown that miRNA-150 was a poor biomarker for diagnosis, decreased circulation concentration demonstrated a high correlation with poor prognosis (Benz et al., 2016). There are many other potential RNA biomarkers for sepsis including long non-coding RNAs and circular RNA. Whilst the potential for long non-coding RNAs has not been fully explored, there is increasing interest in utilising these as potential biomarkers and in therapeutics (Zhang et al., 2017).

There are several genes within immune cells which offer diagnostic evidence for sepsis. These primarily represent immune cell recruitment proteins, and inflammatory response proteins in the presence of an infection. IL-1β, IL-6, IL-8, tumour necrosis factor alpha (TNF-α), Fas ligand (FasL) and monocyte chemoattractant protein-1 (MCP-1 also known as CCL2) mRNAs are all upregulated by the immune system during a systemic infection (Singer, 2013). During infection, cellular death increases causing cells to release DNA and various types of RNA into the blood stream (Hotchkiss et al., 1999). This is assisted by an increase in transcription and translation of FasL and MCP genes, responsible for producing the TNF-α receptor and a monocyte chemoattractant protein respectively (Kobayashi et al., 2006; Chung et al., 2017). Due to the imprecise nature of sepsis, the pro-inflammatory genes are transcribed at vastly higher rates whilst cell death is widespread. Therefore, the

comparative concentration of both the mRNA and the associated proteins is much higher in the blood compared to patients without sepsis (Singer, 2013).

1.4 Biomarker capture

There are several different methods by which biomarkers are captured from bio-fluids in both research and medicine. The capture methods vary in their ability to handle complex samples. As bio-fluids are often complex mixtures, several methods to study biomarkers require varying degrees of isolation and processing (Hottenstein *et al.*, 2017). In medical tests, for example blood sampling, isolation of different blood components is required for further analysis. This involves separating plasma (small molecules), red blood cells (erythrocytes), and white blood cells (leukocytes). This is done using centrifugation, and each fraction can then be further processed for molecule or cell type isolation (NHS, 2020). In research there are a variety of different isolation techniques which are utilised dependant on the pathogenesis under investigation (Hottenstein *et al.*, 2017).

With small molecules (metabolomics), mass spectrometry is often utilised to study the presence and concentration of a biomarker (Kuhn *et al.*, 2018). This uses a separation technique such as liquid or gas chromatography to separate out complex mixtures in samples for analysis based on their physical properties (Sinclair and Dudley, 2019). The drawbacks to this technique are the preparation of the sample, time required to run samples, and difficulty in interpreting results (Matsuda, 2016). For a direct analysis of single compounds of interest, more targeted methods have been developed to study singular biomarkers with less sample preparation. These techniques use biochemical principles to isolate and measure biomarkers (Solier and Langen, 2014; Gilboa *et al.*, 2020).

1.4.1 Methods for Biomarker Capture

1.4.1.1 Antibodies

Antibodies are a collection of proteins which are produced by B-cells in the blood in response to an assumed threat to the immune system (Litman *et al.*, 1993; Janeway *et al.*, 2005). These include foreign proteins, peptides or cell membrane fragments presented to the B-cell by T-cells through a major histocompatibility complex (MHC) class 2 receptor. These foreign fragments induce the B-cell to produce antibodies specific to the fragment presented. The antibodies then assist immune cells to fight the threat (Liszewski *et al.*, 1996; Medzhitov and Janeway, 1997). Antibodies have a mass of 150 kDa (approximately 10 nm in size) and are made up of four polypeptide chains, as highlighted in Figure 1.3 (Reth, 2013). Two of these are identical heavy chains, and two light chains, connected via disulphide bonds. Each end of each chain has a variable region which varies to bind the antigen presented to the B-cell. These are the antigen binding regions (Woof and Burton, 2004; Kotiw *et al.*, 2012; Solier and Langen, 2014). In biomarker capture research, this variable region has been used to diagnostic advantage. Highly purified antigens are injected into laboratory animals, or presented to cancerous B-cell line hybrids, to initiate an intense immune response of antigen-specific antibodies, which can then be separated from the blood or growth medium. This allows for the production of a high concentration of antibodies specific to a single protein (Hanly *et al.*, 1995; Dangi *et al.*,

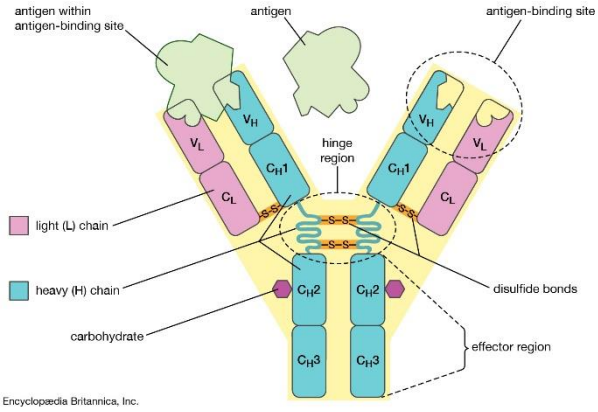

Figure 1.3. Antibody structure. Image taken from Encyclopedia Britannica (The Editors of Encyclopedia Britannica, 2020)

2018). These can then be used to study protein or peptide fragments from patient samples using techniques such as enzyme linked immunosorbent assay (ELISA) or immunohistochemistry (Koopmann *et al.*, 2004; Suganuma *et al.*, 2008).

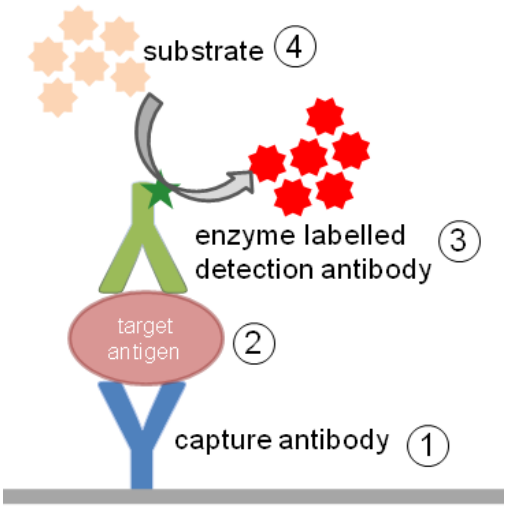

Figure 1.4. Diagram illustrating the ELISA technique. Image taken from the British society for immunology (Horlock, 2009).

ELISAs are carried out in an antibody absorbent 96 well plate. The most common ELISA performed is a sandwich ELISA as illustrated in Figure 1.4. Firstly, the primary antibody is conjugated to the plate surface. Secondly, the antigen under investigation, cell lysate, or patient sample, is added into the well and incubated to allow conjugation. Incubation often occurs at different temperatures dependant on the antibody used. This ranges from 4°C for an overnight reaction to 37°C for 1 – 2 hours. Often these reactions take place at room temperature. Thirdly, a secondary antibody, specific to the same antigen, or enzyme linked antibody, is added and incubated. If an unlinked antibody is used, an enzyme complex is then added and incubated. The enzyme complex is typically a horseradish peroxidase or an alkaline phosphatase. Finally, a dye, such as 3,3′,5,5′-Tetramethylbenzidine (TMB), is added into the wells. The dye is colourless but becomes coloured due to the action of the enzyme. The amount of colour change is directly proportional to the amount of target antigen present and allows for accurate concentration determination. Between each step of the assay, wash steps are performed to ensure no non-specific conjugations occur (Wild *et al.*, 2013; Jeong, 2014). Whilst not widely used for clinical diagnostics, Battaglia *et al.,* used a human

procalcitonin specific ELISA to detect the sepsis marker in equine and canine samples with a limit of detection of 56 and 11 ng/mL respectively (Battaglia *et al.*, 2020).

Variations on antibodies are the use of Fab fragments (antigen-binding fragment). These are the antigen binding regions of the antibody with the tail (Fc fragment) removed through papain digestion. These are much smaller than full antibodies and allow for better antigen localisation, or capture within a smaller space (Flanagan and Jones, 2004).

1.4.1.2 Affimers and aptamers

Other methods for antigen isolation include the use of affimers and aptamers. Both capture methods work on similar principles but use different monomer types for assembly. Affimers are small polypeptides (approximately 1.6 nm) which are designed to mimic the antigen binding region of an antibody, whilst being much smaller in size (12 - 14 kDa). Affimers are comprised of an alpha helix on top of an anti-parallel beta sheet with two peptide loops. It is the peptide loops which are altered to allow the binding of different proteins. Affimers are produced through plasmid bacterial expression and purified (Tiede *et al.*, 2017).

Aptamers are small (3 - 5 nm), single stranded polynucleotides which fold into a specific shape when heated to allow the binding of a single antigen. These are designed through a random sequence generator *in silico*, and modelling used to predict folding properties and antigen specificity (Mallikaratchy, 2017). Aptamers are produced similarly to oligonucleotide synthesis, using protection groups to add single nucleotides then the cleavage of this group to continue adding bases to the sequence (Sigma-Aldrich, 2018). In addition to specific sequences, modified nucleotides can be included

to ensure that the secondary and tertiary shapes fold correctly or add additional sites for modification. The sequences are then trialled *in vivo* against the predicted antigens. Systematic Evolution of Ligands by Exponential enrichment (SELEX) is a method by which aptamers are tested for activity and specificity. Initially a library of randomly sequenced DNA and RNA oligonucleotides are folded and added to surface bound analytes of interest. The aptamers which do not bind the analyte are washed from the surface, whilst the bound sequences remain. The bound sequences are then eluted from the analyte and amplified before repeating the procedure, typically between 5-15 times (Tuerk and Gold, 1990; Kruspe and Giangrande, 2017). Aptamers are extremely useful for biomarker capture as they can be designed for any biomarker of interest, produced rapidly, and are cost effective. However, there are few which are commercially available for purchase. Whilst antibodies and affimers are primarily used for protein and peptide capture, aptamers can be designed against a range of different biomarkers, including; metabolites, lipids, and proteins (Mallikaratchy, 2017).

1.5 Diagnostics

Diagnostics is the field of understanding, practicing, and developing diagnostic tests for conditions based on unique characteristics to a disease. Across all areas of diagnostics there is an increasing need to improve the diagnostic method due to the increased understanding of pathogenesis (Peeling *et al.*, 2019; Gomez-Marquez and Hamad-Schifferli, 2021). This also includes the improved understanding of disease resistance to treatment and the interaction between different systems within cells and the body as a whole (Korman, 2020). Within biomarker discovery research, new biomarkers, for a variety of diseases, are found every year. However, few of these are rapidly implemented into routine medical tests for a variety of reasons (Parlato and

Cavaillon, 2015). These often concern the available specialist equipment or time required to perform the assays. Developing a method for using one or several of these biomarkers in an “easy to use” point-of-care device could improve the time from bench to implementation (Stenman, 2016).

Recently 178 potential biomarkers for sepsis have been identified, however, few have been used clinically for diagnostics. It has been suggested that a combination of these biomarkers could be the most effective way to improve the accuracy of diagnosis and decrease treatment time (Parlato and Cavaillon, 2015).

1.5.1 Developments in Sepsis Diagnostics

Due to the rapid advancement of sepsis, it is vital that a diagnosis is rapidly confirmed to improve patient outcome. Investigation into rapid sepsis diagnosis has facilitated the development of an automated procalcitonin chemiluminescence immunoassay. Due to the specificity of procalcitonin as a systemic bacterial infection marker, the determination of early blood procalcitonin concentration is vital for monitoring the efficiency of antibiotic therapy. The LIAISON® BRAHMS PCT® II GEN test utilises two monoclonal antibodies which coat magnetic beads to detect the presence of procalcitonin and a reference molecule from patient plasma. The technique allows for rapid quantitation of procalcitonin, over 16 minutes, through chemiluminescence by interaction of the two substrates. This test has a limit of detection (LoD) of 0.02 ng/mL. Utilising this technique clinically could allow for early and reliable diagnosis as well as early antibiotic treatment (Hubl et al., 2003; Fortunato, 2016).

Recent work in diagnostic tests for sepsis has led to the development of a microelectrode to detect IL-6 in real-time. An array of eight 50 μm gold electrodes were

fabricated onto a needle-like substrate with immobilised IL-6 antibody on the electrode surface. By use of electrochemical impedance spectroscopy and differential pulse voltammetry it was possible to detect the protein at physiological levels (LoD of 25 pg/mL) in 2.5 minutes. It is believed that this technique could offer point of care (PoC) testing or be inserted into blood vessels for continued monitoring (Russell *et al.*, 2019).

It is well known that multi-biomarker detection increases the reliability of diagnosis. Research by Kemmler into the simultaneous detection of CRP, IL-6, procalcitonin neopterin (NPT) led to the development of a biochip for multi-analyte detection and measurement. The biochip uses 10 – 75 µL of human plasma, injected through a microfluidic channel, to perform an immunofluorescence assay using a sandwich format for IL-6 and PCT, and binding inhibition for NPT and CRP. The fluorophores are excited at 638 nm and fluorescence read at 670 nm using an 8-bit fire wire CCD camera. Using this method, it was possible to achieve LODs for IL-6 of 0.27 ng/mL, PCT at 0.34 ng/mL, CRP at 900 ng/mL, and NPT at 1 ng/mL. This takes 25 minutes for a single assay. Whilst the sensitivity is not high enough yet to detect clinically relevant levels of the proteins, and blood plasma requires isolation, the results are encouraging towards the development of a multi-biomarker detection system (Kemmler *et al.*, 2014).

1.6 Aims and Objectives

In this research there were several ideals that were aimed to be accomplished. The primary aim was to develop a new method for constructing a DNA based multi-biomarker capture system, compatible with nanopipette DNA translocation for rapid diagnostics. Through this aim, the development of a purification technique to isolate only the biomarker capture system would be necessary. To achieve these aims, firstly,

the development of a robust method for consistent nanopipette production with a pore size of 10 nm for high resolution data acquisition was necessary. Secondly, characterisation of unmodified DNA translocation would be required for comparison against modified constructs. Finally, characterisation of both bound and unbound modified constructs would be undertaken. This would allow for the ability to distinguish these structures and determine if the method could be used for rapid sepsis diagnosis.

1.7 Thesis outline

Chapter 2 will introduce DNA modification and assembly techniques. This chapter reports the successful production of new DNA structures with functional biomolecule probes. The adaptation of existing techniques to confirm the biomolecule capture properties are also discussed. It has been shown that it was possible to create a range of structures of different sizes which could specifically capture proteins of interest through gel electrophoresis and modified ELISA techniques.

Chapter 3 introduces the concepts behind nanopore sensing and discusses the instrumentation used to perform these types of experiments. The exploration of different protocols for the production of nanopipettes and the advantages and limits of those pipettes produced is also discussed. It has been shown that it is possible to create nanopipette diameters ranging from 10 - 30 nm and above. This chapter provides conformation of experimental conditions for further research to be conducted.

Electronic noise and methods for its limitation are discussed in Chapter 4. With one of the aims of this research being the translocation and detection of the modified DNA structures, the methods explored to limit the noise during experiments are further discussed and analysed. It has been shown that it was possible to significantly reduce

both high and low frequency noise through shielding data transfer cables, electrode anodization and length reduction, use of a granite anti-vibration table, removing main-line power supply, and use of filters for incoming signal.

Chapter 5 introduces the expected electronic signature of linear, folded, and knotted DNA translocation. This chapter also discusses the results from DNA translocation experiments in comparison to literature values. Translocation results from created DNA structures are also discussed. Comparison of bare and modified DNA translocation experiment results allowed for the presence of probe structures conjugated to the DNA backbone to be confirmed.

Chapter 6 and 7 conclude the thesis by discussing the overall results and progress made through the research. The future developments and optimisation are then discussed in relation to biosensing and medical application. Whilst not contributing a significant amount to the overall thesis, Appendix 1 discusses several projects conducted alongside the main body of work: the development of a simulated translocation code is discussed in relation to the identification of subevent peaks; a proof-of-concept experiment for DNA isolation using electrophoretic separation concludes that it is possible to isolate DNA using this technique; initial experiments show that while this technique did not have a high yield, it was possible to isolate DNA from the solution; and finally, experiments to image the modified DNA are discussed in Appendix 1. While images produced are not conclusive to confirm the structures produced, it was possible to acquire low resolution images of the structures. These images also provide some evidence to confirm the presence of probe structures on the DNA backbone.

Appendix 2 contains the DNA sequences used for all DNA assembly experiments, and the random sequence generator code developed. Appendix 3 contains the code developed for conductance data analysis. Appendix 4 contains the codes developed for noise analysis. Appendix 5 contains the code for the analysis of translocation data, and subevent analysis. Appendix 6 contains the code developed to simulate DNA translocation data.

1.8 References

Abe, R., Hirasawa, H., Oda, S., Sadahiro, T., Nakamura, M., Watanabe, E., et al. (2008) Up-Regulation of Interleukin-10 mRNA Expression in Peripheral Leukocytes Predicts Poor Outcome and Diminished Human Leukocyte Antigen-DR Expression on Monocytes in Septic Patients. J Surg Res 147: 1–8.

Alsaif, M., Guest, P.C., Schwarz, E., Reif, A., Kittel-Schneider, S., Spain, M., et al. (2012) Analysis of serum and plasma identifies differences in molecular coverage, measurement variability, and candidate biomarker selection. Proteomics - Clin Appl 6: 297–303.

Battaglia, F., Meucci, V., Tognetti, R., Bonelli, F., Sgorbini, M., Lubas, G., et al. (2020) Procalcitonin detection in veterinary species: Investigation of commercial elisa kits. Animals 10: 1–12.

Brock, T.D., and Freeze, H. (1969) Thermus aquaticus gen. n. and sp. n., a nonsporulating extreme thermophile. J Bacteriol 98: 289–297.

Broza, Y.Y., Zhou, X., Yuan, M., Qu, D., Zheng, Y., Vishinkin, R., et al. (2019) Disease Detection with Molecular Biomarkers: From Chemistry of Body Fluids to Nature-Inspired Chemical Sensors. Chem Rev 119: 11761–11817.

Centers for Disease Control and Prevention (2018) What is sepsis? | Sepsis | CDC. Centers Dis Control Prev.

Dangi, A.K., Sinha, R., Dwivedi, S., Gupta, S.K., and Shukla, P. (2018) Cell line techniques and gene editing tools for antibody production: A review. Front Pharmacol 9: 630.

Dellinger, R.P., Levy, M.M., Rhodes, A., Annane, D., Gerlach, H., Opal, S.M., et al. (2013) Surviving Sepsis Campaign. Crit Care Med 41: 580–637.

Deutschman, C.S., and Tracey, K.J. (2014) Sepsis: Current dogma and new perspectives. Immunity 40: 463–475.

Dunne, P.W., Roberts, D.L., Quinones, M.P., Velligan, D.I., Paredes, M., and Walss-Bass, C. (2017) Immune markers of social cognitive bias in schizophrenia. Psychiatry Res 251: 319–324.

Erlich, H.A. (2015) PCR technology: Principles and applications for DNA amplification. 1st ed., Palgrave Macmillan, 7–22.

Faqi, A.S. (2017) A Comprehensive Guide to Toxicology in Nonclinical Drug Development. 2nd ed., Elsevier, 326-496.

FDA-NIH Biomarker Working Group (2016) BEST (Biomarkers, EndpointS, and other Tools) Resource. Updat Sept 25 55 https://www.ncbi.nlm.nih.gov/books/NBK326791/. Accessed July 26, 2021.

Flanagan, R.J., and Jones, A.L. (2004) Fab antibody fragments: Some applications in clinical toxicology. Drug Saf 27: 1115–1133.

Gao, F., Gao, Y., Zhang, S.J., Zhe, X., Meng, F.L., Qian, H., et al. (2017) Alteration of plasma cytokines in patients with active epilepsy. Acta Neurol Scand 135: 663–669.

Gardner, A., Carpenter, G., and So, P.-W. (2020) Salivary Metabolomics: From Diagnostic Biomarker Discovery to Investigating Biological Function. Metab 2020, Vol 10: 47.

Gazquez, C., Ribal, M.J., Marín-Aguilera, M., Kayed, H., Fernández, P.L., Mengual, L., and Alcaraz, A. (2012) Biomarkers vs conventional histological analysis to detect lymph node micrometastases in bladder cancer: A real improvement? BJU Int 110: 1310–1316.

Gilboa, T., Garden, P.M., and Cohen, L. (2020) Single-molecule analysis of nucleic acid biomarkers – A review. Anal Chim Acta 1115: 61–85.

Gomez-Marquez, J., and Hamad-Schifferli, K. (2021) Local development of nanotechnology-based diagnostics. Nat Nanotechnol 16: 484–486.

Hack, C.E., Hart, M., Schijndel, R.J.M.S. Van, Eerenberg, A.J.M., Nuijens, J.H., Thijs, L.G., and Aarden, L.A. (1992) Interleukin-8 in sepsis: Relation to shock and inflammatory mediators. Infect Immun 60: 2835–2842.

Hanly, W.C., Artwohl, J.E., and Bennett, B.T. (1995) Review of Polyclonal Antibody Production Procedures in Mammals and Poultry. ILAR J 37: 93–118.

Hansson, O., Finnebraaten, M., Heitmann, I.K., Ramse, M., and Bouzga, M. (2009) Trace DNA collection-Performance of minitape and three different swabs. Forensic Sci Int Genet Suppl Ser 2: 189–190.

Haschek, W.M., Rousseaux, C.G., Wallig, M.A., Bolon, B., and Ochoa, R. (2013) Haschek and Rousseaux's Handbook of Toxicologic Pathology. 1st ed., Elsevier Inc., 9-20.

Hirsch, M.S., and Watkins, J. (2020) A comprehensive review of biomarker use in the gynecologic tract including differential diagnoses and diagnostic pitfalls. Adv Anat Pathol 27: 164–192.

Hottenstein, C., Szapacs, M., Fuller, K., and Evans, C. (2017) Platforms and techniques used for biomarker assays: where are we now? https://doi.org/104155/bio-2017-0107 9: 1029–1031.

Hsu, W.Y., Chen, C.J., Huang, Y.C., Tsai, F.J., Jeng, L. Bin, and Lai, C.C. (2013) Urinary nucleosides as biomarkers of breast, colon, lung, and gastric cancer in Taiwanese e81701. PLoS One 8: e81701.

Janeway, C., Travers, P., Walport, M., and Capra, J. (2005) Immunobiology: the immune system in health and disease. 5th ed., New York: Garland Science, 1-62.

Jeong, J.A. (2014) ELISA technical guide and protocols_ thermo. Appl Mech Mater 597: 421–424.

Jia, H., Guo, Y., Zhao, W., and Wang, K. (2014) Long-range PCR in next-generation sequencing: Comparison of six enzymes and evaluation on the MiSeq sequencer. Sci Rep 4: 1–8.

Kasteren, P.B. van, Veer, B. van der, Brink, S. van den, Wijsman, L., Jonge, J. de, Brandt, A. van den, et al. (2020) Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. J Clin Virol 128: 104412.

Kemmler, M., Sauer, U., Schleicher, E., Preininger, C., and Brandenburg, A. (2014) Biochip point-of-care device for sepsis diagnostics. Sensors Actuators, B Chem 192: 205–215.

Khan, T.K. (2016) Clinical Diagnosis of Alzheimer's Disease BT - Biomarkers in Alzheimer's Disease. 1st ed., Academic Press, 27-48.

Kilty, C.G., Keenan, J., and Shaw, M. (2007) Histologically defined biomarkers in toxicology. Expert Opin Drug Saf 6: 207–215.

Koopmann, J., Fedarko, N.S., Jain, A., Maitra, A., Iacobuzio-Donahue, C., Rahman, A., et al. (2004) Evaluation of osteopontin as biomarker for pancreatic adenocarcinoma. Cancer Epidemiol Biomarkers Prev 13: 487–91.

Korman, N.J. (2020) Management of psoriasis as a systemic disease: what is the evidence? Br J Dermatol 182: 840–848.

Kotiw, M., Johnson, M., Pandey, M., Fry, S., Hazell, S.L., Netter, H.J., et al. (2012) Immunological response to parenteral vaccination with recombinant hepatitis B virus surface antigen virus-like particles expressing Helicobacter pylori KatA epitopes in a murine H. pylori challenge model. Clin Vaccine Immunol 19: 268–276.

Kruspe, S., and Giangrande, P.H. (2017) Aptamer-siRNA chimeras: Discovery, progress, and future prospects. Biomedicines 5: 5–45.

Kuhn, M., Sühs, K.W., Akmatov, M.K., Klawonn, F., Wang, J., Skripuletz, T., et al. (2018) Mass-spectrometric profiling of cerebrospinal fluid reveals metabolite biomarkers for CNS involvement in varicella zoster virus reactivation. J Neuroinflammation 15: 1–15.

Kumar, Y., Qunibi, M., Neal, T.J., and Yoxall, C.W. (2001) Time to positivity of neonatal blood cultures. Arch Dis Child - Fetal Neonatal Ed 85: 182F – 186.

Laterza, O.F., Hendrickson, R.C., and Wagner, J.A. (2007) Molecular biomarkers. Drug Inf J 41: 573–585.

Li, X., Xu, Z., Pang, X., Huang, Y., Yang, B., Yang, Y., et al. (2017) Interleukin-10/lymphocyte ratio predicts mortality in severe septic patients. PLoS One 12: e0179050.

Lin, G.L., McGinley, J.P., Drysdale, S.B., and Pollard, A.J. (2018) Epidemiology and Immune Pathogenesis of Viral Sepsis. Front Immunol 9: 2147.

Liszewski, M.K., Farries, T.C., Lublin, D.M., Rooney, I.A., and Atkinson, J.P. (1996) Control of the complement system. Adv Immunol 61: 201–283.

Litman, G.W., Rast, J.P., Shamblott, M.J., Haire, R.N., Hulst, M., Roess, W., et al. (1993) Phylogenetic diversification of immunoglobulin genes and the antibody repertoire. Mol Biol Evol 10: 60–72.

Mallikaratchy, P. (2017) Evolution of complex target SELEX to identify aptamers against mammalian cell-surface antigens. Molecules 22: 22–215.

Matsuda, F. (2016) Technical Challenges in Mass Spectrometry-Based Metabolomics. Mass Spectrom 5: S0052–S0052.

Medicine, I. of (2004) Physiological Biomarkers for Predicting Performance. In Monitoring Metabolic Status: Predicting Decrements in Physiological and Cognitive Performance. National Academies Press (US), 103.

Medzhitov, R., and Janeway, C.A. (1997) Innate immunity: impact on the adaptive immune response. Curr Opin Immunol 9: 4–9.

Moore, K.W., Waal Malefyt, R. De, Coffman, R.L., and O'Garra, A. (2001) Interleukin-10 and the interleukin-10 receptor. Annu Rev Immunol 19: 683–765.

NHS (2020) Blood tests - NHS. NHS https://www.nhs.uk/conditions/blood-tests/. Accessed July 26, 2021.

NHS (2021a) Interleukin - 6 (IL-6) - North West London Pathology. https://www.nwlpathology.nhs.uk/tests-database/interleukin-6-il-6/. Accessed June 1, 2022.

NHS (2021b) C-reactive protein (CRP) - North West London Pathology. https://www.nwlpathology.nhs.uk/tests-database/c-reactive-protein-crp/. Accessed June 1, 2022.

NHS (2021c) Procalcitonin - North West London Pathology. https://www.nwlpathology.nhs.uk/tests-database/procalcitonin/. Accessed June 1, 2022.

Nordström, A., and Lewensohn, R. (2009) Metabolomics: Moving to the Clinic. J Neuroimmune Pharmacol 2009 51 5: 4–17.

Parlato, M., and Cavaillon, J.M. (2014) Host response biomarkers in the diagnosis of sepsis: A general overview. Methods Mol Biol 1237: 149–211.

Peeling, R.W., Murtagh, M., and Olliaro, P.L. (2019) Epidemic preparedness: why is there a need to accelerate the development of diagnostics? Lancet Infect Dis 19: e172–e178.

Ramazzotti, D. (2014) An Observational Study: The Effect of Diuretics Administration on Outcomes of Mortality and Mean Duration of I.C.U. Stay. 1–75.

Raymond, J.J., Oorschot, R.A.H. van, Gunn, P.R., Walsh, S.J., and Roux, C. (2009) Trace evidence characteristics of DNA: A preliminary investigation of the persistence of DNA at crime scenes. Forensic Sci Int Genet 4: 26–33.

Reth, M. (2013) Matching cellular dimensions with molecular sizes. Nat Immunol 14: 765–767.

Russell, C., Ward, A.C., Vezza, V., Hoskisson, P., Alcorn, D., Steenson, D.P., and Corrigan, D.K. (2019) Development of a needle shaped microelectrode for electrochemical detection of the sepsis biomarker interleukin-6 (IL-6) in real time. Biosens Bioelectron 126: 806–814.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science (80-) 230: 1350–1354.

Selleck, M.J., Senthil, M., and Wall, N.R. (2017) Making Meaningful Clinical Use of Biomarkers. Biomark Insights 12: 1–7.

Sigma-Aldrich (2018) DNA Oligonucleotide Synthesis. https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/genomics/pcr/dna-oligonucleotide-synthesis. Accessed July 26, 2021.

Sinclair, K., and Dudley, E. (2019) Metabolomics and Biomarker Discovery. Adv Exp Med Biol 1140: 613–633.

Smith, J.J., Sorensen, A.G., and Thrall, J.H. (2003) Biomarkers in imaging: Realizing radiology's future. Radiology 227: 633–638.

Solier, C., and Langen, H. (2014) Antibody-based proteomics and biomarker research-current status and limitations. Proteomics 14: 774–783.

Srinivas, P.R., Kramer, B.S., and Srivastava, S. (2001) Trends in biomarker research for cancer detection. Lancet Oncol 2: 698–704.

Stein, U., Meller, A., Rozevsky, Y., Gilboa, T., Kooten, X.F. Van, Kobelt, D., and Huttner, D. (2020) Quantification of mRNA expression using single-molecule nanopore sensing. ACS Nano 14: 13964–13974.

Stenman, U.H. (2016) Biomarker development, from bench to bedside. Crit Rev Clin Lab Sci 53: 69–86.

Strimbu, K., and Tavel, J.A. (2010) What are Biomarkers? Curr Opin HIV AIDS 5: 463.

Suganuma, M., Yamaguchi, K., Ono, Y., Matsumoto, H., Hayashi, T., Ogawa, T., et al. (2008) TNF-α-inducing protein, a carcinogenic factor secreted fromH. pylori, enters gastric cancer cells. Int J Cancer 123: 117–122.

Taeb, A.M., Hooper, M.H., and Marik, P.E. (2017) Sepsis: Current definition, pathophysiology, diagnosis, and management. Nutr Clin Pract 32: 296–308.

The Editors of Encyclopedia Britannica (2020) antibody | Definition, Structure, Function, & Types | Britannica. https://www.britannica.com/science/antibody. Accessed August 2, 2021.

The UK Sepsis Trust (2019) The UK SEPSIS Trust FAQs.

Tiede, C., Bedford, R., Heseltine, S.J., Smith, G., Wijetunga, I., Ross, R., et al. (2017) Affimer proteins are versatile and renewable affinity reagents. Elife 6: 24903.

Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science (80-) 249: 505–510.

U.S. Food & Drug Administration (2019) About Biomarkers and Qualification. fda.gov https://www.fda.gov/drugs/biomarker-qualification-program/about-biomarkers-and-qualification. Accessed July 26, 2021.

Vliet, E.A. van, Puhakka, N., Mills, J.D., Srivastava, P.K., Johnson, M.R., Roncon, P., et al. (2017) Standardization procedure for plasma biomarker analysis in rat models of epileptogenesis: Focus on circulating microRNAs. Epilepsia 58: 2013–2024.

Wild, D., John, R., and Sheehan, C. (2013) The Immunoassay Handbook Theory and applications of ligand binding , Edited by Associate Editors. 4th ed., Elsevier Ltd, 29–242.

Woof, J.M., and Burton, D.R. (2004) Human antibody–Fc receptor interactions illuminated by crystal structures. Nat Rev Immunol 2004 42 4: 89–99.

Yoshio Hirao, Tsugiyasu Kanda, Yoshimasa Aso, Masato Mitsuhashi, and Isao Kobayashi (2000) Interleukin-8—An Early Marker for Bacterial Infection. Lab Med 31: 39–44.

Zinellu, A., Arru, F., Vito, A. De, Sassu, A., Valdes, G., Scano, V., et al. (2021) The De Ritis ratio as prognostic biomarker of in-hospital mortality in COVID-19 patients. Eur J Clin Invest 51: e13427.

Chapter 2

Sterically Controlled Nuclease Enhanced DNA Assembly

Synopsis: This chapter introduces DNA modification and assembly techniques. The primary aim of this project was the development of a new DNA based technique for multi-biomarker capture. This chapter discusses the method developed to achieve this aim along with the techniques used and modified to confirm structure assembly. The integration of a patient sample isolation technique has also been considered but is further discussed in Appendix 1.2.

2.1 Introduction

For the development of a rapid, multi-biomarker capture system for sepsis diagnostics, it is possible to utilise and improve upon several existing techniques and exploit naturally occurring processes. The natural processes and established techniques which are later developed upon are discussed in this introduction.

2.1.1 Naturally occurring DNA modifications

All biological systems contain an array of different modifications which contribute to overall cellular activity (Nowak and Corces, 2004; Mateus-Pinheiro *et al.*, 2011; Harmel and Fiedler, 2018; Schvartzman *et al.*, 2018). With proteins, post translational modification is a method used to activate, deactivate, and localise proteins. This is also an important mechanism for regulating certain cellular processes dependant on cellular health (Levine, 2002; Konstantinova *et al.*, 2008; Olszewski *et al.*, 2010). These modifications include sub-unit cleavage, phosphorylation, glycosylation, ubiquitination, methylation, acetylation, and lipidation. Modifications are regulated by a collection of enzymes in the proteome, which are themselves regulated by cellular conditions (Berg *et al.*, 2002).

DNA modification is a natural process which occurs to regulate the expression of genes and has been linked to the pathogenesis of many diseases from cancer to neurological disorders (Liyanage *et al.*, 2014). The field of studying these modifications is epigenetics. There are several modifications by which DNA is directly modified including methylation, hydroxymethylation, carboxylation, and formylation (Dupont *et al.*, 2009; Golla *et al.*, 2014; Wen and Tang, 2014; Klungland and Robertson, 2017; Kumar *et al.*, 2018). Gene expression is also regulated through the modification of histones, around which DNA is wrapped in the nucleus (Farrelly *et al.*, 2019). This is a

relatively new field of exploration, and research has primarily focused on understanding the effect of methylation on gene expression (Dupont *et al.*, 2009).

DNA methylation is regulated by a collection of methyl transferase enzymes (Bestor, 2000). These enzymes use S-Adenosyl methionine (SAM) as a methyl donor for transfer to a DNA base, as illustrated in Figure 2.1. The most common base modified by these enzymes is cytosine. DNA contains regions with a high percentage of cytosine and guanine bases called CpG islands, which are the primary location for mammalian DNA methylation (Berg *et al.*, 2002).

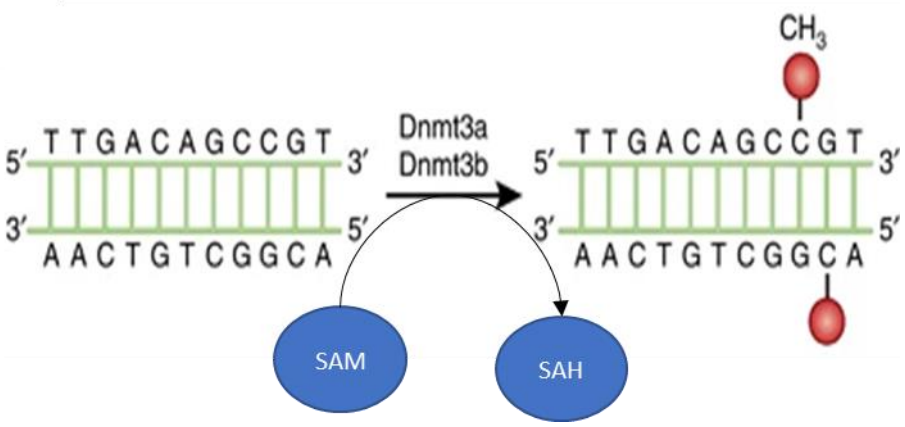

Figure 2.1. Simplified diagram of DNA methylation, where SAM donates it methyl group to become S-Adenosyl-L-homocysteine (SAH) through the action of DNA methyl transferases (Dnmt) 3a and 3b. Figure adapted from Figure 1 (Moore et al., 2013).

The *Thermus aquaticus* extremophile produces a methyl transferase enzyme (M.TaqI) which recognises TCGA sites, transferring the methyl group to the adenosine residue (Nelson and McClelland, 1987). Transfer of the methyl group does not always occur symmetrically on both strands, often DNA is hemimethylated. When this occurs, it has been linked with structural changes to the backbone, preventing the binding and recognition of sites by methyl transferases (Woodcock *et al.*, 1997; Couldrey *et al.*, 2015; Welsh *et al.*, 2017).

There has been much research into modifying DNA for sensing and analytics. These modifications have included a range of molecules including antibodies, fluorescent tags, aptamers, and ssDNA fragments.

In research conducted by Wand *et al.*, a bacterial genome extract was modified at specific M.TaqI sites and utilised to specifically “barcode” the DNA. The fluorescent reporter used was an Atto 647N-labelled NHS ester. This allowed for identification of the genome via fluorescence imaging. As each species’ genome has a different pattern of M.TaqI sites, it was possible to differentiate the species by the pattern of fluorescence. Due to the specificity of the enzyme and that the dibenzocyclooctyne (DBCO) - Atto 647N cannot bind to unmodified DNA, the data acquired highlights the efficiency of the DNA to DBCO- Atto 647N conjugate (Wand *et al.*, 2019).

2.1.2 Click chemistry

Click chemistry is the term for a highly efficient, stereospecific, water-soluble reaction often occurring between an azide and an alkyne. The two main types of reaction are copper catalysed and copper free cycloaddition. Copper catalysed reactions are preferable for molecular synthesis due to single isomer formation and can be performed at room temperature (Kolb *et al.*, 2001). However due to the interaction of copper and DNA, potential cleavage, cellular death, or structure alteration is possible when performing this type of reaction with biosystems (Devaraj and Finn, 2021). Therefore, it is necessary to perform a copper-free reaction when performing click reactions with biosystems. In biochemistry, click chemistry can utilise native biological systems to introduce intermediate, click-ready molecules into the environment, such as modified nucleotides (for example 5-ethynyl-2’-deoxyuridine (EdU)) in culture medium. Cellular systems treat the analyte as it would its natural counterpart, and in

the case of EdU, as this is a thymidine analogue, is incorporated into the DNA of a dividing cell. As EdU has a alkyne group incorporated into the molecule, it is then possible to introduce a complementary azide which then will only react at the EdU locations (Kharas *et al.*, 2008; Chehrehasa *et al.*, 2009; Dorsett *et al.*, 2009). Other groups often used to create “click” sites in biological systems are modified SAM molecules containing an azide group and DBCO as illustrated in Figure 2.2. The use of modified SAM molecules allows researchers to utilise methyl transferase function. While typical neat reactions require heat input to complete the click reaction, the use of DBCO overcomes this obstacle, allowing reactions to occur at physiological conditions (Devaraj and Finn, 2021).

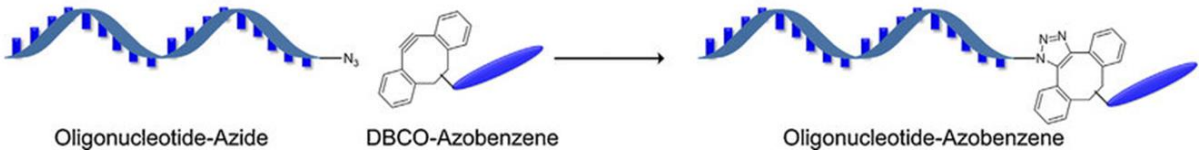

Figure 2.2. Diagram illustrating a copper free click reaction between an azide modified oligonucleotide and DBCO-Azobenzene (a photo-switchable compound). Image adapted from Figure 2 (Gu *et al.*, 2015)

In research conducted by Gong *et al.,* an antibody-DBCO to azide oligonucleotide conjugate was created to detect specific proteins through nucleic acid detection. Once conjugation has been performed as previously described, the antibody-oligonucleotide complex binds to a protein of interest. The oligonucleotides have a 6 bp complementary region, allowing 2 of the sequences to bind to one another. The addition of DNA synthesis enzymes then allows for the sequences to be extended before qPCR is performed. This amplification and detection then allows for the relative quantification of the protein of interest (Gong *et al.*, 2016). This research highlights the ability to create functional DNA-capture probe conjugates utilising a click reaction.

2.1.3 Methods for DNA assembly

2.1.3.1 DNA-Assay-on-a-string

There are several different methods which have been employed to incorporate functional biomolecular probes using DNA as the carrier strand. Henceforth, the term probes will refer to any biomolecule capture system as discussed in section 1.4, unless stated. Many of these methods rely on single stranded genomic DNA hybridisation to short oligonucleotides with designed overhangs or binding groups. In research conducted by Keyser and Bell, a M13mp18 ssDNA genome (7.2 kbp) was hybridised with 190 oligonucleotides with a length of 38 bases each. A select few short strands included an additional thymidine overhang motif to prevent dimerization creating an oligonucleotide with a length of 46 bases. As the 8 additional bases are not recognised by the ssDNA genome sequence, an overhang region is created which protrudes from the newly formed dsDNA strand. The overhang regions were also modified to contain biotin groups such that streptavidin can be bound, and the protrusions detected using nanopore sensing (Bell and Keyser, 2015; Plesa *et al.*, 2015). In research conducted by Loh *et al.,* this concept was further investigated utilising the single stranded overhangs as direct genetic capture probes, hereto also referred to as probes. The probes are comprised of specific DNA sequences complementary to an 88 base long biomarker sequence, as illustrated in Figure 2.3. It was also suggested that the use of antibodies and aptamers in combination to capture complementary proteins and metabolites could be employed. The research concluded that it was possible to differentiate bound and unbound probes through nanopore sensing (Loh *et al.*, 2018).

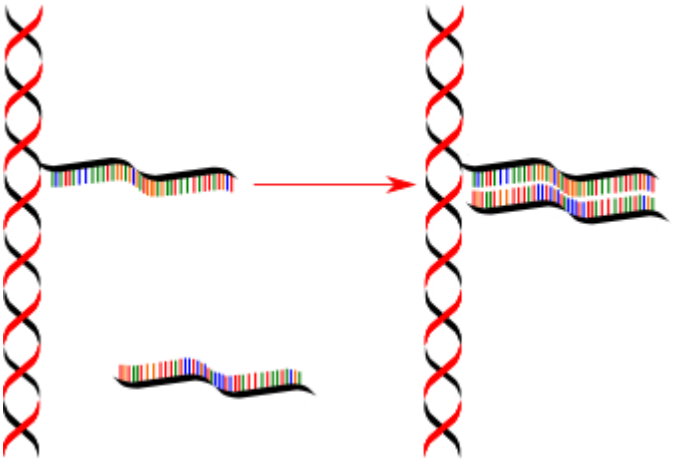

Figure 2.3. Illustration of the probe ready to accept the complimentary strand for IL-6.

Further developments to incorporate biomarker capture altered the form in which probes were attached to the DNA carrier. In research conducted by Sze *et al.,* λ-DNA was used as the carrier containing 12 base overhangs on the 5’ ends. To these overhangs, nucleotide aptamers were conjugated and functionalised. It was also possible to add an additional third probe on one end of the DNA backbone by using a portion of the sequence from the initial aptamer for hybridisation. The research concluded that it was possible to capture thrombin at all three positions and detect this using nanopore sensing (Sze *et al.*, 2017).

The detection of biomolecules using nanopore sensing will be further discussed in Chapter 5.

2.1.3.2 Gibson assembly

Gibson assembly is a well-known technique for adjoining blunt ended DNA fragments rather than utilising enzyme restriction sites, as is typical. This technique employs three enzymes, a 5’ exonuclease, a DNA polymerase, and a DNA ligase as illustrated in Figure 2.4. The blunt ended DNA must contain specific overlapping sequences from the 3’ end as the 5’ exonuclease digests up to 100 bp on the 5’ end of the DNA fragment. The now single stranded regions of the DNA fragments overlap reforming dsDNA. The DNA polymerase then fills in any gaps left during the digestion step and the DNA ligase joins breaks in the phosphodiester backbone. The blunt DNA fragments

are limited to greater than 100 bp. This is attributed to the stability of the DNA fragments used as well as the activity of the 5' exonuclease. Below 100 bp, DNA fragments are often entirely digested due to the action of the exonuclease, rendering it impossible to create a stable fragment capable of further assembly (Gibson *et al.*, 2009).

The research conducted by Gibson *et al.,* developed a technique to adjoin multiple large fragments of DNA together. The technique demonstrates that it is possible to adjoin blunt ended DNA together by creating sticky ended overhangs through 5' exonuclease digestion (Gibson *et al.*, 2009).

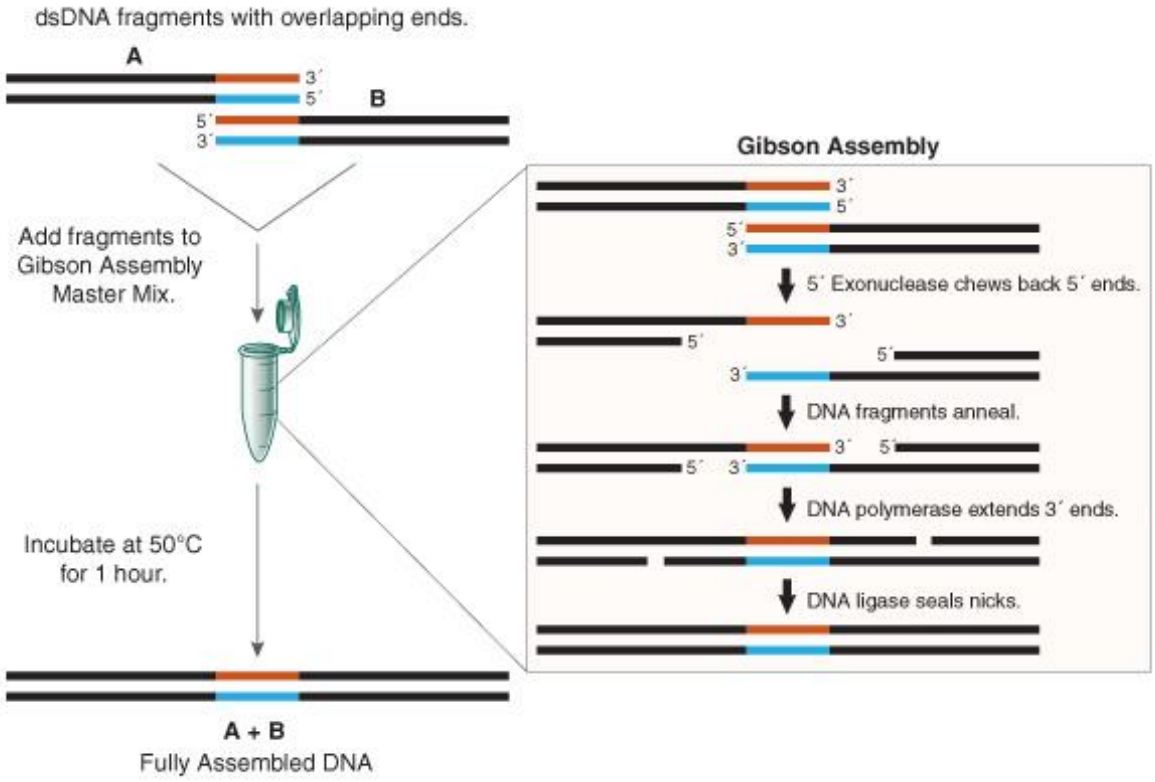

Figure 2.4. Simplified diagram of the Gibson assembly process highlighting the assembly of 2 different, blunt ended DNA fragments (A and B) Figure acquired from NEB (Gibson and Russello, 2010)

2.1.4 Furthering sepsis diagnostics

Given the complexity and variety of biomarkers associated with sepsis, there is a clear need for a technique with the capability to detect multiple analytes simultaneously and quickly (Kemmler *et al.*, 2014). Advances in other fields, for example cancer diagnostics, have allowed development of more advanced biosensors, based on genetic testing (Pujol *et al.*, 2021). However, the necessity for high concentrations of DNA within samples has led to the further requirement of PCR amplification, thus increasing the time to diagnosis (Vuković *et al.*, 2021). This issue is further compounded by the fact that some hospitals do not have access to genetic laboratories, meaning samples must be sent to a larger facility, further increasing time to diagnosis (BSGM, 2021). Combinations of electrochemistry with DNA technology presents an innovative solution potentially enabling rapid and accurate detection of biomarkers critical in sepsis, and other diseases (Pellitero *et al.*, 2020). Recent improvements in DNA technology, in combination with electrochemistry, have allowed for the design of rapid and accurate detectors. A bioelectrode utilising single stranded DNA with specific end strands, encoding mutated BRCA1 genes, adhered onto a chitosan-co-polyaniline and indium-tin-oxide surface detected 3.4 ng/µL of the mutated gene. This is in comparison to normal DNA testing which usually requires enrichment if the initial concentration is below 4 ng/µL (Tiwari and Gong, 2009).

When using DNA as a backbone to create a multi-biomolecule sensor, there is a need to create an ordered system for directional capture and sensing, allowing for the positioning of the bound or unbound probe, by the sensor, to be determined. This could impact the diagnosis or prognosis dependant on the biomarkers used for sensing and

is most significant in nanopore sensing where DNA can translocate both forwards and backwards. This is further discussed in section 5.1.

As discussed in this chapter, it is possible to create DNA based sensors for biomolecules, however these are often at specific enzymatic sites, contain a random distribution of sensing probes, or are limited to a single sensor per DNA backbone strand. To further diagnostic capability of these techniques, there is a need to create sensors that can be of any DNA length and have multiple different capture probes in a specific order.

2.1.5 *Sterically Controlled Nuclease Enhanced DNA Assembly*

Sterically controlled nuclease enhanced (SCoNE) DNA assembly is a technique that has been developed during this project for the preparation of multi-functional DNA nanostructures. This technique incorporates elements from enzymatic DNA modification, click chemistry and Gibson assembly, to create an overall “one-pot” reaction. As is illustrated in Figure 2.5, a technique capable of being adapted for biomarker-based diagnostics has been developed, so long as a biomarker for a particular disease is known and a probe can be developed against the analyte. Two different types of DNA strands have been designed to construct the final capture system. These are separated into two categories; short probe strand DNA (34 – 100 bp) containing a single M.TaqI modification site, and spacer strand DNA (>100 bp), referred to as sDNA. The short probe strand DNA and sDNA have a minimum 15 base long overlapping region on the 3’ ends which is complementary to the corresponding strand. These strands are designed such that the short probe strand DNA strand can only adjoin to the corresponding sDNA strand. Aminated probes including, but not limited to, aptamers, antibodies, and fab fragments are conjugated to DBCO via an

amide bond, leaving the alkyne bond available for further reactions, as illustrated in Figure 2.5.A. The short probe DNA strands are modified using an azide containing cofactor analogue and M.TaqI to conjugate the azide group to the fragment, as illustrated in Figure 2.5.B. An alkyne-azide reaction is then performed between azide-DNA intermediate and the probes of interest, as illustrated in Figure 2.5.C, and henceforth referred to as pDNA. pDNA was also designed to be used as an isolation fragment. To do this a DBCO-biotin molecule was conjugated to the azide-DNA intermediate to create isolation DNA fragments, referred to as iDNA. Due to our interest in sepsis, the capture and analysis of IL-6 and procalcitonin was focused on. Each pDNA and iDNA is created individually for maximum control over probe-strand conjugation. To create the final SCoNE structure all pDNA, iDNA and sDNA strands are added to a single reaction vessel along with a T5'-exonuclease, DNA ligase, and DNA polymerase enzyme mixture, as illustrated in Figure 2.5.D and E.

Due to the adaptable nature of the technique, pDNA can be created and stored without conjugation to a probe. This allows rapid production of the final construct with a collection of probes, specific to the disease under investigation. Through the current experimental conditions, full assembly and analysis by resistive pulse sensing can be performed in under 24 hours. The starting reagents are also designed such that PCR amplification can be performed on the sDNA strands, reducing overall cost and allowing large volumes of starting material to be stockpiled ready for use. Fully constructed SCoNE structures can also be frozen and stored up to several weeks prior to use.

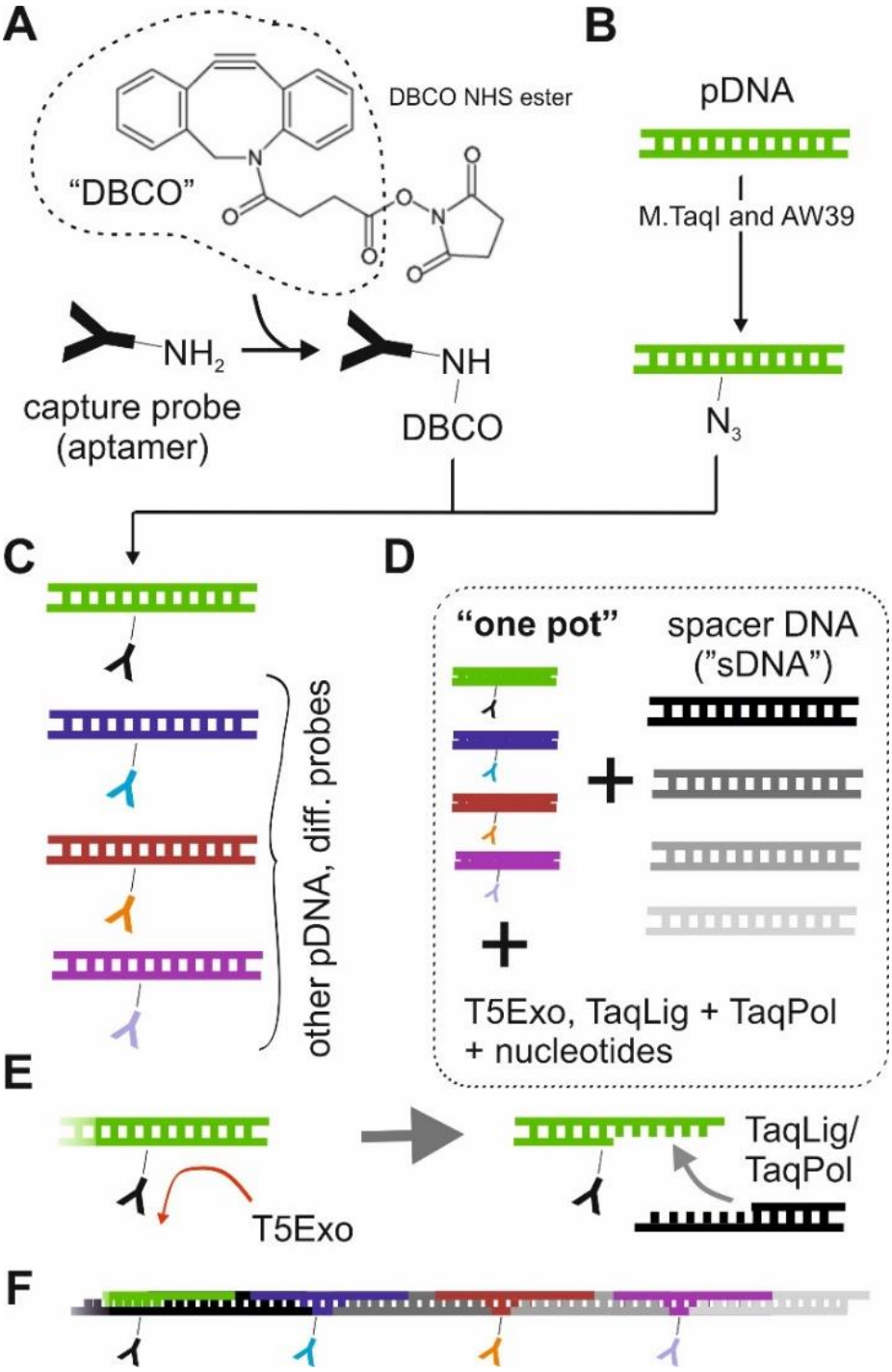

Figure 2.5. The SCoNE DNA assembly technique. Aminated capture probe/s are conjugated to DBCO (A). pDNA strands are azidated using M.TaqI and AdoHcy-azide (AW39) (B). Independently, modified probes and their respective undergo a copper free click reaction to create the final pDNA strand (C). Functionalised pDNA strands are combined into a single vial along with sDNA fragments, nucleotides, and the enzymes T5 exonuclease, Taq DNA Ligase, and Taq DNA polymerase (D). The 5' ends of the p and sDNA are digested creating sticky ends specific to their corresponding strands. The T5 exonuclease is unable to navigate across the modified probe attachment and therefore releases the DNA (E). DNA fragments are therefore able to form sequence specific alignment, forming one continuous dsDNA strand. DNA polymerase fills in any gaps created during digestion and DNA ligase seals the strands together forming phosphodiester bonds (F).

2.2 Methodology

To create a SCoNE structure, the technique is split into several individual stages with all DNA fragment and primer sequences available in Appendix 2.2 and 2.3. All reactions and incubations were performed in low DNA binding PCR tubes (ThermoFisher), unless otherwise stated, and dilutions using Nuclease-Free Water for Molecular Biology (Merck) unless otherwise stated. All purifications were performed using a PCR clean-up kit (GenElute^{TM}, Sigma-Aldrich), unless stated. First, the wash solution was prepared by adding 12 mL of the wash solution concentrate to 48 mL of 100% ethanol (HPLC grade, Sigma-Aldrich). The mini spin column was inserted into the provided collection tube. To the column, 0.5 mL of the column preparation solution was added and centrifuged at 13,500 x g for 1 minute. For this procedure, all centrifugation steps were performed at 13,500 x g. The solution was discarded, and the column re-inserted into the collection tube. To the sample, 5 times its volume of binding buffer was added and mixed well. The solution was then transferred into the column and centrifuged for 1 minute. The elute was then discarded and the column re-inserted into the collection tube. 0.5 mL of the diluted wash solution was then added to the column and centrifuged. The elute was then discarded, the column re-inserted into the collection tube and centrifuged for 2 minutes. The column was transferred to a new collection tube and incubated with 20 µL of nuclease-free water (Merck) for 1 minute. This was then centrifuged for 1 minute, the elute collected in a separate collection tube, and stored at -20°C until use. All enzymes used were acquired from NEB. DNA concentration determination was performed using 2 µL of sample and UV-Vis spectroscopy, measuring the absorbance at 260 nm (BioSpec-nano,Shimadzu).

2.2.1 sDNA preparation

sDNA was acquired in plasmid form from GeneArt (Invitrogen), using primarily randomly generated DNA sequences with designed overhang regions on both 5' and 3' ends. The code to generate these sequences is provided in Appendix 2.1. sDNA was digested using SfiI to excise the 1000 bp sDNA fragment from the plasmid using the conditions provided in Table 1.

Table 1. Conditions for SfiI digestion of sDNA fragments

Reagent	Volume (µL)
SfiI (20,000 units/mL)	1
DNA (5µg/mL)	2
10X Cutsmart buffer (NEB)	5
Nuclease free water	42

The reaction was performed at 50°C for 1 hour then purified, as described in section 2.2, prior to PCR amplification. This was performed to isolate the 1 kbp sDNA fragment, increasing the efficiency of the PCR reaction, and ensuring only the 1 kbp fragment of interest was amplified. Concentration was determined using UV-visible spectroscopy, as described in section 2.2. A 25 µL PCR reaction was performed using a Taq PCR Kit (NEB in a 0.5 mL PCR tube). The volumes of reaction buffers and components were added into a single reaction vial, as described in Table 2. Due to differing yields of DNA from the SfiI digestion, the volume of DNA and nuclease free water was adjusted to compensate for this difference.

Table 2. Component and volume requirements for PCR

Component	Volume to add in a 25µL reaction (µL)	Final concentration
10X standard Taq rection buffer	2.5	1X
10mM dNTPs	0.5	200 µM
10µM forward primer	0.5	0.2 µM
10µM reverse primer	0.5	0.2 µM
Template DNA	0.8-1.2	10 nM
Taq DNA polymerase	0.125	1.25 enzymatic units
Nuclease free water	19.675 -20.075	/

This solution was gently homogenised using a wide pipette, to prevent breaking of the DNA fragments, before being gently spun down in a microcentrifuge. The PCR tube was placed in a thermocycler (PrimeG, Techne) under the conditions described in Table 3. Where rows are highlighted in light grey, these steps were cycled together in the order written in the table.

Table 3. PCR cycle protocol

Cycle step	Temperature (°C)	Time (s)	Cycles performed
Denaturation (initial)	95	30	1
Denaturation	95	20	30
Annealing	50	30	
Extension	68	60	
Extension (final)	72	300	1
Hold	4	Until ready to remove	

The DNA was performed using a PCR clean-up kit (GenElute™, Sigma-Aldrich), and final DNA concentration was determined using UV-vis spectroscopy, as described in section 2.2. Additional PCR reactions were performed until a minimum final concentration of 600 ng/µL was achieved.

2.2.2 SCoNE DNA assembly

2.2.2.1 Probe preparation

Aminated aptamers (5'NH_2, Cambio) were briefly centrifuged and resuspended in Resuspension Buffer (#RTW0001, Cambio) using 11.2 µL and 24.9 µL for the human procalcitonin aptamer and the human IL-6 aptamer, respectively, for a 100X working concentration (ATW0060-GM3-25 and ATW0035-GM3-25 respectively, Cambio). Aptamers were selected based on commercial availability and high affinity (26.6 nM and 19 nM respectively). These were aliquoted and stored at -20°C until use.

To prepare DBCO, 1 mg of DBCO-NHS-ester (Sigma-Aldrich) was added to 1 mL of HPLC grade Dimethyl sulfoxide (DMSO, Sigma-Aldrich) to create a 2.5 mM stock. To 49 µL of nuclease free water (Merck), 1 µL of the 2.5 mM DBCO solution was added to create a working stock of 50 µM.

Aptamers were diluted to a 10X working concentration (10 µM) in aptamer folding buffer (RTW0003, Cambio), and heated to 95°C for 5 minutes, then left to cool to room temperature for 15 minutes prior to use. 2 µL of the prepared aptamer solution was added into a new PCR tube and diluted with 18 µL of a 50 µM DBCO-NHS ester solution and homogenised well. This was incubated at room temperature for a minimum of 1.5 hours on a shaker to allow conjugation of the probe to the DBCO. The reaction schematic is illustrated in Figure 2.5.A.

Probes such as antibodies, affimers, and fluorescent tags, containing a free amine group can be used in place of the aptamer, however this has not yet been explored.

2.2.2.2 pDNA azidation

Each pDNA was prepared in a separate vial under the same conditions. The reaction vessel contained 1 µL of the pDNA (IDT) strand at 600 ng/µL, 2 µL Cutsmart buffer 10X, 1 µL AW39 azide donor at 100 µM (illustration of the molecule as shown in Figure 2.6, synthesised by the Neely group at UoB (Wilkinson *et al.*, 2020)), 0.5 µL M.TaqI (NEB) and 15.5 µL of nuclease free water (Merck), adding the enzyme last. Each vial was incubated at 40°C for 1.5 hours. After which, 0.5 µL proteinase K (800 units/mL, NEB) was added and incubated at 40°C for 1 hour. Vials were removed from the heat and allowed to cool to room temperature for 20 minutes. The solution was then purified using a PCR clean-up kit (GenElute^{TM}, Sigma-Aldrich), as described in section 2.2, and stored at 4°C until required.

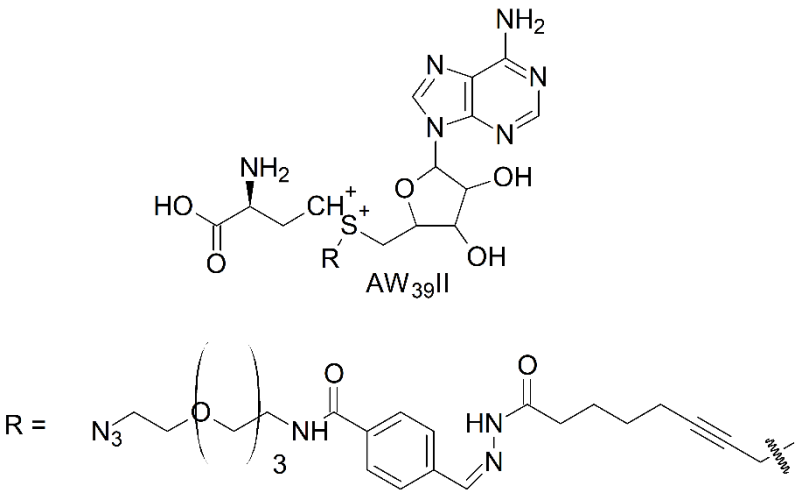

Figure 2.6. Illustration of AW39, with the SAM molecule modified with the azide containing R group as opposed to the standard methyl group. Image provided by the Neely group.

2.2.2.3 pDNA and iDNA assembly and conjugation

Assembly of the probe strands was performed in separate vials. To each reaction vial, 2 µL of the desired probe solution was added to 4 µL of the azidated pDNA strand. This was incubated at room temperature on a shaker for a minimum of 1.5 hours to allow the click reaction to take place. To prepare iDNA strands, the same method was

used replacing the prepared aptamer with 2 µL of 50 µM DBCO-dPEG®12-biotin (Sigma-Aldrich). The vials were stored at room temperature for up to 2 days prior to use.

2.2.2.4 Final assembly

All final assembly steps were performed in the same vial. sDNA fragments were diluted 1:5 to create a 1X working solution using nuclease free water (Merck). sDNA fragments and pDNA fragments were added into a single vial at equal concentrations (1:1 ratio), to achieve a final 1 pM solution, and homogenised well using a wide pipette tip. 10 µL of this solution was then removed and pipetted into a separate vial. 5 µL of a Gibson assembly master mix (NEB) and 5 µL of nuclease free water (Merck) added and homogenised gently. Vials were then incubated at 40°C for a minimum of 1.5 hours. SCoNE structures were assembled in a variety of different conformations, from dimers (2 pDNA or iDNA and 2 sDNA fragments) to decamers (10 pDNA or iDNA and 10 sDNA fragments). The reaction schematic is illustrated in Figure 2.5.D and E.

An in-house enzyme mix was also developed for use, including a Taq T5' exonuclease (NEB), a Taq ligase (Sigma-Aldrich), Taq polymerase (Sigma-Aldrich), and nucleotides (NEB) to replace the Gibson assembly master mix. The volumes and concentrations used are highlighted in Table 4. Due to the higher activity of the Taq DNA polymerase, a lower concentration of units was used, respective to the other added enzymes.

Table 4. Gibson assembly enzyme replacement reaction mix composition

Component	Starting concentration used	Volume added (µL)
T5 exonuclease	5000 Units	1
New England biolabs buffer 4	10X	5
Taq DNA ligase	2500 Units	2
DNA ligase reaction buffer	10X	5
Taq DNA polymerase	1500 Units	2
PCR buffer	10X	5
Nuclease free water	N/A	4
Nucleotide mix	10mM	1

2.2.3 Gel electrophoresis

A 1% agarose gel was used for all fragment separation and analysis. To 100 mL of 1X Tris-acetate-EDTA buffer (TAE, Tris (hydroxymethyl) aminomethane 1g, EDTA, tetrasodium < 1g, Water 98 mL, Disodium EDTA ~1g, 1,3-Propanediol, 2-amino-2-(hydroxymethyl)-, acetate (salt) ~1g, Fisher Scientific), 1 g of agarose powder (Sigma-Aldrich) was added. This was microwaved for 2 minutes until the agarose had completely dissolved. The solution was allowed to cool to 50°C then poured into a gel tray with a well comb in place and incubated at room temperature for 30 minutes for the gel to set. To 5 µL of each DNA sample, and DNA ladder (GeneRuler 1 kbp DNA Ladder, ThermoFisher), 1 µL of 6X DNA loading buffer (ThermoFisher) was added and homogenised. The gel was then placed into a Mini-Sub Cell GT Cell (BIORAD), and this filled with 1X TAE buffer until the gel was covered. The well comb was then removed and the first well filled with 5 µL of a DNA ladder (GeneRuler 1 kbp DNA Ladder, ThermoFisher). The remaining wells were filled with 5 µL of the DNA samples under analysis or DNA standards. After connecting the unit to the power pack, the gel was run at 75 V for 45 minutes. Once finished, the gel was then placed in 1X GelRed

(Sigma-Aldrich) for 45 minutes before imaging in a UV illumination box (BIORAD). The gel images were then analysed using Fiji-ImageJ.

2.2.4 DNA isolation

During development of the SCoNE protocol, two different extraction methods were used. Initially gel extraction was used, then with the development of the iDNA fragments, isolation was performed using Dynabeads™ MyOne™ Streptavidin C1 beads (ThermoFisher).

2.2.4.1 Gel extraction

The SCoNE DNA band of interest was excised from the gel using a sterile scalpel under UV visualisation. The DNA was extracted from the gel using a GenElute™ Gel Extraction Kit (Sigma-Aldrich). All centrifugation steps were performed at 13,500 x g using a microcentrifuge (MiniSpin, Eppendorf). The wash solution was prepared by adding 12 mL of the wash solution concentrate to 48 mL of 100% ethanol (HPLC grade, Sigma-Aldrich). The gel fragment was submerged in 1 mL of the gel solubilisation solution and incubated at 55°C for 10 minutes, vortexing every 2 minutes, until the gel had completely dissolved. The binding column was prepared by first being placed in a 2 mL collection tube and 500 µL of column preparation solution added. The tube and column were centrifuged for 1 minute. The flow-through liquid was then discarded. To the dissolved gel vial, 1 mL of 100% isopropanol (Sigma-Aldrich) was added and homogenised well. In 500 µL aliquots, the dissolved gel and isopropanol solution was loaded into the column and centrifuged for 1 minute, discarding the flow-through each time. 700 µL of the diluted wash solution was then added to the column and centrifuged for 1 minute. The flow through was then discarded. The column was then transferred to a clean collection tube and incubated with 20 µL of nuclease free water for 1 minute.

This was then centrifuged for 1 minute and the flow through transferred to a clean sample vial for storage. Concentration was then determined using UV-vis spectroscopy measurements, as described in section 2.2, and the final product stored at -20°C until use.

2.2.4.2 Streptavidin beads

SCoNE structures were extracted using Dynabeads™ MyOne™ Streptavidin C1 beads (ThermoFisher). Three assembly reaction vials were combined, and contents were added to 70 µL of Dynabeads™. This was incubated at room temperature on a shaker for 10 minutes. This was then placed in a magnetic vial holder and incubated on the benchtop for 3 minutes. The vial was removed from the magnetic holder, the supernatant discarded, and beads washed with 200 µL of 80% EtOH, and placed back on the shaker. The wash step was repeated 3 times. Following the third wash, the vial was opened, and the contents dried in air for 5 minutes. To elute the SCoNE DNA, 50 µL of nuclease free water was added and incubated on the shaker for 10 minutes. The vial was then incubated on the magnetic holder for 3 minutes and the supernatant collected and stored at -20°C prior to use.

2.2.5 ELISA

To determine the ability of the probe groups binding analytes of interest, standard, and modified ELISA experiments were used and developed, as illustrated in Figure 2.7. ELISA experiments were split into four separate categories: standard sandwich ELISA, direct biotin binding ELISA, counter protein ELISA, and missing protein ELISA. For the first, third, and fourth ELISAs, streptavidin (Sigma-Aldrich) was added in 3X excess to bind biotin extraction groups, to prevent nonspecific measurements. Assembled SCoNE structures were incubated at room temperature with human IL-6 (Sigma-

Aldrich), procalcitonin (Sigma-Aldrich), or both at 50 ng/µL for 20 minutes on a shaker. SCoNE structures were then re-isolated using Dynabeads™ as described in section 2.2.4.2, with 300 µL of nuclease free water (Merck) used for isolation. This was done to create enough volume to cover the ELISA well plate base. For non-streptavidin blocking experiments, SCoNE structures could not be isolated using this technique, therefore additional washing steps were added to the ELISA protocol.

For IL-6 and procalcitonin experiments, ELISA kits were obtained from Stratech (orb390920-BOR-96Tests) and Merck (RAB0037-1KT) respectively.

Protein standards were set up in a serial dilution from 300 pg/mL to 4.69 pg/mL using the sample dilution solution, provided in each ELISA kit, as the blank. To each well of a precoated 96 well plate, 100 µL of protein standard and/or isolated SCoNE structures were added, the wells sealed, and incubated at 37°C for 90 minutes. The cover was then removed, the supernatant discarded, the wells blotted onto paper towels, and then washed with 0.01 M phosphate buffer solution (PBS, pH 7.4) three times. For non-streptavidin blocking experiments, the plate was washed 5 times. After each wash step the wells were blotted onto paper towels. To each well, 100 µL of the biotinylated secondary antibody was added and incubated at 37°C for 60 minutes. The supernatant discarding and wash step was then repeated before addition of 100µL of an avidin-biotin-peroxidase (ABC) solution to each well and incubated at 37°C for 30 minutes. The wells were then washed using 0.01 M PBS five times, allowing the PBS to sit in the wells for 2 minutes on each wash prior to blotting on paper towels. To each well, 90 µL of TMB solution was then added and incubated at 37°C for 20 minutes. After incubation, 100 µL of TMB stop solution was added to each well, and the absorbance

read at 450 nm on a microplate reader (Infinite 200 PRO microplate reader, Tecan Trading AG, Switzerland).

To measure the ability of SCoNE structures to bind analytes of interest, modifications were made to standard ELISA methods. To measure a trimer SCoNE structure with its biotin groups blocked using streptavidin, a standard ELISA was used (ELISA A). Here SCoNE bound protein was bound to a primary IL-6 antibody (2.7.A.1). Secondary IL-6 antibody was then bound to the protein (2.7.A.2). ABC was then bound to the biotin labelled antibody (2.7.A.3), and TMB converted to a coloured compound for absorbance measurement (2.7.A.4). To measure the direct interaction with the biotin groups on the SCoNE structure, a direct biotin binding ELISA was developed, using a trimer SCoNE structure (ELISA B). The SCoNE bound protein was bound to a primary IL-6 antibody (2.7.B.1), but the secondary antibody was removed from the experimental procedure (2.7.B.2). ABC was then bound to the biotin labelled SCoNE structure (2.7.B.3), and TMB converted to a coloured compound for absorbance measurement (2.7.B.4). A modified ELISA using the tetramer SCoNE structure, where the primary antibody binds procalcitonin and the secondary binds IL-6, was developed as a counter protein ELISA (ELISA C). This was done to measure the ability of the SCoNE structures to capture multiple analytes. SCoNE bound protein was bound to a primary procalcitonin antibody (2.7.C.1). A secondary IL-6 antibody was added, binding to the IL-6 protein (2.7.C.2). ABC was then bound to the biotin labelled IL-6 antibody (2.7.C.3), and TMB converted to a coloured compound for absorbance measurement (2.7.C.4). Finally, to determine blocking of the biotin groups using streptavidin, specificity of the aptamer, and specificity of the secondary antibody, a missing protein ELISA was developed (ELISA D). A SCoNE structure was incubated

with procalcitonin only, leaving the IL-6 aptamer unbound. SCoNE bound protein was bound to primary procalcitonin antibody (2.7.D.1). A secondary IL-6 antibody was added but was unable to bind (2.7.D.2). Due to the lack of free biotin groups, the ABC cannot bind (2.7.D.3), and therefore with the addition of TMB, it was hypothesised that no colour change would occur (2.7.D.4).

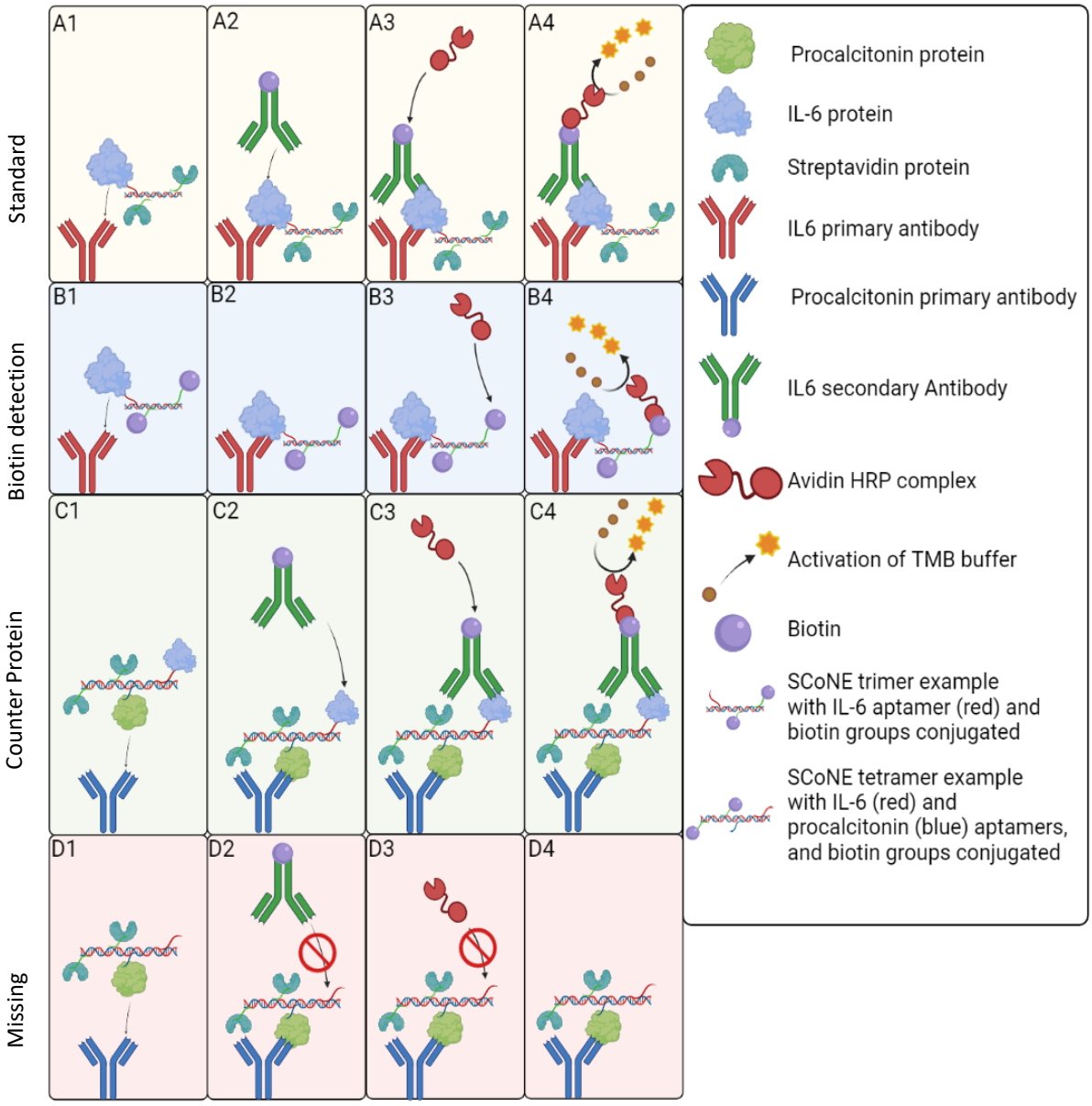

Figure 2.7. ELISA examples for standard sandwich ELISA (A.1-4), direct biotin detection ELISA (B.1-4), counter protein ELISA (C.1-4), and missing protein ELISA (D.1-4).

As protein concentration was determined through use of a single aptamer bound to a SCoNE structure, the concentration of SCoNE structures was calculated based on a 1:1 molarity ratio. This value was converted to pg/mL using the total molecular mass of the structure (2080103 g/mol for a trimer and 2090503 g/mol for a tetramer). This was then converted to ng/µL, the dilution factor applied, and the final values compared against those obtained from UV-vis spectroscopy measurements.

2.3 Results and discussion

2.3.1 Preparation of SCoNE structures

From initial results it was shown that it was possible to create SCoNE structures, overcoming the fragment size limitations of the Gibson assembly technique. Initial experiments were performed without the iDNA strands, as this section of the method was still in development during the initial experimental phase. Using gel electrophoresis for size determination was difficult due to the low concentrations used for assembly. Therefore, it was often necessary to overexpose the gels during UV visualisation to see the DNA bands, and dilute the standards used by 10X. This issue was later overcome through improvement of the isolation and preparation methods, increasing overall DNA concentration.

From the initial gel experiment conducted, it can be observed that, as illustrated in Figure 2.8, it was possible to assemble 2 sDNA and 2 pDNA fragments into a single structure. This was performed 3 times, and the image shown is representative of the results observed. As can be tentatively observed in lanes 5 and 8, highlighted in the yellow boxes of Figure 2.8, there is a mass shift to above the 2 kbp standard fragment. This is further confirmed by the DNA ladder in lane 1. Due to the addition of pDNA fragments to the 2 sDNA fragments an increase of approximately 100 bp causes retardation of the DNA movement. The pDNA fragments contributing 60 bp, and the remaining 40 bp from the aptamers. The folded nature of the aptamers can also contribute to slowing the movement through the gel, increasing the shift observed. SCoNE structures in lane 5 were assembled using a final concentration of 1pM.

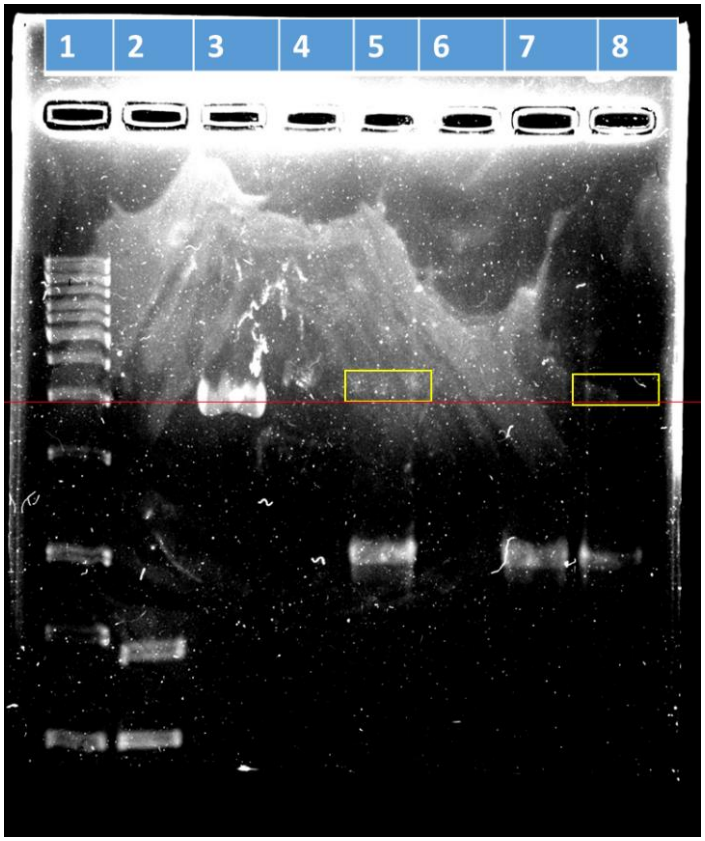

Figure 2.8. 1% agarose gel run at 75V for 45 minutes illustrating the ability to assemble, and extract dimer scone structures (yellow box) with lane 1 being the Gene Ruler 1kb, lane 2 a low range Gene Ruler, lane three a 2kbp DNA fragment, lane 4 being a short P1 strand, lane 5 a SCoNE assembly at 1pM, lane 6 a water blank, lane 7 a negative control containing only the starting materials prior to assembly, and lane 8 a SCoNE assembly at 5pM. The red line shows the 2kbp position across all lanes, and the yellow boxes show the formation of dimer SCoNE fragments.

SCoNE structures assembled in lane 8 used a final concentration of 5 pM to determine if more constructs could be assembled with more starting reagent. Due to the observation of a decrease in the intensity in lane 8, it was determined that increasing the concentration potentially hindered the formation of SCoNE structures. However due to the quality of the gel, this would be further explored.

Further SCoNE structures were assembled using increased number of fragments. Lanes 6 and 7 of Figure 2.9 highlights that it was possible to construct dimer, pentamer and decamer structures, highlighted in red boxes. This was performed 20 times, and the image shown is representative of the results. The successful construction of the SCoNE structures is further confirmed by comparison to the DNA standards (lanes 1, 2, and 3). Into the reaction vessel, 10 sDNA and 10 pDNA fragments were added, however depletion of the starting sDNA fragment, highlighted in the yellow box, indicates that the reaction was not able to fully complete. This was further confirmed by comparison to the negative control, lane 8,

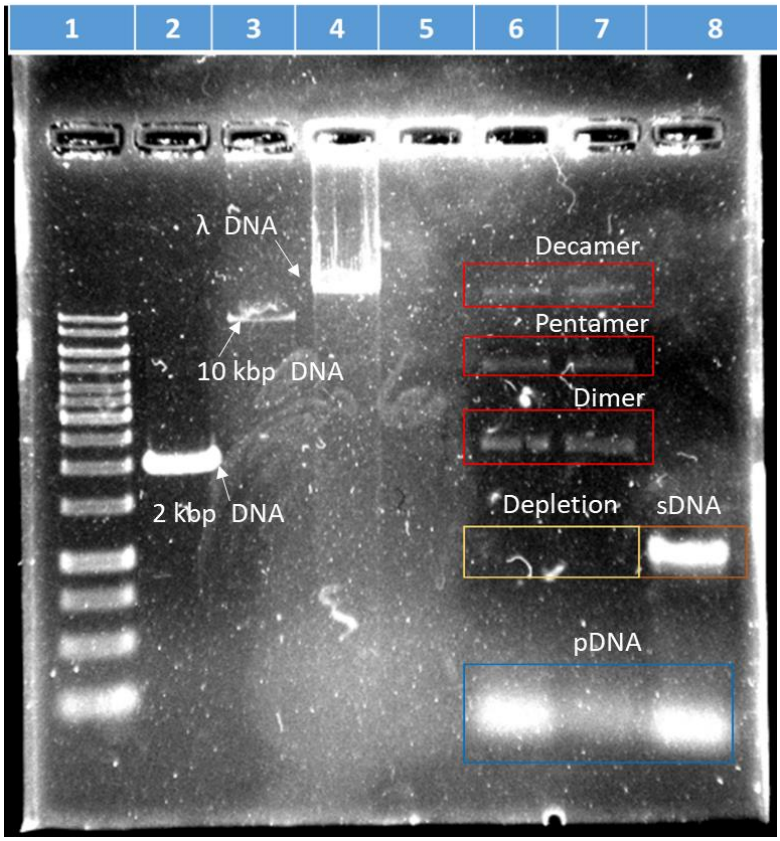

Figure 2.9. A 1% agarose gel showing the ability to create larger SCoNE structures. This highlights the ability to create decamer SCoNE structures (lanes 6 and 7). Lane 1 Gene Ruler 1 kbp, lane 2 2 kbp fragment, lane 3 10 kbp fragment, lane 4 whole λ DNA, lane 5 PCR amplified construct, lane 6 and lane 7, dimer, pentamer and decamer SCoNE structures at 1 pM (N= 1 and 2 respectively), and lane 8 negative control of only starting materials. Red boxes highlight the SCoNE structures created, the yellow box highlights the depletion of sDNA fragments, the orange box shows the starting sDNA fragments, and the blue box shows unincorporated pDNA fragments.

and the presence of the sDNA fragments (highlighted in orange). This is thought to be the reason for the creation of the multimer fragments, as opposed to a single decamer SCoNE structure. The remaining pDNA fragments, highlighted in the blue box, further reinforced the idea that it is the sDNA fragment concentration which needed to be optimised. Created SCoNE structures, were PCR amplified using primers only for the 5'-3' direction against pDNA fragment 1 and loaded into lane 5. From the lack of any band observed in the lane, it was determined that the presence of the probe strands prohibits the activity of the DNA polymerase amplifying the DNA. Thus, the maximum fragment size which could be produced would be equal to the distance between the 5' end and the probe conjugation site (20bp). It was also determined that the high temperatures required for PCR denatured the SCoNE fragments and the aptamer folding structures. Due to this, it is thought that the structures tangle, and are unable to reform their initial double stranded nature. It is also

thought that due to the probe sites being so close to the site of polymerase and ligase activity, correct annealing and repair to the backbone is not possible, making the structures unstable at high temperatures. Absence of fragments above 1 kbp in the negative control, further confirms the assembly of the SCoNE structures.

Further optimisation of the assembly protocol relating to DNA concentration, potential use of a different enzyme mix, and alteration of the probe binding group for increased stability and biocompatibility, was then explored.

2.3.2 Optimisation

2.3.2.1 DNA concentration

From the previous experiments conducted, it was determined that 1 pM starting concentration was the maximum that could be used with the Gibson Assembly master mix. To determine whether the assembly efficiency could be improved, a serial dilution experiment was performed from 1 pM to 0.125 pM of starting fragment concentration. From the results of the dimer assembly, as illustrated in Figure 2.10, decreasing the concentration to 0.5 pM did not decrease assembly efficiency. This was performed 3 times, and the image shown is representative of the results observed. It was also found that while starting fragment concentration decreased, the dimer concentration was unchanged through grey value comparison. Below this concentration it was not possible to observe any assembly of fragments. From this result, it was decided that fragment concentration, for the assembly step, could be kept below 1 pM without affecting the efficiency of the assembly.

However due to issues with imaging gels with low DNA concentrations, it was decided that the assembly reaction steps would be performed at 1 pM of total DNA concentration.

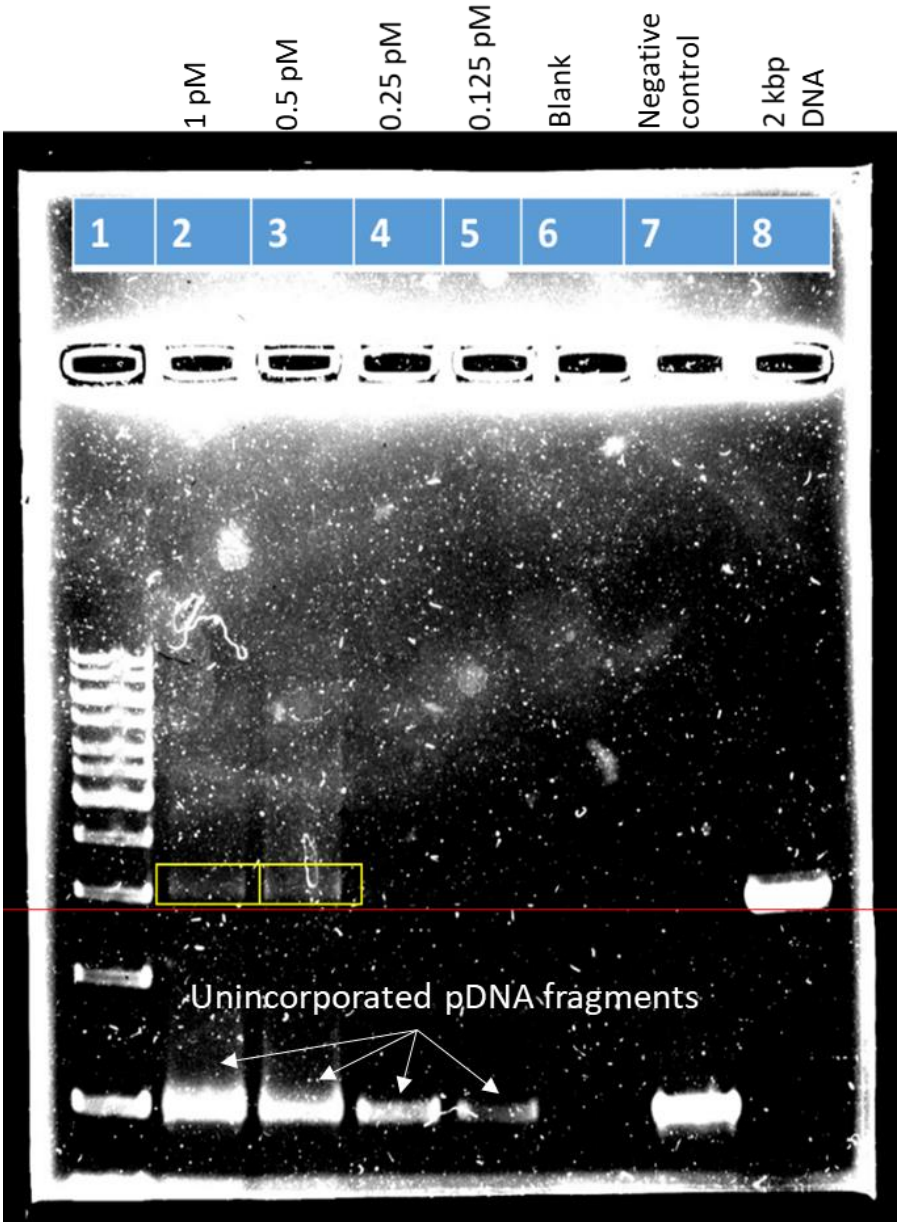

Figure 2.10. A 1% agarose gel showing the efficiency of SCoNE structure assembly with decreasing starting concentration. Lane 1 Gene ruler 1 kbp, lane 2 SCoNE assembly at 1 pM, lane 3 SCoNE assembly at 0.5 pM, lane 4 SCoNE assembly at 0.25 pM, lane 5 SCoNE assembly at 0.125 pM, lane 6 blank, lane 7 negative control of the starting material at 1 pM, lane 8 2 kbp DNA fragment. The red line shows the 2 kbp position across all lanes, and the yellow boxes show observable SCoNE structure formation.

2.3.2.2 Enzyme mixture

Due to the limitations and expense of using the Gibson Assembly master mix, it was decided that an enzyme mixture would be developed to overcome the limitations. The assembly protocol for this experiment was set up to produce decamer structures using 10 pDNA and 10 sDNA fragments. The assembly of the larger SCoNE fragment was used to determine if altering the enzyme mix used increased the assembly of larger fragments. From the results illustrated in Figure 2.11, it can be seen that in comparison to the Gibson assembly master mix (lanes 4 and 5), the in-house enzyme mix was both unable to assemble the fragments and led to the degradation of the starting reagents (lanes 6 and 7). The results from lanes 4 and 5 also highlight the increased reliability of the original assembly protocol. Depletion of the sDNA and pDNA fragments, highlighted in yellow boxes, indicate that all starting material was included in the assembly. This also indicated why there are no larger fragments observed. Due to the depletion, it is likely that significant concentrations of larger structures could not be assembled. Whilst it is possible to suggest that larger fragments were assembled due to previous results, it is not possible to confirm this assembly from this experiment. Further concentration balancing of the starting fragments is required to improve assembly. This was confirmed by comparison to the negative control, in lane 8, highlighted in blue boxes. The formation of dimer and tetramer structures are highlighted in lanes 4 and 5 by red boxes.

Unfortunately, due to time constraints, it was not possible to explore the development of an in-house enzyme mix further. Whilst the combined enzyme activity should perform similarly to the Gibson assembly master mix, the balancing of the concentration of enzyme units requires further exploration.

Figure 2.11. A 1% agarose gel showing the efficiency of SCoNE structure assembly with decreasing starting concentration. Lane 1 Gene ruler 1 kbp, lane 2 2 kbp DNA fragment, lane 3 10 kbp DNA fragment, lane 4 SCoNE assembly at 1 pM using Gibson assembly master mix (N=1), lane 5 SCoNE assembly at 1 pM using Gibson assembly master mix (N=2), lane 6 SCoNE assembly at 1pM using inhouse enzyme mix (N=1), lane 7 SCoNE assembly at 1 pM using inhouse enzyme mix (N=2), and lane 8 negative control of the starting material at 1 pM. The red boxes show the formation of dimer and tetramer SCoNE structures, the yellow and green boxes show the depletion of the sDNA and pDNA starting reagents using Gibson assembly mastermix and the in-house enzyme mix respectively, and the blue boxes show the starting sDNA and pDNA fragments in the negative control.

2.3.2.3 Binding of probe to the azidated DNA backbone

When storing samples and performing experiments with SCoNE DNA fragments, nuclease free water was used as the solvent. DBCO is not soluble in water, and therefore DMSO was used to prepare DBCO stock solutions. DMSO at high

concentrations can denature DNA, therefore it was decided to explore a water soluble DBCO construct, sulfo-DBCO. Whilst the overall conjugation reaction is the same, the stability of the linker between the aminated probe and the azidated DNA backbone was explored to determine if reaction yield was affected. The chemical structures are illustrated in Figure 2.12.

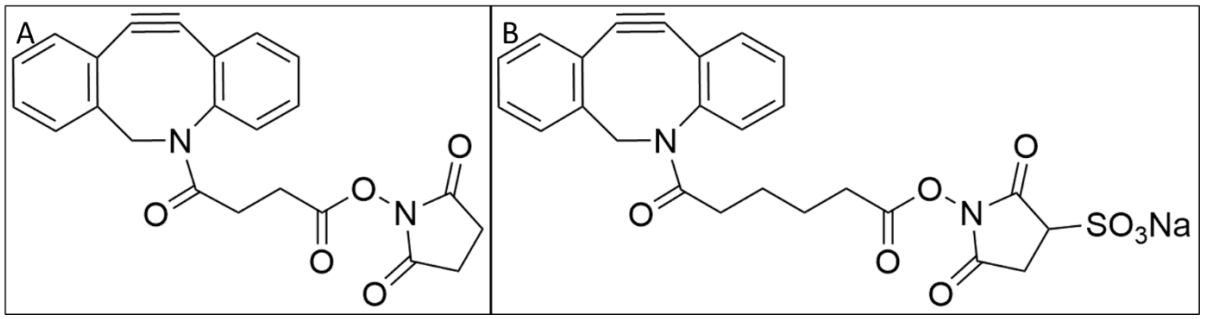

Figure 2.12. Chemical structures of DBCO (A) and sulfo-DBCO (B) NHS ester linker groups.

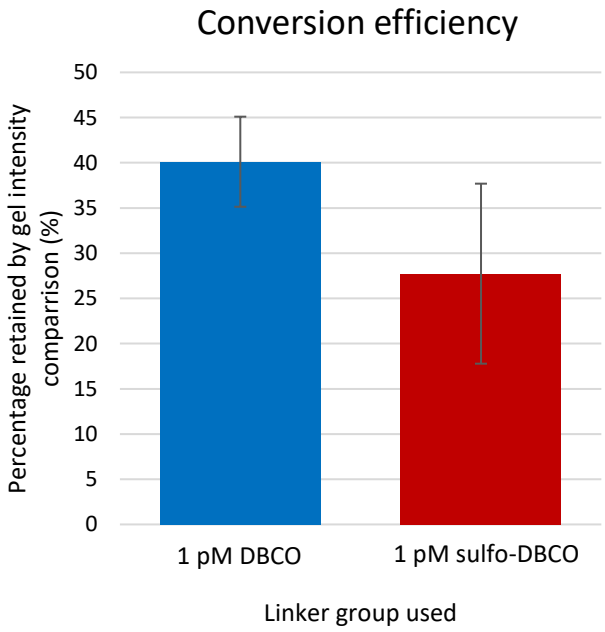

Figure 2.13. Percentage of pDNA retained when using DBCO and sulfo-DBCO as linker groups when compared to the starting unmodified pDNA fragment (N = 3, p = 0.28).

To determine the yield of pDNA formation, a gel experiment was performed comparing the intensity of the conjugated pDNA band (N=3) against the starting concentration of the short DNA fragment used to create the pDNA. From the results of the intensity profile, an average 40 ± 5 and 28 ± 10 % of the conjugated pDNA fragments were retained after modification using DBCO and sulfo-DBCO linker groups respectively, as shown in Figure 2.13. The presence of the negatively charged sulfonate group inhibits

interaction with the DNA, decreasing the conjugation efficiency. The use of high salt solutions could overcome this by electrostatically shielding the sulfonate group, however this could denature the aptamer structures. From statistical analysis (ttest) performed, a p value of 0.28 was obtained, indicating no significance between the groups. However, due to the average decrease in full pDNA formation when using the sulfo-DBCO linker group, it was decided that standard DBCO would be used for further experiments.

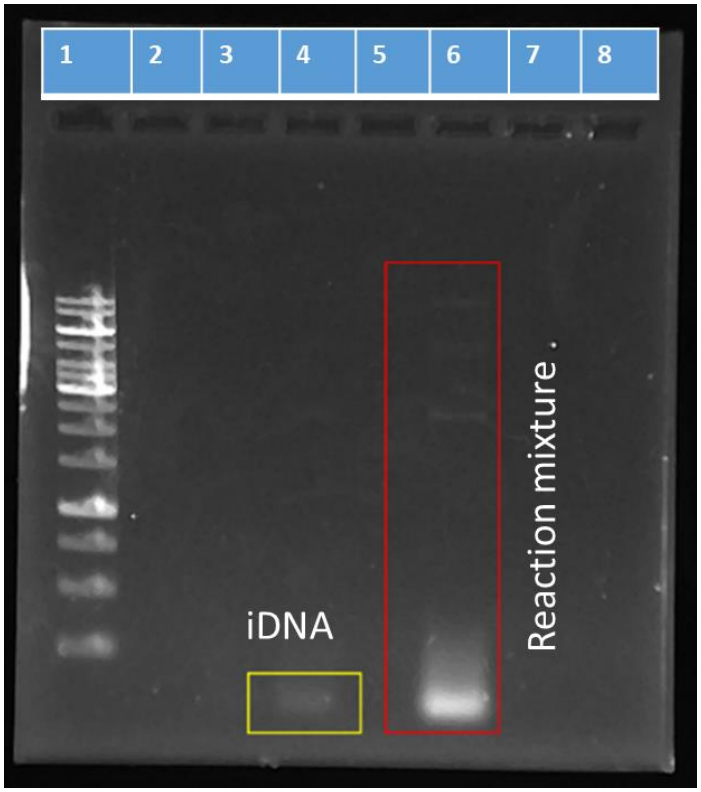

Figure 2.14. 1% agarose gel run at 75V for 45 minutes showing biotin labelled p strands (yellow box) can be extracted from the reaction mixture (red box) with lane 1, gene ruler, lane 4 biotin extracted p strands, lane 6 reaction mixture prior to assembly. All other lanes were left empty.

2.3.2.4 Biotin extraction

The development of the iDNA fragments, allowed for isolation of only biotin containing fragments. Fragments which did not include the specific iDNA fragment, could not be isolated, and were therefore discarded during the isolation process. Initially the ability to conjugate and test the stability of the iDNA strands was investigated. The experiment conducted compared a decamer assembly (~11 kbp) using only pDNA strands against a sample containing 8 pDNA and 2 iDNA but no sDNA fragments using the assembly method described in section 2.2.2.4. From the results of this experiment, it was possible to isolate the iDNA strands, highlighted in the yellow box Figure 2.14,

from the pDNA strands, highlighted in the red box Figure 2.14. This result also highlights that the pDNA and iDNA strands are not compatible for assembly due to the lack of higher weight bands in lane 4, Figure 2.14. The intensity between the low weight bands were compared in terms of relative intensity. The decrease in intensity between the assembly pDNA band and the iDNA band is consistent with the relative change in concentration between the starting fragments, and the isolate. Figure 2.14 also highlights the adaptable nature of the SCoNE assembly method with the ability to alter the probe groups, whilst retaining functionality. This experiment also confirmed that iDNA could not be digested by the T5 exonuclease.

Once it was confirmed that the biotin linker and Dynabeads™ extraction technique was compatible and could be integrated with SCoNE assembly, the position of the group within the structure was investigated. Biotin positioning analysis was performed to determine the effect on the achieved yield. The grey values were obtained from Figure 2.15, highlighted in the yellow box, and compared across all biotin group arrangements. The structures contained 2 sDNA fragments, either 1 or 2 pDNA fragments, and either 1 or 2 iDNA fragments. These were named di/trimers for the purposes of the experiment. These experiments were compared against 2 standard dimer experiments.

From the results of the experiment, it is possible to see that the assembly of di/trimers was successful, as illustrated in Figure 2.15, highlighted in the yellow box. When compared to the standard assembly reaction, it can be seen that the sDNA and pDNA

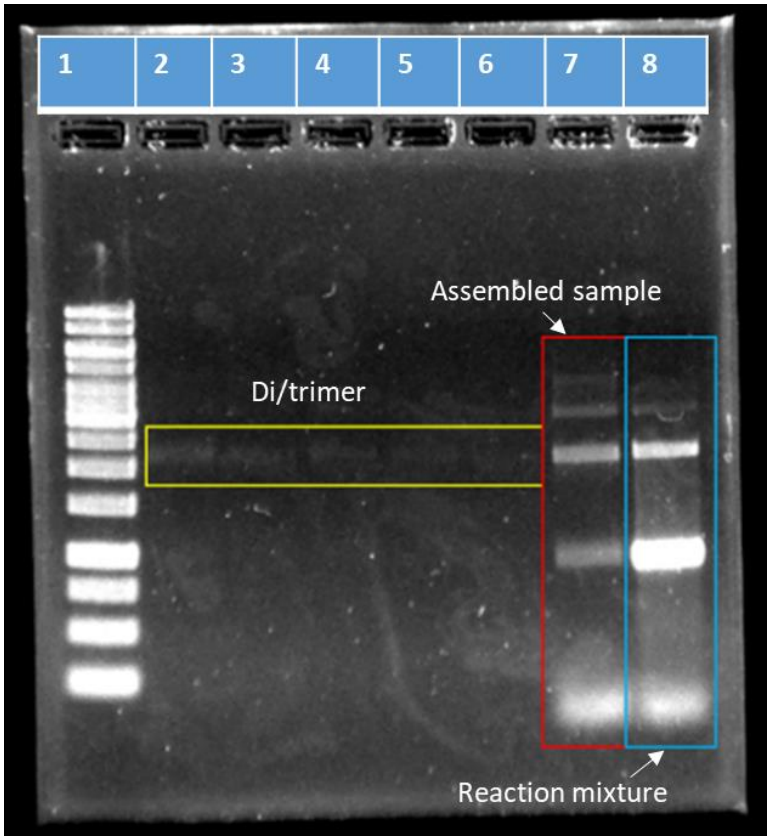

Figure 2.15. 1% agarose gel run at 75V for 45 minutes illustrating our ability to assemble and extract di/trimer scone structures (yellow box) from an assembled sample (red box). This also shows the need for an extraction method comparing to the starting reaction mixture (blue box). The bands visible above the dimer assembly (Red and Blue boxes) are plasmid fragments arising from incomplete SfiI digestion.

fragments are absent from the gel. This confirmed that the assembly is successful, and the biotin integration is vital for removal of unwanted fragments. The lack of bands above the fragments of interest further confirms that SCoNE assembly is a highly specific technique, only allowing for the designed fragments to assemble in an organised manner.

The analysis of the grey values suggests that the position of the biotin group affects the ability to isolate the SCoNE structures. This difference shows that biotin in positions 1 and 2 provides the highest efficiency in extraction, as shown in Figure 2.16. It is thought that the location of the biotin probe allows for multiple bindings to streptavidin during extraction. The relative closeness of the 2 biotin groups could allow for the DNA to bend, assisting the binding of the SCoNE structure. When comparing the intensities across the different structures, a slight

decrease of 2.60 ± 0.05%, 2.88 ± 0.05%, 1.21 ± 0.05%, and 2.48 ± 0.05% respective to 2S,1P_3,2$B_{1,2}$ can be observed. Due to time constraints, it was not possible to repeat this experiment, therefore this was performed to N = 1.

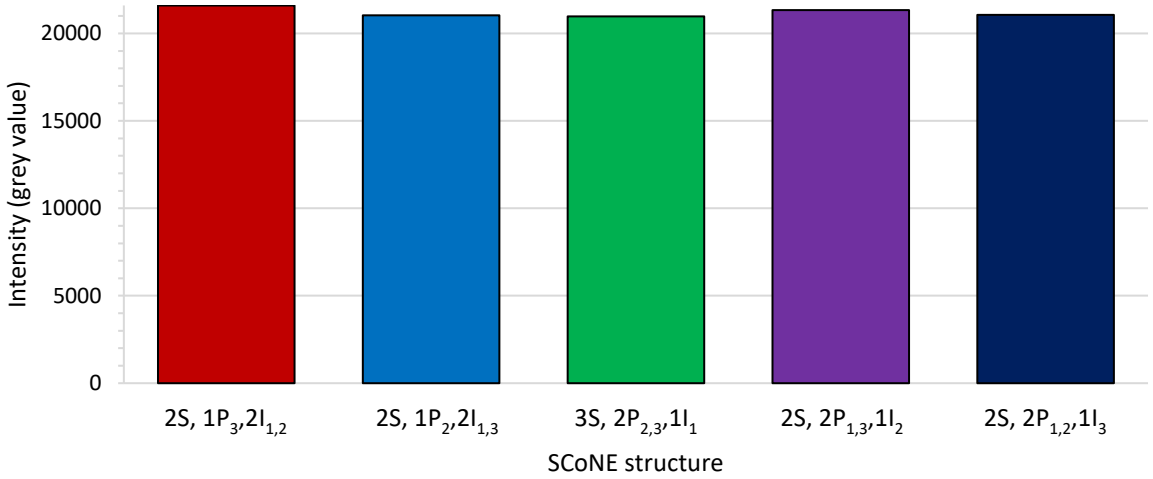

Figure 2.16. Respective intensity analysis compared to biotin group positioning where, from left to right, 1 pDNA fragment is in position 3 and 2 iDNA fragments are in positions 1 and 2, 1 pDNA fragment is in position 2 and 2 iDNA fragments are in positions 2 and 3, 2 pDNA fragments are in positions 2 and 3 and 1 iDNA fragment is in position 1, 2 pDNA fragments are in positions 1 and 3 and 1 iDNA fragment is in positions 2, and 2 pDNA fragments are in positions 1 and 2 and 1 iDNA fragment is in positions 3 with each pDNA and iDNA fragment separated by an sDNA strand.

2.3.3 SCoNE DNA full constructs

After initial optimisation, it was possible to produce and isolate SCoNE structures in a systematic and reliable manner. As is highlighted in Figure 2.17, it is possible to see the production of dimer (~2.2 kbp, lane 6), tetramer (~4.4 kbp, lane 7) and decamer structures (~11 kbp, lane 8). To perform this experiment, the gel was poured using half the typical amount of agarose gel to control gel casting and improve permeation of the gel dye. Figure 2.17, lane 5, also highlights that when performing the same assembly reaction as the decamer, but using unmodified pDNA strands, the assembly reaction does not take place, and no band is visible below the 1 kbp band. This indicated that the unmodified pDNA strand was unable to withstand the action of the T5' exonuclease. The pDNA was enzymatically digested to the point at which they could

not be observed in the gel. The result from this experiment further reinforces that the method developed is robust and adaptable.

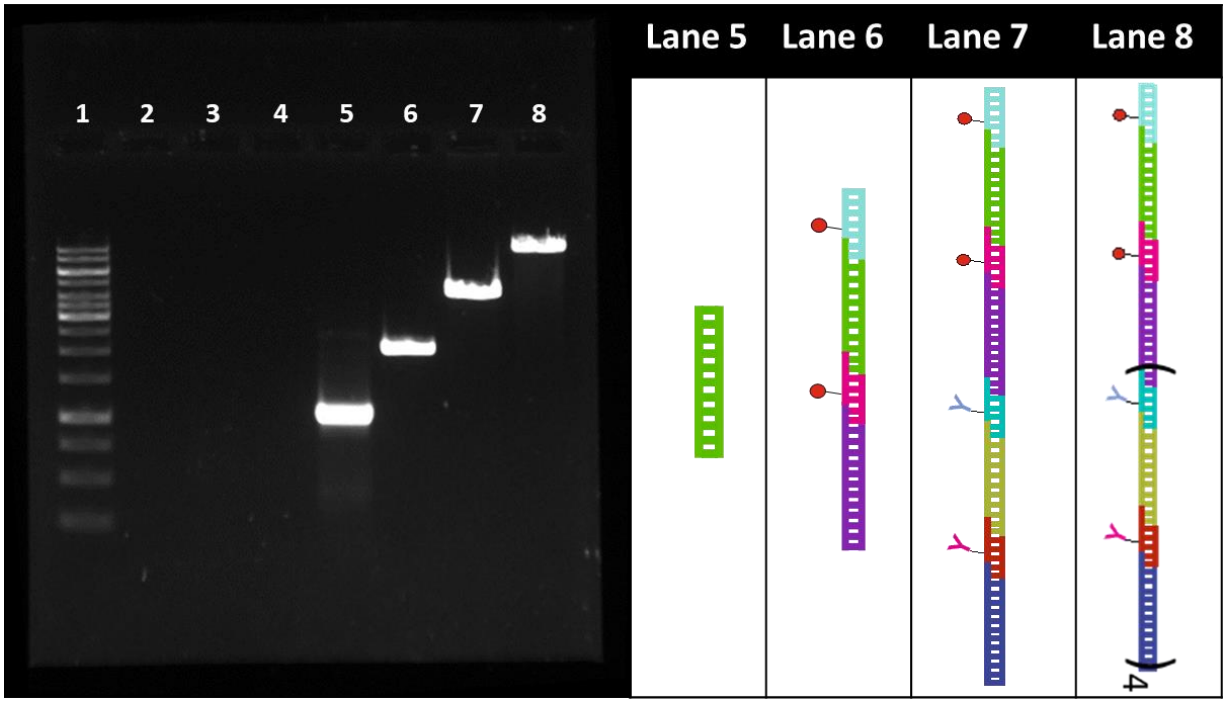

Figure 2.17. A 1% agarose gel run at 75 V for 45 minutes illustrating the ability to assemble different SCoNE structures, and that without modifications to the pDNA strands, assembly is not possible. Lane 1 Gene ruler 1 kbp, lane 2-4 blank, lane 5 Gibson assembly using unmodified pDNA, lane 6 dimer SCoNE assembly, lane 7 tetramer SCoNE assembly, lane 8 decamer SCoNE assembly. The formed structures are illustrated for lane 5-8 with red circles representing the biotin linker groups and the Y shapes representing aptamers.

To calculate the size of the DNA fragments, the distance migrated by each fragment was measured from the centre of the initial well to the centre of the DNA band. The DNA ladder was used to create a calibration curve, applying an exponential fit with a 95% confidence band, as shown in Figure 2.18. The equation of the line was then used to estimate the size of the SCoNE structures and the sDNA fragment. Unfortunately, due to time constraints, this experiment only had an N = 1.

From the results of this analysis, it was possible to estimate that the size of the fragments was 904 ± 198 bp for sDNA, 2005 ± 176 bp for the dimer structure, 4142 ± 211 bp for the tetramer structure, and 11128 ± 524bp for the decamer structure.

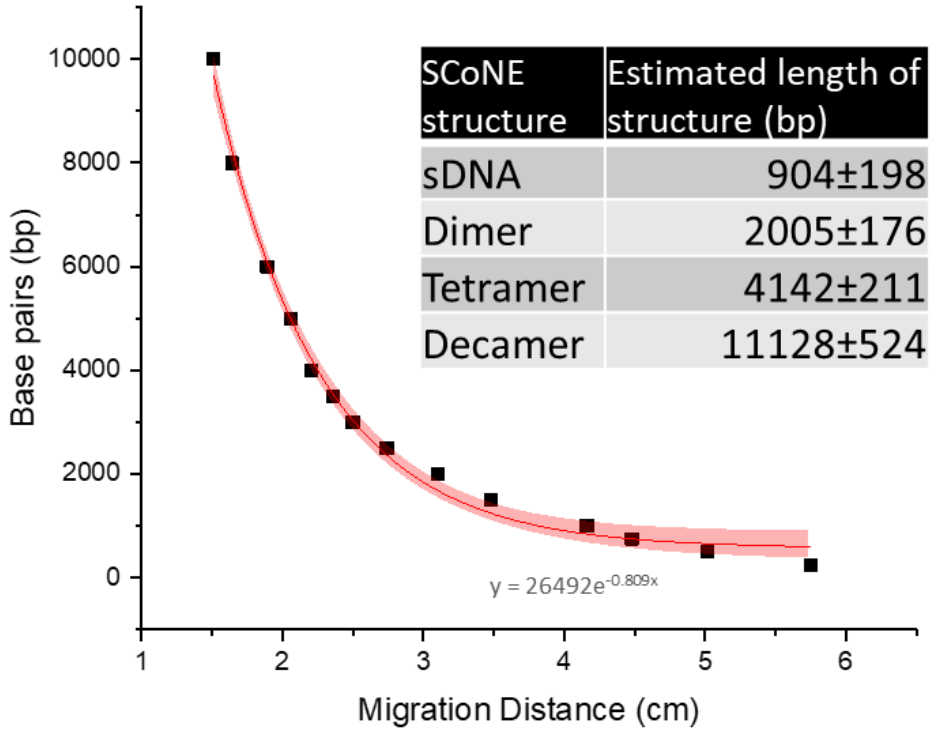

Figure 2.18. An illustration of the calibration curve calculated from the 1kbp GeneRuler ladder with an exponential fit line and 95% confidence intervals, and the resultant sDNA and SCoNE fragment size calculated from this fit, N = 1.

A comparison plot between the predicted bp length and the determined values was generated. A linear fit line was applied, and the gradient noted, as shown in Figure 2.19. Significance was determined using a one-way ANOVA. The slope was calculated to be 1.007 and the significance at 0.960. Both results indicate significant similarity between the predicted and measured values. This result signifies that the structures generated are of approximately the size predicted, further suggesting the successful assembly of SCoNE structures.

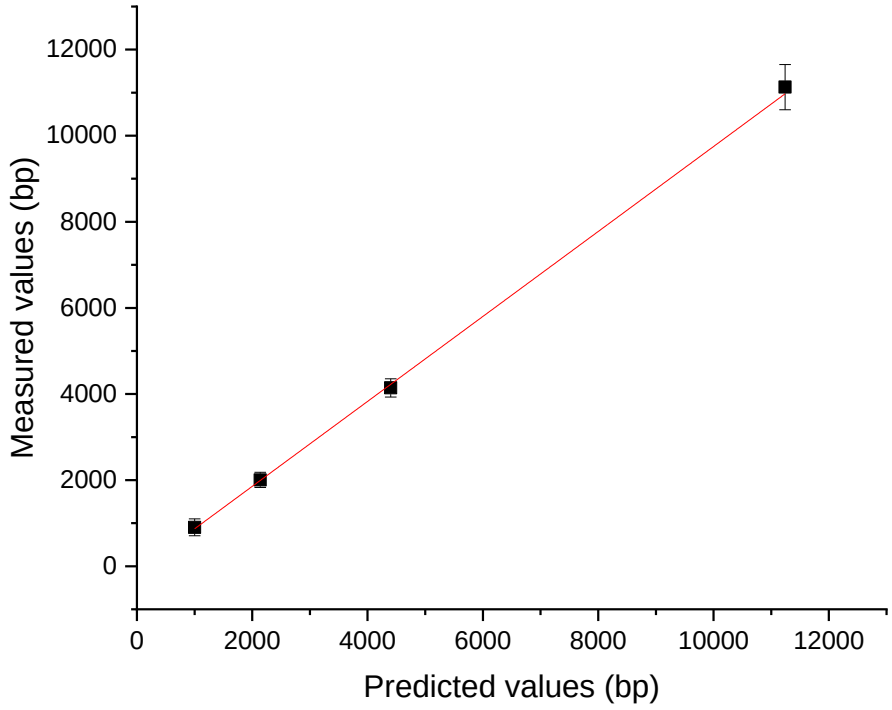

Figure 2.19. A comparison between the predicted and measured sizes of the SCoNE structures illustrating high similarity

2.3.4 ELISA results

To determine the binding ability of assembled SCoNE structures, sandwich and modified ELISA protocols were used. All experiments were completed to N=6. From the results of the ELISAs performed, it was possible to determine that SCoNE structures with functional capture groups could be successfully isolated. As previously described in section 2.2.5 4 different ELISA techniques were used to determine the protein binding capability of SCoNE structures. ELISA A represents a standard sandwich ELISA capturing IL-6, ELISA B removes the secondary antibody but uses the ABC complex directly for sensing captured IL-6, ELISA C captured both IL-6 and procalcitonin utilising the primary antibody for one protein and the secondary antibody for the other protein (this was performed using both IL-6 and procalcitonin as the primary protein in separate experiments), and ELISA D captured IL-6 and used a secondary antibody for procalcitonin. As is highlighted in Figure 2.20, ELISA A captured 443.1 ± 64.6 pg/mL of IL-6 using a trimer structure, ELISA B captured 419.3

± 22.6 pg/mL of IL-6 using a trimer structure, ELISA C using an IL-6 secondary measured 126.0 ± 33.1 pg/mL of IL-6 using a tetramer structure, ELISA C using a procalcitonin secondary captured 103.4 ± 18.9 pg/mL of procalcitonin using a tetramer structure, and ELISA D measured 2.8 ± 0.6 pg/mL when a tetramer structure was incubated with IL-6. Due to the signal generated when no antibody is present, it is possible to suggest that that not all biotin sites were fully occupied, therefore for protein detection analysis, an additional error of 2.8 pg/mL could be applied to all values obtained.

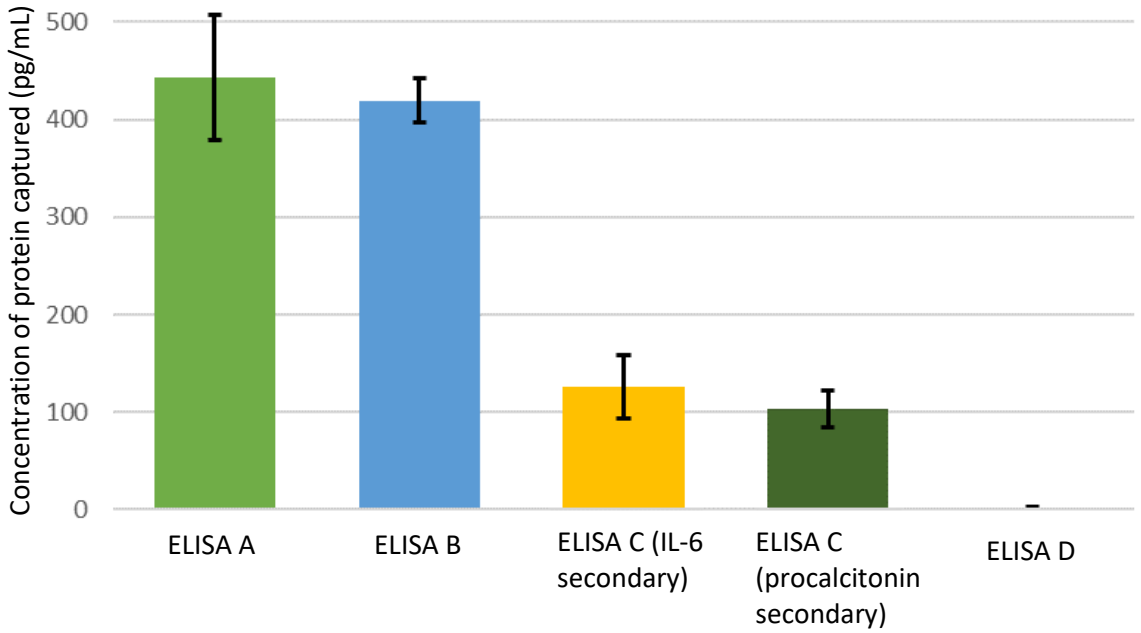

Figure 2.20. Representation of the concentration of protein captured by SCoNE structures for each type of ELISA performed (pg/mL).

As previously discussed in section 2.2, the final concentration of SCoNE structures was determined by UV-vis spectroscopy. For each of the samples, the concentrations of unbound SCoNE structures obtained were 15.82 ± 1 ng/µL for ELISA A and B, 7.58 ± 1 ng/µL for ELISA C (IL-6 secondary), 9.61 ± 1 ng/µL for ELISA C (procalcitonin secondary), and 9.61 ± 1 ng/µL ELISA D. A comparison between the UV-vis

spectroscopy values and the concentration determined from the ELISA results was then performed. The molarity of SCoNE structures to bound protein was assumed to be 1:1 due to the presence of only 1 binding region for 1 specific protein. Therefore, the protein concentration could be converted into SCoNE molarity, using the specific mass of each structure, for direct comparison with the results of the UV-vis spectroscopy results. From the conversion from protein concentration to SCoNE molarity, ELISA A had an experimental final concentration of 11.66 ± 1.70 ng/µL, ELISA B had 11.03 ± 0.60 ng/µL, ELISA C (IL-6 secondary) had 3.33 ± 1.10 ng/µL, ELISA C (procalcitonin secondary) had 2.73 ± 1.32 ng/µL, and ELISA D had 0.07 ± 0.02 ng/µL. The difference in molarity highlights a decrease in binding efficiency between the trimer and tetramer structures. As is illustrated in Figure 2.21, when comparing the relative UV-vis spectroscopy concentration against ELISA results, it is possible to see that when capturing and analysing a single analyte, there is a greater efficiency than when analysing a multiple analyte sample. This effect could be due to several reasons. Due to the SCoNE structure, the flexibility of the backbone could have affected the binding potential when adding the secondary antibody, leading to a lower yield observed. Although ELISA kits were matched as closely as possible for reaction conditions, the combination of different ELISA kit reagents could have affected binding conditions. Notably, the procalcitonin kit presented difficulties in optimisation, often with either the primary or secondary antibody not binding correctly, leaving control wells blank. Unfortunately, due to time constraints, further optimisation of the modified ELISA's was not possible. However, when considering the overall findings from the experiments, it is possible to conclude that assembled SCoNE structures are able to bind and detect multiple analytes from incubation with a single sample vial. This was

not previously possible using a single capture system. It is also possible to conclude that the preparation and isolation method developed does not cause unfolding or

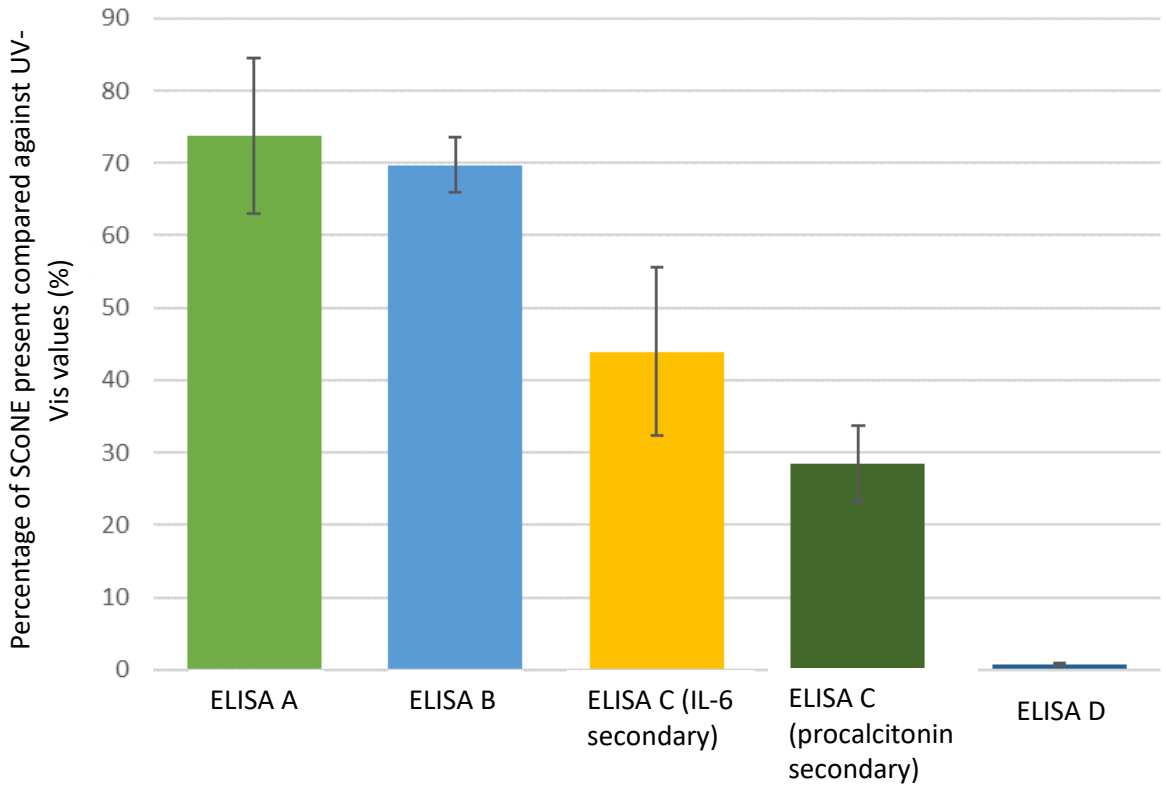

Figure 2.21. Percentage comparison between SCoNE concentration values obtained from ELISA and UV-vis spectroscopy results, separated via the type of ELISA performed.

degradation of the structures.

Whilst multiplex ELISA allows for multiple protein analysis, the combination of SCoNE and modified ELISA techniques has allowed for the direct detection of a protein based on the presence of a secondary protein in a single reaction vial, which has previously not been possible. This is further enhanced by the specificity of the aptamers to their respective proteins not allowing for nonspecific binding.

2.4 Conclusion

From the results of the SCoNE assembly and protein binding ELISA experiments, it is possible to conclude that the new method developed to construct DNA structures

capable of capturing analytes was successful. It is now possible to construct DNA fragments of less than 100 bp in a specific order with functional probes which was not previously possible, highlighting the novelty of this technique. Most significant is that it was possible to show that probes could be altered as necessary to the experiments without the need to significantly alter the overall protocol. This also allows for user design, where capture probes can be replaced dependent on sensing requirement. This indicates that a library of different structures with a variety of probes can be created and stored for any number of analyte-capture requirements. The development and addition of iDNA also provides additional post capture functionality for isolation or streptavidin linked sensing techniques. The development of new ELISA techniques further confirmed that SCoNE structures were capable of selectively binding specific proteins. ELISA results also highlight that for trimer structures approximately 70% of the expected SCoNE structures in solution were measured with captured proteins. This decreases to 45% when measuring tetramer structures with IL-6 secondary antibodies and 28% for the procalcitonin secondary antibodies. Due to issues faced when using the procalcitonin secondary antibody, this could account for the decreased yield. It is also possible to suggest that with the addition of more DNA fragments the shape of the DNA could inhibit secondary antibody binding, leading to the decreased yield observed.

2.5 References

Bell, N.A.W., and Keyser, U.F. (2015) Specific protein detection using designed DNA carriers and nanopores. J Am Chem Soc 137: 2035–2041.

Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002) Biochemistry. 9th ed., W H Freeman, 227-260.

Bestor, T.H. (2000) The DNA methyltransferases of mammals. Hum Mol Genet 9: 2395–2402.

BSGM (2021) List of Genetic Clinics - The British Society for Genetic Medicine. https://bsgm.org.uk/healthcare-professionals/list-of-genetic-clinics/. Accessed May 31, 2022.

Chehrehasa, F., Meedeniya, A.C.B., Dwyer, P., Abrahamsen, G., and Mackay-Sim, A. (2009) EdU, a new thymidine analogue for labelling proliferating cells in the nervous system. J Neurosci Methods 177: 122–130.

Couldrey, C., Brauning, R., Henderson, H. V., and McEwan, J.C. (2015) Genome-wide DNA methylation analysis: No evidence for stable hemimethylation in the sheep muscle genome. Anim Genet 46: 185–189.

Devaraj, N.K., and Finn, M.G. (2021) Introduction: Click Chemistry. Chem Rev 121: 6697–6698.

Dorsett, M., Westlund, B., and Schedl, T. (2009) METT-10, a putative methyltransferase, inhibits germ cell proliferative fate in Caenorhabditis elegans. Genetics 183: 233–247.

Dupont, C., Armant, D.R., and Brenner, C.A. (2009) Epigenetics: Definition, mechanisms and clinical perspective. Semin Reprod Med 27: 351–357.

Farrelly, L.A., Thompson, R.E., Zhao, S., Lepack, A.E., Lyu, Y., Bhanu, N. V., et al. (2019) Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3. Nature 567: 535–539.

Gibson, D.G., and Russello, S. (2010) How Gibson Assembly® is Changing Synthetic Biology | NEB. New Engl BioLabs https://www.neb.com/tools-and-resources/feature-articles/gibson-assembly-building-a-synthetic-biology-toolset. Accessed February 17, 2022.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6: 343–345.

Golla, J.P., Zhao, J., Mann, I.K., Sayeed, S.K., Mandal, A., Rose, R.B., and Vinson, C. (2014) Carboxylation of cytosine (5caC) in the CG dinucleotide in the E-box motif (CGCAG|GTG) increases binding of the Tcf3|Ascl1 helix-loop-helix heterodimer 10-fold. Biochem Biophys Res Commun 449: 248–255.

Gong, H., Holcomb, I., Ooi, A., Wang, X., Majonis, D., Unger, M.A., and Ramakrishnan, R. (2016) Simple Method to Prepare Oligonucleotide-Conjugated Antibodies and Its Application in Multiplex Protein Detection in Single Cells. Bioconjug Chem 27: 217–225.

Gu, R., Lamas, J., Rastogi, S.K., Li, X., Brittain, W., and Zauscher, S. (2015) Photocontrolled micellar aggregation of amphiphilic DNA-azobenzene conjugates. Colloids Surfaces B Biointerfaces 135: 126–132.

Harmel, R., and Fiedler, D. (2018) Features and regulation of non-enzymatic post-translational modifications. Nat Chem Biol 14: 244–252.

Kemmler, M., Sauer, U., Schleicher, E., Preininger, C., and Brandenburg, A. (2014) Biochip point-of-care device for sepsis diagnostics. Sensors Actuators, B Chem 192: 205–215.

Kharas, M.G., Janes, M.R., Scarfone, V.M., Lilly, M.B., Knight, Z.A., Shokat, K.M., and Fruman, D.A. (2008) Ablation of PI3K blocks BCR-ABL leukemogenesis in mice, and a dual PI3K/mTOR inhibitor prevents expansion of human BCR-ABL+ leukemia cells. J Clin Invest 118: 3038–3050.

Klungland, A., and Robertson, A.B. (2017) Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine. Free Radic Biol Med 107: 62–68.

Kolb, H.C., Finn, M.G., and Sharpless, K.B. (2001) Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew Chemie - Int Ed 40: 2004–2021.

Konstantinova, I.M., Tsimokha, A.S., and Mittenberg, A.G. (2008) Role of Proteasomes in Cellular Regulation. Int Rev Cell Mol Biol 267: 59–124.

Kumar, S., Chinnusamy, V., and Mohapatra, T. (2018) Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. Front Genet 9: 640.

Levine, R.L. (2002) Carbonyl modified proteins in cellular regulation, aging, and disease. Free Radic Biol Med 32: 790–796.

Liyanage, V.R.B., Jarmasz, J.S., Murugeshan, N., Bigio, M.R.D., Rastegar, M., and Davie, J.R. (2014) DNA Modifications: Function and Applications in Normal and Disease States. Biology (Basel) 3: 670–723.

Loh, A.Y.Y., Burgess, C.H., Tanase, D.A., Ferrari, G., McLachlan, M.A., Cass, A.E.G., and Albrecht, T. (2018) Electric Single-Molecule Hybridization Detector for Short DNA Fragments. Anal Chem 90: 14063–14071.

Mateus-Pinheiro, A., Pinto, L., and Sousa, N. (2011) Epigenetic (de)regulation of adult hippocampal neurogenesis: implications for depression. Clin Epigenetics 3: 5.

Moore, L.D., Le, T., and Fan, G. (2013) DNA methylation and its basic function. Neuropsychopharmacology 38: 23–38.

Nelson, M., and McClelland, M. (1987) The effect of site-specific methylation on restriction-modification enzymes. Nucleic Acids Res 15 Suppl: r219.

Nowak, S.J., and Corces, V.G. (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet 20: 214–220.

Olszewski, N.E., West, C.M., Sassi, S.O., and Hartweck, L.M. (2010) O-GlcNAc protein modification in plants: Evolution and function. Biochim Biophys Acta - Gen Subj 1800: 49–56.

Pellitero, M.A., Shaver, A., and Arroyo-Currás, N. (2020) Critical Review—Approaches for the Electrochemical Interrogation of DNA-Based Sensors: A Critical Review. J Electrochem Soc 167: 037529.

Plesa, C., Ruitenberg, J.W., Witteveen, M.J., and Dekker, C. (2015) Detection of individual proteins bound along DNA using solid-state nanopores. Nano Lett 15: 3153–3158.

Pujol, P., Barberis, M., Beer, P., Friedman, E., Piulats, J.M., Capoluongo, E.D., et al. (2021) Clinical practice guidelines for BRCA1 and BRCA2 genetic testing. Eur J Cancer 146: 30–47.

Schvartzman, J.M., Thompson, C.B., and Finley, L.W.S. (2018) Metabolic regulation of chromatin modifications and gene expression. J Cell Biol 217: 2247–2259.

Sze, J.Y.Y., Ivanov, A.P., Cass, A.E.G., and Edel, J.B. (2017) Single molecule multiplexed nanopore protein screening in human serum using aptamer modified DNA carriers. Nat Commun 8: 1–10.

Tiwari, A., and Gong, S. (2009) Electrochemical detection of a breast cancer susceptible gene using cDNA immobilized chitosan-co-polyaniline electrode. Talanta 77: 1217–1222.

Vuković, P., Peccatori, F.A., Massarotti, C., Miralles, M.S., Beketić-Orešković, L., and Lambertini, M. (2021) Preimplantation genetic testing for carriers of BRCA1/2 pathogenic variants. Crit Rev Oncol Hematol 157: 103201.

Wand, N.O., Smith, D.A., Wilkinson, A.A., Rushton, A.E., Busby, S.J.W., Styles, I.B., and Neely, R.K. (2019) DNA barcodes for rapid, whole genome, single-molecule analyses. Nucleic Acids Res 47: e68–e68.

Welsh, L., Maleszka, R., and Foret, S. (2017) Detecting rare asymmetrically methylated cytosines and decoding methylation patterns in the honeybee genome. R Soc Open Sci 4: 170248.

Wen, L., and Tang, F. (2014) Genomic distribution and possible functions of DNA hydroxymethylation in the brain. Genomics 104: 341–346.

Wilkinson, A.A., Jagu, E., Ubych, K., Coulthard, S., Rushton, A.E., Kennefick, J., et al. (2020) Site-Selective and Rewritable Labeling of DNA through Enzymatic, Reversible, and Click Chemistries. ACS Cent Sci 6: 525–534.

Woodcock, D.M., Lawler, C.B., Linsenmeyer, M.E., Doherty, J.P., and Warren, W.D. (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. J Biol Chem 272: 7810–7816.

Chapter 3

Nanopipette fabrication and optimisation

Synopsis: This chapter provides an introduction to nanopores and nanopore sensing technology with a discussion of the theory involved with nanopore preparation and electrical detection. One of the aims of this project was to produce nanopipettes capable of high-resolution electrical detection. This chapter discusses the methods for nanopipette preparation and optimisation used to produce nanopipettes for further experiments.

3.1 Introduction

3.1.1 Pipettes

Pipettes are a ubiquitous lab staple for volume measurement, however, the application to nanopipettes is novel. As the need for better imaging and cell manipulation increased, micropipettes were developed to allow for intracellular injections and patch clamp experiments. This led to improvements in transfection of cells with genetic information and the ability to study singular ion channels respectively (Morris *et al.*, 2010).

3.1.2 Nanopores

Nanopores are defined as pores or openings with a diameter of between 1 - 100 nm. Whilst nanopore systems can contain several pores, for this discussion, systems which use a single nanopore will be the primary focus. Single nanopore systems are split into three categories: biological nanopores, chip-based nanopores, and nanopipettes.

Biological pores are self-assembling protein scaffolds formed in cellular membranes. These pores are very reproducible and are stable across a variety of experimental conditions (Deamer *et al.*, 2016). The protein most associated with biological nanopores is α-haemolysin. This is a toxic protein produced by *Staphylococcus aureus* intended to induce osmotic stress to erythrocytes and cause them to burst (Chalmeau *et al.*, 2011). The structure of these pores is such that the pore centre is hollow with size ranging from 1.4 to 4.6 nm (Song *et al.*, 1996). Whilst wide enough at the opening for dsDNA to pass, the narrow middle ring (situated in the phospholipid bilayer) is too narrow. However, ssDNA is narrow enough to pass through the pore unheeded. It is due to this that these pores have been employed in next generation DNA sequencing techniques. These pores have been modified via mutagenesis to control analyte flow,

or enhance specificity (Maglia *et al.*, 2008; Deamer *et al.*, 2016). Mycobacterial porins have also been used for sequencing (Butler *et al.*, 2008; Derrington *et al.*, 2010). Whilst useful for sequencing, the permanent structure of the protein limits its adaptability for measuring other analytes.

Chip based nanopores, are formed in a membrane structure, typically from silicon nitride due to its resistance, thermal and dielectric properties (Li *et al.*, 2001; Storm *et al.*, 2003; Li *et al.*, 2003). Due to the synthetic nature of production, these pores have greater flexibility in size, being produced by high energy ion or electron beam drilling (Chen *et al.*, 2004). Chip-based nanopores have been used to study the translocation of many different biomolecules including proteins, glycans, ssDNA, and dsDNA (Fologea *et al.*, 2005; Iqbal *et al.*, 2007; Xue *et al.*, 2020). Chip-based nanopores are a highly useful tool for studying analyte translocation, however the production of these pores requires the use of a cleanroom and cannot be easily produced (Edel and Albrecht, 2013).

Nanopipettes are produced via mechanical pulling of quartz or borosilicate capillaries with openings of 1 -100 nm (Steinbock *et al.*, 2010; Zhang *et al.*, 2018). The advantage of using quartz for fabrication is that it reduces current leakage observed in borosilicate nanopipettes and can be worked to produce smaller pore sizes (Morris *et al.*, 2010; Wadhawan and Compton, 2013). Nanopipettes can be broken down into five main sections; the pore, the sensing region, the taper, the shoulder and the pipette bulk (Bulbul *et al.*, 2018). Nanopipettes are a cheap (~21p, World Precision Instruments) and easy to produce alternative to solid state and biological nanopores (Levis and Rae, 1993). Due to the mechanical production of nanopipettes, it is often difficult to reproduce pore size and taper length. It is also difficult to produce nanopores with a

diameter of less than 20 nm whilst minimising the taper length. This can affect the resolution during measurement (Wang *et al.*, 2019). Nanopipettes have become useful in many different fields, from delivering drugs to sensing and material deposition (3D printing) (Suryavanshi and Yu, 2007; McAllister *et al.*, 2018; Yu *et al.*, 2019).

It has been shown that direct interaction with a functionalised nanopipette (a nanopipette with an inner coating of a specific binding substrate) allows for specific antigen binding and measurement (Morris *et al.*, 2010). This method also provides the potential for concentration to be measured through current reduction over time. Modifying the inner surface using proteins, such as poly-L-lysine, has allowed for detection and quantification of protons, complementary proteins (via interaction) and DNA (Umehara *et al.*, 2006; Umehara *et al.*, 2009; Fu *et al.*, 2009). Specific aptamer modifications have allowed for the detection of immunoglobulins from patient blood (Ding *et al.*, 2009). Other aptamer modifications have allowed for the detection of bio-toxic compounds and quantified the amount present (Wang *et al.*, 2015). Finally, from using a glass nanopipette with a secondary carbon nanotube, it has been possible to inject 5 mM K^+ directly into the cells. This technique was then used to directly monitor live cell ion channels, as well as the ion channel response to pharmacological agents (Schrlau *et al.*, 2009).

Whilst nanopipettes are not currently used in general medicine, micropipettes are used frequently in *in-vitro* fertilisation. These pipettes are manufactured by heat pulling capillaries, using borosilicate glass as the starting material. These have pore openings of 12-14 µm to allow for the capture of sperm and puncturing of ova (Yaul *et al.*, 2008). There is much in development for using nanopipettes for single cell monitoring and cellular excretion study. It has been suggested that these advances could contribute

to greater understanding of disease at the cellular level and real time monitoring of changes in cellular processing (Feng *et al.*, 2020; Cao *et al.*, 2021).

3.1.3 Operating principles of a nanopore device

The resistive pulse sensing method is a technique used to detect the presence of an analyte passing through a pore using an electrical field in a conductive solution (Zhang *et al.*, 2009; Loh *et al.*, 2018; Albrecht, 2019; Feng *et al.*, 2020). This technique was invented by Wallace H. Coulter in 1953 and led to the development of the Coulter counter (Coulter, 1953). The principle of resistive pulse sensing, in the context of nanopore sensing, is that a nanopore is situated in an insulating membrane in between two electrolyte solution containing chambers. Each chamber contains an electrode, typically silver/silver chloride (Ag/AgCl) and is filled with a conductive solution made of uniformly distributed salt ions, such as lithium chloride (LiCl). Ag/AgCl electrodes are considered to be non-polarisable and Faradaic electrodes, meaning that there is no change in electrostatic potential at the electrode surface. When applying a potential difference (or voltage) to the electrodes, a redox reaction occurs, highlighted in equation 3.1.

$$Ag_{(s)} + Cl^-_{(aq)} \leftrightharpoons AgCl_{(s)} + e^- \qquad (3.1)$$

Silver is oxidised at the anode, reacting with a chloride ion in solution to produce silver chloride and an electron. At the cathode the reverse is occurring, whereby a chloride ion is liberated from the surface of the electrode via the addition of an electron. This in turn establishes a flow of chloride and corresponding cations through the electrolyte, generating a current. This flow can be measured by sensitive ammeters (Edel and Albrecht, 2013).

The conductive solution used to conduct nanopore experiments is often varied between research labs dependant on the intended purpose of the experiment. The most common salts used are KCl, NaCl, and LiCl. In research conducted by Kowalczyk *et al.*, the use of different salt solutions at 1 M concentrations and varying the concentration of LiCl on DNA translocation was investigated. It was found that by decreasing the size of the metal ion, the translocation time and event amplitude increases drastically. This was also observed when increasing the LiCl concentration from 1 M, to 2 M, then to 4 M. This research indicates that by using a high concentration of LiCl, the resolution of the translocation events increases significantly. The higher ionic strength of the lithium ion interacting with the DNA backbone more effectively masks the negative charge than potassium or sodium due to its smaller size. This reduces the overall charge of the DNA, decreasing the effect of the applied current to pull the DNA through the pore (Kowalczyk *et al.*, 2012).

In a system with no analyte present, when a voltage is applied, cations are drawn towards the negative electrode, and anions to the positive electrode, thus creating an ionic current. It is possible to utilize the difference in current measurements when applying different voltages across the pore to generate a current/ voltage trace, as illustrated in Figure 3.1. As voltage is increased, the current increases also. It is possible to use the slope of the line generated, at zero bias, to determine the pore conductance. This can then be further used to estimate the pore size as a function of the electrolyte solution conductivity (Liebes *et al.*, 2010; Frament and Dwyer, 2012). For nanopipette sizing, additional factors play a key role in determining pore size, including taper length of the nanopipette, capillary inner diameter, and electrolyte conductivity. This relationship is described in equation 3.2 where, G is the

conductance of the nanopipette, l is the taper length of the nanopipette (as measured using callipers and optical microscopy), Di is the inner diameter of the capillary (0.5 mm) and g(c) the conductivity of the electrolyte (4 M LiCl + 10% Tris-HCl EDTA (TE) determined to be 173 mS cm^{-1} (Loh, 2017)) (Frament and Dwyer, 2012; Steinbock *et al.*, 2012).

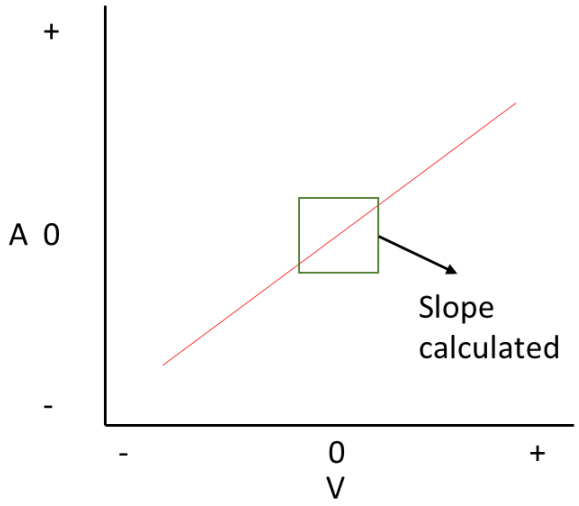

Figure 3.1. Simplified diagram of a conductance experiment where the red line represents the relationship between the voltage applied (X axis) and the current response (Y axis). The green box highlights the area of interest for pore conductance measurement.

$$d_{pore} = \frac{4Gl + \frac{\pi}{2}GD_i}{D_i\pi g(c) - \frac{\pi}{2}G} \qquad (3.2)$$

Analytes can then be added to one of the chambers and will diffuse and move through the solution via Brownian motion. When a bias is applied, charged analytes will tend to follow the electrical gradient towards the opposing charge, however in the bulk of the solution, the electrical field strength is low, so the analytes move primarily via diffusion. Due to the comparative difference in sizes between the electrode and the pore entrance/ exit, the electric field lines converge at the site of the pore, leading to the strong electric field near the pore. When the analyte nears the entrance of the pore, the strength of the electric field increases, directly pulling the analyte towards the pore (Muthukumar, 2007; Wanunu *et al.*, 2010). Once at the entrance, analytes then move from one side, through the pore, towards the other electrode, also referred to as electrophoretic migration, as illustrated in Figure

3.2. As the analytes pass through the pore, the electric field remains approximately constant, and the ionic current flow is reduced or blocked dependant on the structure of the analyte translocating (Kasianowicz *et al.*, 1996; Wanunu *et al.*, 2010; Shi *et al.*, 2017). This has been dubbed the volume exclusion model, describing the relationship between the volume occupied by the analyte as it passes through the pore and the resulting inhibition of ionic flow (DeBlois and Bean, 1970; Talaga and Li, 2009). This model was further refined to reflect the structure and charge of the analyte and the influence this has on ionic flow (Kim *et al.*, 2014; Si and Aksimentiev, 2017; Wilson *et al.*, 2019; Huo *et al.*, 2021). Due to nanopore resistance being in the mega-ohm to giga-ohm range, the addition of analytes to one chamber has a limited effect on current flow. Therefore, it is at the nanopore entrance where notable effect on current can be observed (Wanunu, 2012). By measuring the current across the channel, it is possible to determine the event frequency and the relative size of the analyte compared to a reference (Yang and Yamamoto, 2016).

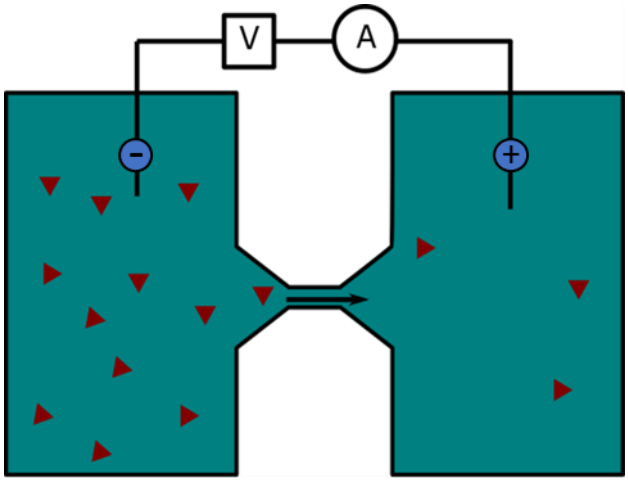

Figure 3.2. Simplified illustration of a nanopore sensing system where the triangles represent analytes, and the arrow represents analyte translocation direction.

In terms of nanopipettes, electroosmosis is the movement of liquids induced by the application of an electric field between the negatively charged surface and the positively charged ions in solution, along the surface wall. Using an example of a quartz nanopipette and LiCl as the salt solution, the negatively charged pipette surface is balanced by the lithium ions on its surface. This leads to an excess of cations at the

surface, in comparison to the bulk solution. With the application of an electric field, cations and anions in the solution will follow the potential gradient, taking their solvent shell with them, and resulting in the flow of liquid in the direction of the field (Edel and Albrecht, 2012; Gubbiotti *et al.*, 2022).

It is also important to consider the effect of the nanopore shape itself for translocation. In a nanopipette, the desired shape is conical, allowing for high resolution detection of the analyte translocated (Edel and Albrecht, 2012). The size of the pore is also vital for analyte translocation analysis. As pore size increases, the conductance increases, and the sensitivity of analyte detection decreases. With larger pores current is able to flow around the analyte, decreasing the impact on the current measurement (Kowalczyk *et al.*, 2011; Edel and Albrecht, 2012). The taper length also plays a key role in the sensing ability of a nanopipette. With increasing taper length often comes the elongation of the sensing region, requiring the analyte to travel further through the nanopipette before the current can recover. For analysis purposes, this can artificially increase the translocation duration, and inhibit any exploration of analyte structure (Yu *et al.*, 2019). It is therefore necessary when using nanopipettes for analyte sensing to construct them with small pores and short tapers to maximise the resolution of the data acquired.

3.1.4 Uses of nanopipettes

Many different instruments employ nanopores for imaging. Scanning probe microscopy is a broad term for a sub-category of microscopy involving scanning electrochemical microscopy, scanning ion-conductance microscopy, and near-field scanning optical microscopy (Morris *et al.*, 2010). Within these techniques, nanopipettes have been used as the probe in addition to a current feedback mechanism to determine various

electrochemical properties of an interface under investigation. The incorporation of nanopipettes into these techniques provides significant advantages compared with other measurement methods, such as biocompatibility and the potential for single molecule measurements. Using nanopipettes in contact-scanning electrochemical microscopy has been utilised to study ion transfer between two different ion containing solutions. This method has been used in biological investigations into ion-channel mediated transport of vital ions across membranes (Shevchuk *et al.*, 2001; Rothery *et al.*, 2003; Rodgers *et al.*, 2010).

3.2 Methodology

All chemicals used in the following procedures were obtained from Sigma-Aldrich® unless otherwise stated.

3.2.1 Electrolyte solution preparation

4 M LiCl solution was made using 16.96 g of LiCl anhydrous powder, 90 mL of ultra-pure water (MilliQ®, 18.2 MΩ) and 10 mL of TE (10mM Tris-HCl (pH 8.0), 0.1mM EDTA, ThermoFisher) buffer. This solution was then filtered through a 0.2 µm filter to remove any debris. This created the working 4 M LiCl 10% TE buffer.

3.2.2 Silver/ Silver chloride electrode preparation

Silver wire (0.25 mm, 99.99% purity, Goodfellow) was cut at a length of 8 cm and half submerged in 37.5% nitric acid for 20 seconds to oxidise. The wires were then removed and placed in a 4M LiCl solution and anodized at 1 mA for 20 minutes using a gold (99.99% purity, Goodfellow) counter electrode. These were then removed from the solution, washed with ultra-pure water (MilliQ®, 18.2 MΩ) and soldered to gold contact pins.

3.2.3 Nanopipette fabrication

A P2000 pipette puller (World Precision Instruments) was used alongside quartz capillaries (outer diameter: 1 mm, inner diameter: 0.5 mm. length: 7.5 cm, Sutter Instruments®, Novato, USA). Capillaries were cleaned using a plasma cleaner for 20 minutes at 0.8 Torr. There are five parameters to optimise on the P2000 before pulling can commence. These include “Heat” which determines the power of the laser, “Filament” represents the surface area of the quartz capillary that is heated by the laser. The “Velocity” value determines the speed at which the puller bar is moving, and

“Delay” states the time between the laser heating and the execution of the pull. The final parameter, “Pull”, indicates the force pulling the capillary apart. All units for the P2000 are represented as AU.

3.2.4 Optimisation of the pulling process

Optimisation of the pore size was achieved through the altering of several different parameters, including temperature and pull force of the instrument. Five capillaries were cleaned and pulled for each of the parameters highlighted in Table 5. The conductance of each were then measured and the pore size determined.

Table 5. Initial parameters for the production of 10 nm nanopores.

Pulling program	Pull	Heat	Filament	Velocity	Delay
99_1	200	800	4	20	100
	200	900	1	20	160
99_2	200	800	4	20	100
	160	900	1	20	160
99_3	200	800	4	20	100
	220	900	1	20	160
99_4	220	800	4	20	100
	200	900	1	20	160

3.2.5 Optimisation for taper length

Additional parameters were used during pipette preparation due to the need to minimise pipette taper length. Program 59 on the P2000 was used as defined by the manufacturer and modified to observe the effect on taper and pore size. Five capillaries were cleaned and pulled for each of the parameters highlighted in Table 6. The conductance of each were then measured and the pore size determined.

Table 6. Taper length optimised pipette pulling parameters

Pulling program	Pull	Heat	Filament	Velocity	Delay
59	75	700	5	35	150
	200	700	0	15	128
59_1 Heat modified	75	575	5	35	150
	200	900	0	15	128
59_2 Delay modified	75	700	5	35	145
	200	700	0	15	128

3.2.6 Conductance experiments

To prepare the liquid cell, a 3 mL brown glass sample vial was first filled with 2 mL of ethanol (HPLC grade, Sigma-Aldrich) and sonicated for 10 minutes, then the solvent discarded. The vial was then filled with 2 mL of ultra-pure water (MilliQ®, 18.2 MΩ) and sonicated again for 10 minutes, then the solvent discarded. The plastic lid of the vial was drilled twice using a 0.5 mm drill bit to allow for the nanopipette to be inserted, along with the electrode. The inner rubber sealing cap was used to hold the pipette stationary during experiments. The liquid cell was then filled with 2 ml of the 4 M LiCl 10% TE solution. The pipette was backfilled using the 4 M LiCl 10% TE solution and a Microfil™ needle, ensuring no air bubbles were present. To ensure complete filling of the pipette, the tip was submerged in the vial whilst backfilling. The structure of the setup is illustrated in Figure 3.3.

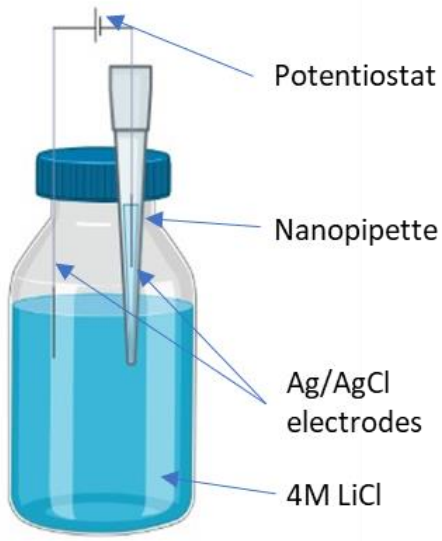

Figure 3.3. Simplified diagram of the conductance experimental set up.

The electrodes were connected to a potentiostat (CompactStat, Ivium) with one electrode as the working electrode (inside the pipette) and one as the counter electrode (in the solution), all within a Faraday cage.

The software controls were set up with a 4 mV scan rate, 5 scans and a range of -0.5 to 0.5 V. The scan files were then imported into MATLAB. A custom written script was created to extract the average gradient of the line (conductance). This value was than applied to equation 3.2 to calculate the size of the nanopore. The analysis code is available in Appendix 3.

Microscopy images of pipette tips were taken using a Nikon Type 105 optical microscope, with a 4X objective lens. The inbuilt digital measuring tape was used to determine pipette taper length, measuring from shoulder (the point at which the bulk pipette begins to narrow) to observable tip, as illustrated in figure 3.4.

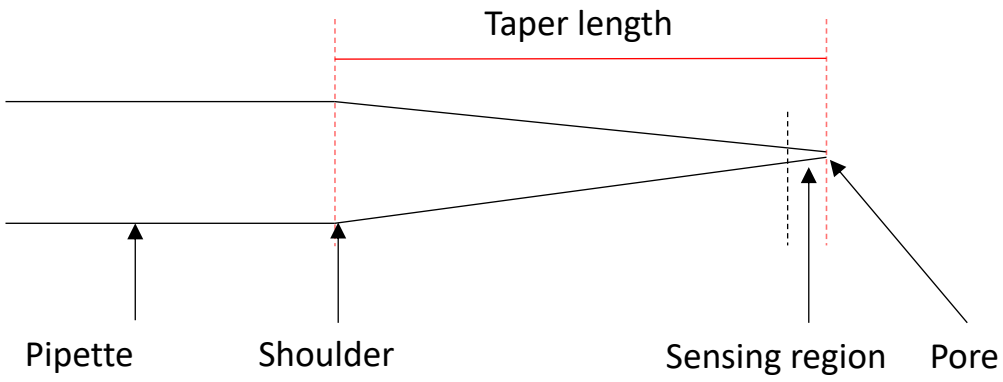

Figure 3.4. Schematic of a nanopipette for taper length measurement.

It was not possible during this work to image pipettes using scanning or transmission electron microscopy, however previous work has confirmed the validity of the model used, equation 3.2, for determining nanopore sizes (Yue *et al.*, 2017).

3.3 Results and discussion

3.3.1 Characterisation of primary pipette pulling parameters

Initial parameters were chosen and adapted from previous work conducted within the Albrecht group. Previous work had been able to generate pores of approximately 20 nm (Fraccari *et al.*, 2016; Yue *et al.*, 2017; Loh *et al.*, 2018). Due to the ability of quartz to remain stable under high heat and large pull force as discussed in section 3.1.2, it was decided that altering the pull force could have a significant effect on pore size. Starting parameters used a pull of 200/200 prior to optimisation, and statistical significance determined using One-Way ANOVA. It was found that decreasing secondary pull increased pore size by 320% from ~11 ± 3 to 35 ± 3 nm (p = <0.001) and decreased taper length by 34% from 6433 ± 301 to 4265 ± 244 µm (p = <0.001). Increasing the initial pull force decreased pore size on average 60% from ~11 ± 3 to ~6 ± 2 nm (p = 0.10) and decreased the taper length by 6% from 6433 ± 301 to 6079 ± 167 µm (p = 0.05). Increasing the secondary pull force increased pore size 35% from ~11 ± 3 to 15 ± 4 nm (p = 0.01) and increased taper length by 2% from 6433 ± 301 to 6531 ± 309 µm (p = 0.57), as illustrated in Figure 3.5.

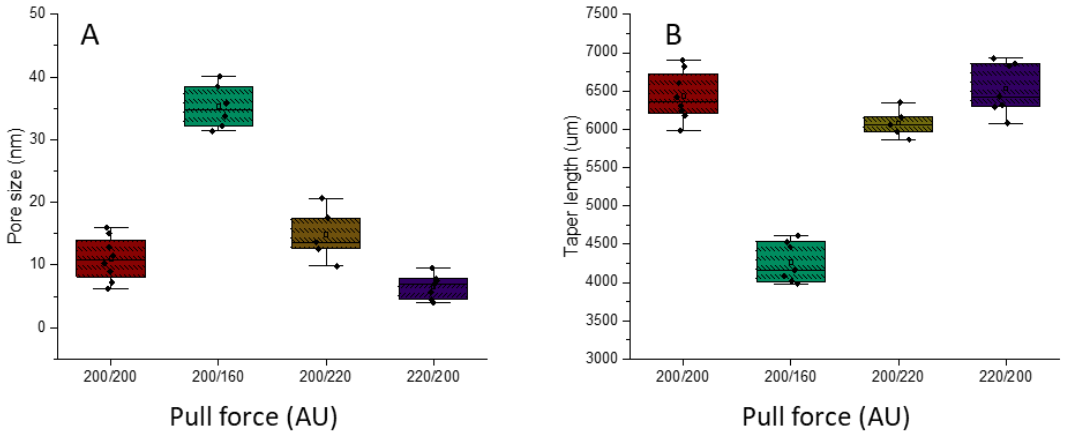

Figure 3.5. Effect of changing pulling parameters on pore diameter (A) and taper length (B) showing a decrease in pore size and an increase in taper length with increased pull force (N=8, 6, 5, and 8 respective to pull parameter used).

The effect of changing pull force also altered the number of useable pipettes. These were determined by observable breakages during microscopy measurements, and whether a current could be measured during characterisation. For the four groups measured, the number of broken pipettes observed was 2, 4, 5, and 2 of a total 10 repeats, respective to Figure 3.5. This suggests a correlation between altering the current recommended pulling parameters and an increase in the likelihood of pore failure. Increasing the secondary pull over 200 AU seems to increase the likelihood of failed pore formation.

The effect of altering the pull force can be visually observed in representative microscope images, as illustrated in Figure 3.6. Due to the length of the pipette tapers, it was necessary to join three images taken along the pipette taper. It can be seen that when primary pull force is increased (from 200 to 220 AU), the taper increases in length and decreases in diameter. Whilst this decreases overall pore size, the size of the sensing region increases, limiting the potential resolution of data acquisition for DNA translocation.

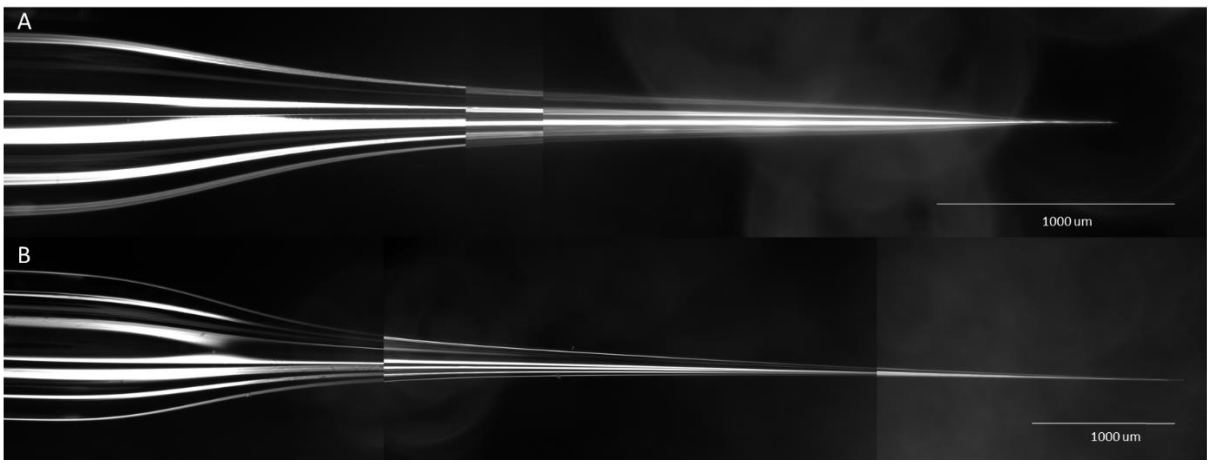

Figure 3.6. Comparison between pipette taper shapes under less primary pull force (A) and a greater primary pull force (B) as described in Figure 30.

From the results obtained it can be observed that increasing the initial pull decreases pore size, however the secondary pull has the most significant impact on pore size, taper length, and taper shape. While the initial pull creates the bulk of the taper, the secondary pull is responsible for the final formation of the pore. With a reduced secondary pull, the taper cannot be pulled as far, increasing pore size, but decreasing taper length.

It was found during further experiments that the ambient temperature of the laboratory had an effect on pore diameter. To achieve the required size several different pulling parameters were used dependant on the laboratory temperature, as highlighted in Table 7.

Table 7. Alteration of pulling conditions based on laboratory temperature.

Pulling Conditions	Pull	Heat	Filament	Velocity	Delay
<18 °C Lab temperature	220	800	4	20	100
	200	900	1	20	160
18-20 °C Lab temperature	200	800	4	20	100
	200	900	1	20	160
>20 °C Lab temperature	190	800	4	20	100
	200	900	1	20	160

Due to the long taper length observed in initial parameters, it was necessary to further optimise pipette pulling parameters to reduce taper length. Initial parameters used were those of program 59, as defined in section 3.2.5, Table 6. Variation on temperature and pull delay were applied to observe the effect on the pore size and taper length. Higher temperatures allow for a more fluid quartz capillary at the stretching point. Decreasing the delay parameter allows the capillary to cool less before the pull is applied. Both conditions should allow a smaller pore at reduced taper

length to form. Statistical analysis was performed using a One-Way ANOVA. As is highlighted in Figure 3.7, it was found that reducing the initial heat, and increasing the secondary heat reduced the pore size significantly by 1114% from ~18 ± 8 to 2 ± 1 nm ($p = <0.001$). Taper length, on average, increased by 4.08% from 3064 ± 431 to 3189 ± 63 µm ($p = 0.60$), however this result is not significant. When reducing the initial delay parameter, it was found that the pore size decreased by 400% from ~18 ± 8 to 5 ± 1 nm ($p = 0.01$) and taper length decreased by 9% from 3064 ± 431 to 2799 ± 105 µm ($p = 0.29$).

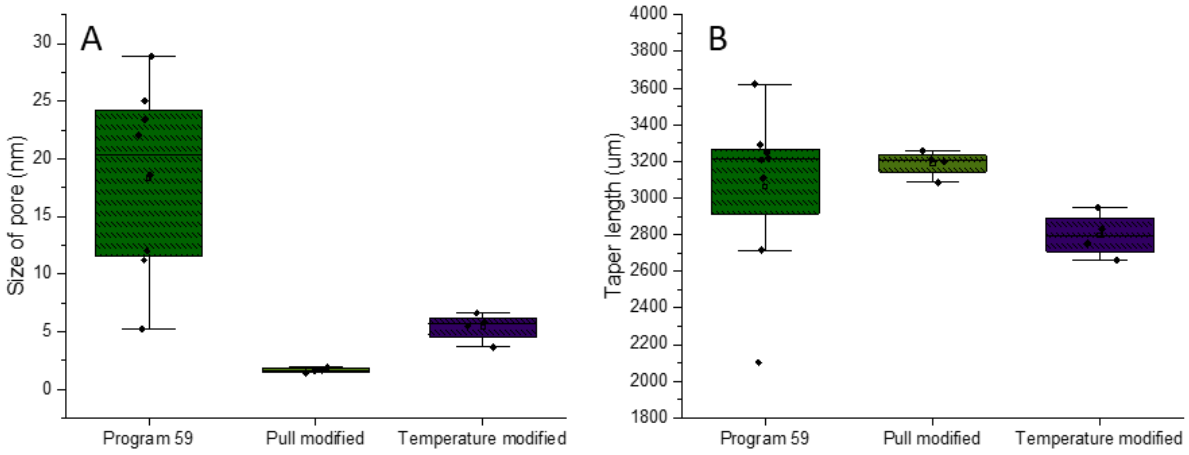

Figure 3.7. Effect of changing temperature and delay parameters on pore diameter (A) and taper length (B) (N=8, 4, and 4 respective to parameter alteration).

The effect on changing pulling parameter conditions also altered the number of useable pipettes. For the three conditions measured, the number of "failed" pipettes was 2, 6, and 6 from 10 samples, respective to Figure 3.7. When altering temperature, the number of "blocked" pores increased. This indicates that the higher secondary temperature potentially caused the pores to close after the final heat and pull. Decreasing the primary delay also increased the number of broken pores, indicating that altering the delay and allowing the quartz to cool more prior to the pull increased the fragility of the pores to manipulation.

3.3.2 Nanopipette characterisation for further experiments

It was determined that program 59 would be used for further experiments due to the size range of pores produced, the taper length, and the reduced frequency of pore failure. An example of a pipette used for further experiments is illustrated in Figure 3.8.A with the corresponding conductance experiment (3.8.B) and a linear line of best fit applied to the average conductance measurements to extract the conductance value for pore size determination (3.8.C).

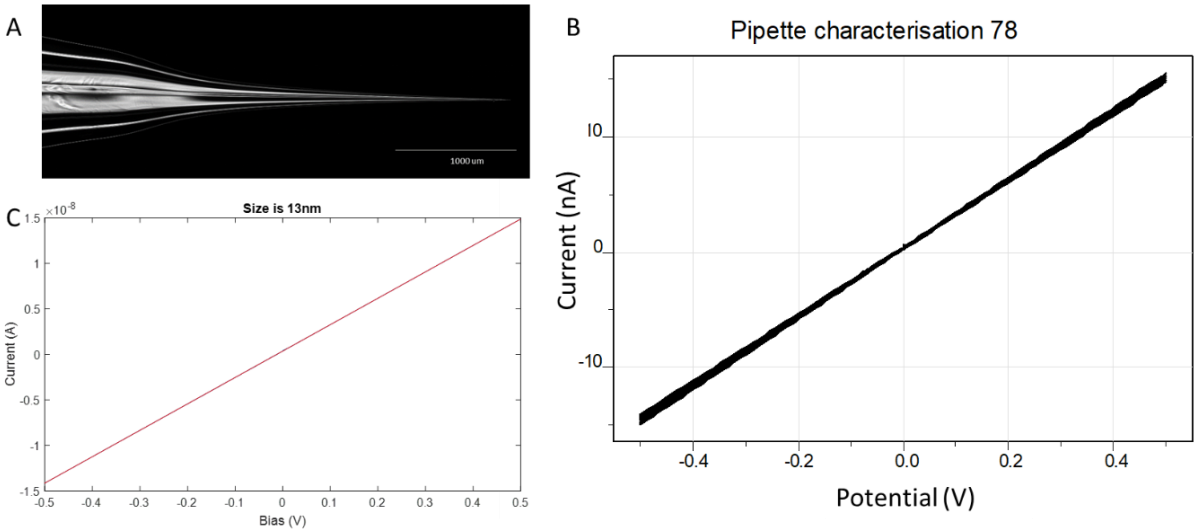

Figure 3.8. A nanopipette pulled using program 59 with a taper length of 3208 ± 0.005 µm (A) and a pore size of ~10 ± 1 nm determined by applying a linear line of best fit (C) to the average of the conductance measurements (B).

3.4 Conclusion

It is possible to conclude that a series of pulling parameters to synthesise nanopipettes with pores sizing greater than 30 nm, less than 30 nm, and less than 10 nm has been developed. When preparing nanopipettes, significant differences in pore size are observed through slight modification of pulling conditions. First, alteration of the primary pull parameter allowed for a great reduction in pore size, however drastically increased taper length. Secondly, decreasing secondary pull increased pore size and

decreased taper length. Thirdly, decreasing primary pull temperature and increasing secondary pulling temperature decreased pore size to less than 5 nm, however this had little effect on taper length. Finally decreasing primary pull decreased pore size below 10 nm and decreased taper length. It can be concluded that further optimisation of these parameters could ensure pore size reproducibility whilst minimising taper length. From the results of this optimisation, it was possible to identify the parameters necessary to conduct translocation experiments. Through observation of environmental conditions, it was also possible to note that changing laboratory conditions altered the ability to reliably pull pipettes. This was then incorporated into the parameters used to conduct further experiments.

3.5 References

Albrecht, T. (2019) Single-Molecule Analysis with Solid-State Nanopores. Annu Rev Anal Chem 12: 371–387.

Bulbul, G., Chaves, G., Olivier, J., Ozel, R., and Pourmand, N. (2018) Nanopipettes as Monitoring Probes for the Single Living Cell: State of the Art and Future Directions in Molecular Biology. Cells 7: 55.

Butler, T.Z., Pavlenok, M., Derrington, I.M., Niederweis, M., and Gundlach, J.H. (2008) Single-molecule DNA detection with an engineered MspA protein nanopore. Proc Natl Acad Sci U S A 105: 20647–20652.

Cao, Y., Yu, Y.J., Jiang, Q.Y., Sun, Y., Mao, Z., Wang, J., and Chen, F. (2021) The Advance of Plasmonic-Electric Nanopipette Sensing in Single Cells. Curr Pharmacol Reports 7: 55–66.

Chalmeau, J., Monina, N., Shin, J., Vieu, C., and Noireaux, V. (2011) α-Hemolysin pore formation into a supported phospholipid bilayer using cell-free expression. Biochim Biophys Acta - Biomembr 1808: 271–278.

Chen, P., Gu, J., Brandin, E., Kim, Y.R., Wang, Q., and Branton, D. (2004) Probing single DNA molecule transport using fabricated nanopores. Nano Lett 4: 2293–2298.

Coulter, W.H. (1953) Means for counting particles suspended in a fluid. US Patent, 2656508. United States Pat Off Patentiert am 20: 1953.

Deamer, D., Akeson, M., and Branton, D. (2016) Three decades of nanopore sequencing. Nat Biotechnol 34: 518–524.

DeBlois, R.W., and Bean, C.P. (1970) Counting and sizing of submicron particles by the resistive pulse technique. Rev Sci Instrum 41: 909–916.

Derrington, I.M., Butler, T.Z., Collins, M.D., Manrao, E., Pavlenok, M., Niederweis, M., and Gundlach, J.H. (2010) Nanopore DNA sequencing with MspA. Proc Natl Acad Sci U S A 107: 16060–16065.

Ding, S., Gao, C., and Gu, L.-Q. (2009) Capturing Single Molecules of Immunoglobulin and Ricin with an Aptamer-Encoded Glass Nanopore. Anal Chem 81: 6649–6655.

Edel, J., and Albrecht, T. (2012) Nanopores for Bioanalytical Applications Proceedings of the International Conference. 24-31.

Edel, J.B., and Albrecht, T. (2013) Engineered Nanopores for Bioanalytical Applications: A Volume in Micro and Nano Technologies. Elsevier Inc., 1-172.

Feng, W.N., He, J., Li, S.Y., Yang, D., Li, H.N., Yang, G.C., et al. (2020) Detection of Secretion of Exosomes from Individual Cell in Real-Time by Multifunctional Nanoelectrode-Nanopore Nanopipettes. Chinese J Anal Chem 48: e20061–e20068.

Fologea, D., Gershow, M., Ledden, B., McNabb, D.S., Golovchenko, J.A., and Li, J. (2005) Detecting single stranded DNA with a solid state nanopore. Nano Lett 5: 1905–1909.

Fraccari, R.L., Carminati, M., Piantanida, G., Leontidou, T., Ferrari, G., and Albrecht, T. (2016) High-bandwidth detection of short DNA in nanopipettes. Faraday Discuss 193: 459–470.

Frament, C.M., and Dwyer, J.R. (2012) Conductance-based determination of solid-state nanopore size and shape: An exploration of performance limits. J Phys Chem C 116: 23315–23321.

Fu, Y., Tokuhisa, H., and Baker, L.A. (2009) Nanopore DNA sensors based on dendrimer-modified nanopipettes. Chem Commun (Camb) 4877–4879.

Gubbiotti, A., Baldelli, M., Muccio, G. Di, Malgaretti, P., Marbach, S., and Chinappi, M. (2022) Electroosmosis in nanopores: computational methods and technological applications. Adv Phys X 7: 2036638

Huo, M.Z., Li, M.Y., Ying, Y.L., and Long, Y.T. (2021) Is the volume exclusion model practicable for nanopore protein sequencing? Anal Chem 93: 11364–11369.

Iqbal, S.M., Akin, D., and Bashir, R. (2007) Solid-state nanopore channels with DNA selectivity. Nat Nanotechnol 2: 243–248.

Kasianowicz, J.J., Brandin, E., Branton, D., and Deamer, D.W. (1996) Characterization of individual polynucleotide molecules using a membrane channel. In Proceedings of the National Academy of Sciences of the United States of America. Proc Natl Acad Sci U S A, pp. 13770–13773.

Kim, S.C., Kannam, S.K., Harrer, S., Downton, M.T., Moore, S., and Wagner, J.M. (2014) Geometric dependence of the conductance drop in a nanopore due to a particle. Phys Rev E - Stat Nonlinear, Soft Matter Phys 89: 42702.

Kowalczyk, S.W., Grosberg, A.Y., Rabin, Y., and Dekker, C. (2011) Modeling the conductance and DNA blockade of solid-state nanopores. Nanotechnology 22: 315101.

Kowalczyk, S.W., Wells, D.B., Aksimentiev, A., and Dekker, C. (2012) Slowing down DNA translocation through a nanopore in lithium chloride. Nano Lett 12: 1038–1044.

Levis, R.A., and Rae, J.L. (1993) The use of quartz patch pipettes for low noise single channel recording. Biophys J 65: 1666–1677.

Li, J., Gershow, M., Stein, D., Brandin, E., and Golovchenko, J.A. (2003) DNA molecules and configurations in a solid-state nanopore microscope. Nat Mater 2: 611–615.

Li, J., Stein, D., McMullan, C., Branton, D., Aziz, M.J., and Golovchenko, J.A. (2001) Ion-beam sculpting at nanometre length scales. Nature 412: 166–169.

Liebes, Y., Drozdov, M., Avital, Y.Y., Kauffmann, Y., Rapaport, H., Kaplan, W.D., and Ashkenasy, N. (2010) Reconstructing solid state nanopore shape from electrical measurements. Appl Phys Lett 97: 223105.

Loh, A.Y.Y., Burgess, C.H., Tanase, D.A., Ferrari, G., McLachlan, M.A., Cass, A.E.G., and Albrecht, T. (2018) Electric Single-Molecule Hybridization Detector for Short DNA Fragments. Anal Chem 90: 14063–14071.

Maglia, G., Restrepo, M.R., Mikhailova, E., and Bayley, H. (2008) Enhanced translocation of single DNA molecules through α-hemolysin nanopores by manipulation of internal charge. Proc Natl Acad Sci 105: 19720–19725.

McAllister, A.B., Menchaca, C.X., Freeman, J., Chen, R., and Shen, M. (2018) A Nanopipette Platform for Delivering Nanoliter Volumes of Acetylcholine. J Electrochem Soc 165: G3093–G3098.

Morris, C.A., Friedman, A.K., and Baker, L.A. (2010) Applications of nanopipettes in the analytical sciences. Analyst 135: 2190–2202.

Muthukumar, M. (2007) Mechanism of DNA transport through pores. Annu Rev Biophys Biomol Struct 36: 435–450.

Rodgers, P.J., Amemiya, S., Wang, Y., and Mirkin, M. V. (2010) Nanopipet Voltammetry of Common Ions across the Liquid−Liquid Interface. Theory and Limitations in Kinetic Analysis of Nanoelectrode Voltammograms. Anal Chem 82: 84–90.

Rothery, A.M., Gorelik, J., Bruckbauer, A., Yu, W., Korchev, Y.E., and Klenerman, D. (2003) A novel light source for SICM-SNOM of living cells. J Microsc 209: 94–101.

Schrlau, M.G., Dun, N.J., and Bau, H.H. (2009) Cell Electrophysiology with Carbon Nanopipettes. ACS Nano 3: 563–568.

Shevchuk, A.I., Gorelik, J., Harding, S.E., Lab, M.J., Klenerman, D., and Korchev, Y.E. (2001) Simultaneous measurement of Ca2+ and cellular dynamics: combined scanning ion conductance and optical microscopy to study contracting cardiac myocytes. Biophys J 81: 1759–1764.

Shi, W., Friedman, A.K., and Baker, L.A. (2017) Nanopore Sensing. Anal Chem 89: 157–188.

Si, W., and Aksimentiev, A. (2017) Nanopore Sensing of Protein Folding. ACS Nano 11: 7091–7100.

Song, L., Hobaugh, M.R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J.E. (1996) Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. Science (80-) 274: 1859–1866.

Steinbock, L.J., Lucas, A., Otto, O., and Keyser, U.F. (2012) Voltage-driven transport of ions and DNA through nanocapillaries. Electrophoresis 33: 3480–3487.

Steinbock, L.J., Otto, O., Chimerel, C., Gornall, J., and Keyser, U.F. (2010) Detecting DNA folding with nanocapillaries. Nano Lett 10: 2493–2497.

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

Suryavanshi, A.P., and Yu, M.-F. (2007) Electrochemical fountain pen nanofabrication of vertically grown platinum nanowires. Nanotechnology 18: 105305.

Talaga, D.S., and Li, J. (2009) Single-molecule protein unfolding in solid state nanopores. J Am Chem Soc 131: 9287–9297.

Umehara, S., Karhanek, M., Davis, R.W., and Pourmand, N. (2009) Label-free biosensing with functionalized nanopipette probes. Proc Natl Acad Sci U S A 106: 4611–4616.

Umehara, S., Pourmand, N., Webb, C.D., Davis, R.W., Yasuda, K., and Karhanek, M. (2006) Current rectification with poly-L-lysine-coated quartz nanopipettes. Nano Lett 6: 2486–2492.

Wadhawan, J.D., and Compton, R.G. (2013) Electrochemistry. 12th ed., Royal Society of Chemistry, 198-234.

Wang, Y., Montana, V., Grubišić, V., Stout, R.F., Parpura, V., and Gu, L.Q. (2015) Nanopore sensing of botulinum toxin type B by discriminating an enzymatically cleaved peptide from a synaptic protein synaptobrevin 2 derivative. ACS Appl Mater Interfaces 7: 184–192.

Wang, Z., Liu, Y., Yu, L., Li, Y., Qian, G., and Chang, S. (2019) Nanopipettes: A potential tool for DNA detection. Analyst 144: 5037–5047.

Wanunu, M. (2012) Nanopores: A journey towards DNA sequencing. Phys Life Rev 9: 125–158.

Wanunu, M., Morrison, W., Rabin, Y., Grosberg, A.Y., and Meller, A. (2010) Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. Nat Nanotechnol 5: 160–165.

Wilson, J., Sarthak, K., Si, W., Gao, L., and Aksimentiev, A. (2019) Rapid and Accurate Determination of Nanopore Ionic Current Using a Steric Exclusion Model. ACS Sensors 4: 634–644.

Xue, L., Yamazaki, H., Ren, R., Wanunu, M., Ivanov, A.P., and Edel, J.B. (2020) Solid-state nanopore sensors. Nat Rev Mater 5: 931–951.

Yang, L., and Yamamoto, T. (2016) Quantification of virus particles using nanopore-based resistive-pulse sensing techniques. Front Microbiol 7: 1500.

Yaul, M., Bhatti, R., and Lawrence, S. (2008) Evaluating the process of polishing borosilicate glass capillaries used for fabrication of in-vitro fertilization (iVF) micro-pipettes. Biomed Microdevices 10: 123–128.

Yu, R.-J., Ying, Y.-L., Gao, R., and Long, Y.-T. (2019) Confined Nanopipette Sensing: From Single Molecules, Single Nanoparticles, to Single Cells. Angew Chemie Int Ed 58: 3706–3714.

Yue, L., Amelia, Y., Cass, T., and Albrecht, T. (2017) A DNA Assay On a String – Fabrication and Sensing.

Zhang, H., Chon, C.H., Pan, X., and Li, D. (2009) Methods for counting particles in microfluidic applications. Microfluid Nanofluidics 7: 739–749.

Zhang, S., Li, M., Su, B., and Shao, Y. (2018) Fabrication and Use of Nanopipettes in Chemical Analysis. Annu Rev Anal Chem 11: 265–286.

Chapter 4

Noise optimisation

Synopsis: This chapter provides an introduction to electronic noise and the methods used to control this factor for experimental purposes. One of the aims of this research was to observe and differentiate bare DNA and modified DNA during translocation analysis. This chapter further discusses the methods used to optimise the signal to noise ratio for further experiments.

4.1 Introduction

4.1.1 Noise

Noise is an intrinsic and fundamental property of a system which is generated from the induction of current flowing through that system. Low frequency intrinsic noise, also known as white noise, is distributed along the frequency spectrum and separated into several categories including flicker, white (thermal, shot, and protonation), dielectric, and capacitive noise, as illustrated in Figure 4.1 (Edel and Albrecht, 2013). Flicker noise, or low-frequency fluctuations, arises from the ionic current generated from the application of a potential difference across a membrane. In nanopore experiments, this noise dominates at very low to low frequencies over thermal and shot noise (Edel and Albrecht, 2013). Thermal noise is unavoidable above absolute zero and is generated by random thermal motion of electrons and dominates at relatively low frequencies (Motchenbacher and Connelly, 1994; Edel and Albrecht, 2013). In a nanopore system, shot noise is introduced due to random fluctuation of ion motion inside the channel (Ott, 1988; Edel and Albrecht, 2013; Fragasso *et al.*, 2020). Dielectric noise is closely associated with the capacitance of the nanopore and its local environment. Some materials which are non-ideal for use in nanopore sensing can transfer some of the electrical energy to heat energy and increase the contribution of thermal noise. This noise dominates the mid frequency

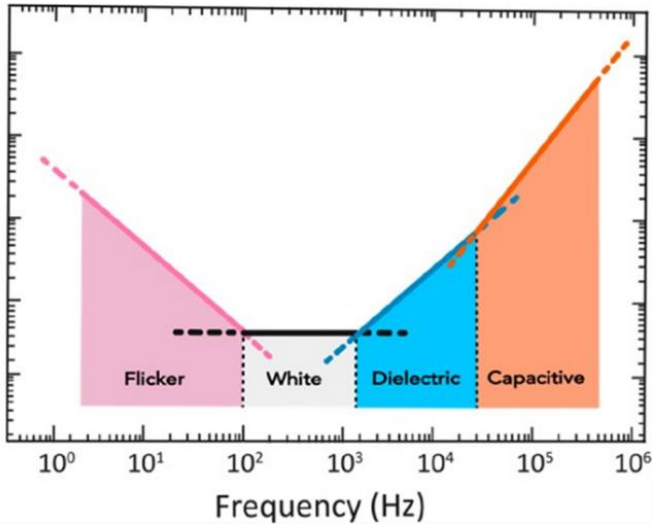

Figure 4.1. Illustration of the noise sources present in ionic current. Adapted from Figure 2 (Fragasso et al., 2020).

spectrum (Edel and Albrecht, 2013; Fragasso *et al.*, 2020). Above ~10 KHz capacitive noise becomes the dominant noise contributor. The total capacitance of the system is calculated as the sum of all of the individual capacitances for each system component (Tabard-Cossa *et al.*, 2007; Edel and Albrecht, 2013; Hartel *et al.*, 2019). Protonation noise in nanopores relates to the fluctuation of wall surface charges, and its effect extends in the mid frequency noise spectrum (Hoogerheide *et al.*, 2009; Edel and Albrecht, 2013).

Extrinsic noise is the induction of unwanted frequencies in the desired electrical signal (Fish, 2017). Noise can be induced in a variety of different ways, from signals in the air, vibrations through the benchtop, the electronics within the device, and/ or from the power supply itself (Motchenbacher and Connelly, 1994; Vasilescu, 2006; Fish, 2017). Noise causes distortions in useful data, and it is therefore vital to minimise noise induction when performing sensitive data acquisition and transfer. This improves the signal to noise ratio during signal acquisition (Vasilescu, 2006; Edel and Albrecht, 2013; Fish, 2017).

Noise influence on a signal is typically characterised through use of power spectral density and root means square (RMS) analysis. When a signal is generated, it is comprised of a multitude of frequencies which contribute to the overall structure of the signal. The more contribution a frequency makes to generating a signal, the increased power that frequency has (Smeets *et al.*, 2008). Fourier transform (FT) is a mathematical operation which transforms a function into the spatial or temporal frequency domain (Boehme and Bracewell, 1966). Applying FT to a time-current dataset allows for the determination of specific frequencies and their relative effect on signal generation (Smeets *et al.*, 2008). It is also possible to obtain the frequency and

relative power spectral density (PSD) contribution. RMS is a mathematical property by which the signal to noise ratio (SNR) is determined. RMS is defined as the square root of the sum of the mean squared current at each time point, divided by the total number of samples used (Daintith, 2009). This equation is highlighted in equation 4.1.A and simplified respective to current modulation in 4.1.B. SNR is then calculated as the absolute current change induced by DNA translocation divided by the RMS noise value determined. This provides a direct measurement of all the noise frequencies experienced as a function of the deviation from the mean. This is significantly important to isolate nanopore translocation events. With high RMS values, it becomes more difficult to separate events from noise influence (Dutta and Horn, 1981; Albu and Heydt, 2003; Parkin and Drndić, 2018; Liang *et al.*, 2020).

$$x_{RMS} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \cdots x_n^2} \qquad (4.1.A)$$

Where x_{RMS} represents the RMS noise value, n represents the number of samples, and $x_{1,\,2,\,n}$ represents the current value at a particular time point.

$$I_{RMS} = \sqrt{\Delta I^2}(t) \qquad (4.1.B)$$

Where I_{RMS} represents the RMS noise value, $\Delta I^2(t)$ represents the change in current over a set time period.

4.1.2 Frequency filtering

To ensure high resolution data, it is often necessary to filter a generated signal (Vasilescu, 2006). Manual signal filters are comprised of a circuit designed to remove undesirable or enhance desired frequencies. Filters work based on differently assembling a combination of resistors, capacitors, and inductors in an electrical circuit.

Low pass filters place a limit on the maximum frequency that can pass through the filter, from 0 Hz to its cut-off frequency, creating a steep decrease in current intensity above this value (Shenoi, 2005; Fish, 2017). In an ideal filter, all the frequencies above the filter threshold would be eliminated, however this would require the signal to be unlimited in its time domain. In practically used filters this is not possible and therefore the signal is attenuated, respective to the order of the filter used. In a first-order filter, the signal amplitude is reduced by half as frequency doubles. In a second-order filter signal amplitude is reduced by a factor of four as frequency doubles. This generates a filtered signal with a “corner” at the filter frequency applied. The issue with increasing the order of a filter is that the desired filter frequency becomes attenuated relative to the order applied. Therefore, the signal strength can be reduced below the desired filter frequency (Fish, 2017).

The two types of filters which utilise these principles are passive and active filters (Davis, 2017). Passive filters are power supply independent, comprising of resistors, capacitors, and inducers. These filters have a specific working frequency range that cannot be altered. If a different type or frequency of filter is required, these can be removed from the circuit and replaced (Ott, 1988; Slichter *et al.*, 2009; Davis, 2017). Active filters are power dependant filters which are able to be tuned to filter different frequencies. Whilst active filters offer more adaptability than inline filters, due to the power required, they can also introduce different frequencies of noise (Ott, 1988; Cochrane *et al.*, 2003).

When using any filter to limit the signal, it is possible to induce signal distortions, most discussed in terms of optic signalling. The application of a single filter often offers the desired effect on bandwidth limitation, however when several filters are applied in

series to a signal, each with the same cut-off frequency, they can narrow the overall bandwidth and cause clipping of the signal. Therefore, limiting the number of filters within a circuit may not eliminate all unwanted frequencies, but can potentially reduce the amount of data loss (Khrais *et al.*, 1996; Otani *et al.*, 1999; Downie *et al.*, 2002).

4.1.3 Methods for noise optimisation

There are several methods employed to reduce noise in a variety of systems. In high sensitivity instrumentation, it is often important to reduce the level of vibration experienced (Olson, 1956; Fish, 2017). For sensitive single molecule sensing equipment, such as scanning tunnelling microscopes, atomic force microscopes, and low noise amplifiers, engineering precautions, including anti-vibration tables, are utilised, counteracting vibrations through the floor (Jeon *et al.*, 2006). The equipment is also stored and used inside a Faraday cage. This is typically a box made of, or surrounded by a highly conductive material, such as copper, which blocks electromagnetic radiation from permeating and effecting signal acquisition (Gamry, 2021). It is also vital to shield any cables running from the equipment to the visualisation device. When transferring an analogue signal from its generation point to its respective interpreter, it is possible for external electromagnetic waves to influence the signal. In a similar manner to a Faraday cage, it is necessary to cover any cables outside of the cage with a conductive material to deflect any ambient signals. These protective layers require a connection to ground, this allows for excess interference to be removed. This also allows for noise travelling along the power supply cable to be negated (DATAFORTH, 2002; Fish, 2017).

It is also possible to digitally filter data to enhance the signal to noise ratio removing unwanted frequencies. One way this is done is by the use of a Bessel filter. This

method is a linear analogue filter applying a low pass filter for post-acquisition signal processing (Rae and Levis, 1992; Srivastava *et al.*, 2021).

4.2 Methodology

PSD and RMS were used for all noise characterisation with PSD used to determine the frequency of noise captured, and RMS for overall noise contribution. To minimise the noise contribution to the signal several factors were investigated. The preparation of electrodes was explored as a potential to minimise noise contribution. To explore the potential of minimising vibrational influences (white noise), altering electrode length and using an anti-vibration table were investigated. To reduce potential extrinsic noise sources, the use of standard shielded, and silver shielded data transfer cables (Pearl cable, AudioQuest) was explored. The use of mains power and battery power to perform sensing experiments was also explored to minimise extrinsic noise. Finally, a comparison between passive and active filters was performed to complement the investigation into the shielding of extrinsic noise.

4.2.1 Low noise amplifier

All noise experiments were conducted using a custom-made low-noise, wide bandwidth current amplifier. The amplifier uses CMOS technology, splitting the ionic current into open pore (DC channel), and nanopore (AC channel) channels. The DC channel measures the ambient current of the system and provides a signal which can be normalised against. The AC channel measures the direct effect of current passing through the nanopore, sampling at 1 MHz.

Liquid cells were prepared as described in section 3.2.6 and placed inside the inner Faraday cage, with the pipette-electrode connected to the low noise amplifier, as illustrated in Figure 4.2. The counter electrode was connected to the power supply and used to produce a specific voltage. A Picoscope 4262 oscilloscope was used for real time monitoring of the system and as a voltage source.

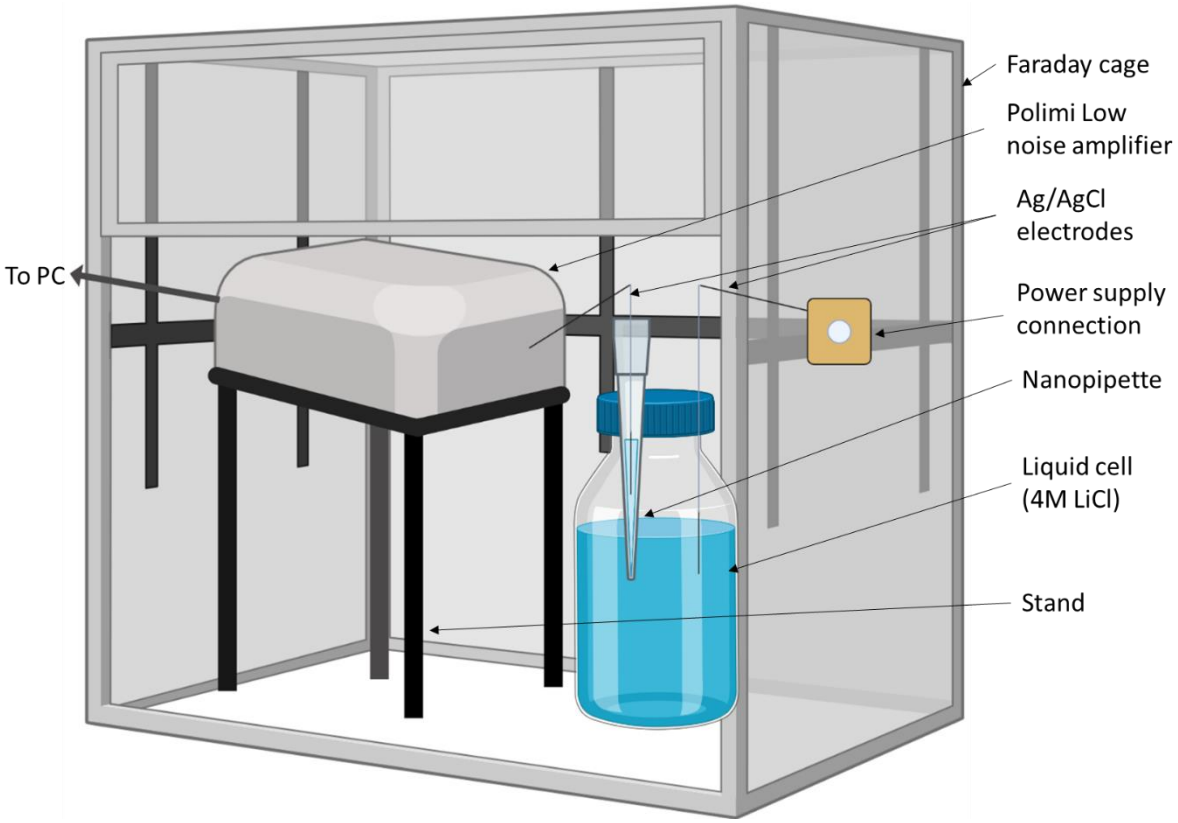

Figure 4.2. Simplified translocation experimental set-up showing the liquid cell connection to the low noise amplifier inside a Faraday cage.

Custom code was written in MATLAB to extract and combine (courtesy of Tim Albrecht) thousands of DNA translocation events and is available in Appendix 5. Initially, the AC channel scans were background corrected, generating a mat.corr file. DNA translocation was then determined by applying a cut-off of 5X the standard deviation (5σ). The event information was extracted, generating two files, one corresponding to information about the event such as max current, and event duration, and the other containing the data points involved in the event. All of the files from each bias were then merged together to generate a mat.all file. These .all files were then further analysed to include additional event information including event characteristics based on using full width at half max and max current as event distribution criteria. This generated a mat.ana file which could then be further analysed.

4.2.2 Electrode preparation

Electrodes were prepared in two different ways, chemical synthesis, and anodization. For anodization, the method described in section 3.2.2 was used. For chemical synthesis, silver wire (0.25 mm diameter, 99.99% purity, Goodfellows) was cut to 8 cm long. Of this, 6 cm was submerged in 37.5% nitric acid for 10 seconds until white in colour, before washing in ultra-pure water (MilliQ®, 18.2 MΩ). The remaining nitric acid was then neutralised using sodium bicarbonate and disposed of with running water. The wire was then soldered to a gold coated connector pin and transferred into a beaker containing sodium hypochlorite in the form of household bleach (Clorox) for one hour, ensuring the wire had changed from white to a purple/ black colour. The wires were then washed in ultra-pure water (MilliQ®, 18.2 MΩ) prior to use. Noise features were characterised at 25 kHz filtered, and electrodes imaged using an optical microscope (Nikon, Type 105).

4.2.3 Electrode length

Electrodes were cut to several different lengths prior to preparation to determine the effect on noise. Electrodes were cut to 12 cm, 10 cm, and 8 cm in length. Above 12 cm in length was not possible due to the space available and the risk of interacting with the Faraday cage, or other wires. Below a length of 8 cm was not possible either due to the length of wire required to reach both power supply and low noise amplifier. These electrodes were all prepared using the anodization method described in section 3.2.2. Noise features were characterised at 100 kHz filtered.

4.2.4 Cable alteration

Several different cable connections were investigated regarding the level of shielding around the cables and the effect on the noise experienced. A standard USB type B

cable was initially used to transfer data from the oscilloscope to the desktop. This was compared against a 5% silver shielded USB type B cable (Pearl cable, AudioQuest). Experimental cables were also tested both unshielded and shielded using copper tape (99.99% purity, RS components) and wrapped in parafilm. Noise features were characterised at 10 kHz filtered.

4.2.5 Anti-vibration table (Granite)

Granite is an excellent material for high frequency noise reduction. Experimental measurements were taken with the experimental set up both on and off a solid 2.5 cm granite block to determine the effect on high frequency noise. Noise features were characterised at 100 kHz filtered.

4.2.6 Altering power supply and effect of grounding

To determine the effect of main line power on the noise characteristics, a lithium-ion battery power pack was created (courtesy of Tim Albrecht) and used to power the amplifier. This was compared against mains power to the device through a power supply (KEYSIGHT, E3602A). Grounding sources were then tested to determine the effect on noise. A shielded copper wire was prepared by attaching crocodile clips to each end. One end was attached to the outer Faraday cage, and the other to; a gas line attached to the wall of the lab, a radiator attached to the lab wall (drained and switched off), or a specific ground line attached to the fuse board for the entire lab. For contributions to noise, this was compared against complete removal of the grounding cable. Noise features were characterised at 100 kHz filtered.

4.2.7 Filtering incoming electrical signal

To filter the data signal, several different approaches were used. Firstly, an active signal filter (KROHN-HITE LP/HP/BP/BR Butterworth/ Bessel dual channel filter, model 3940) was used varying the filter frequency at 10, 25, 100, 240 and 500 kHz. Secondly, passive filters (10, 25, 50, 100, and 240 kHz Low-Pass, BNC Feedthrough filters, ThorLabs) were used, inserted between the low noise amplifier and the oscilloscope.

4.2.8 FT and noise analysis

FT analysis and noise analysis was performed using a custom written MATLAB code (courtesy of Tim Albrecht), available in Appendix 4. Frequency peak detection and selection was performed manually using the inbuilt MATLAB data cursor to highlight areas of interest.

4.3 Results and discussion

Translocation of analytes is often not observed at low biases, therefore when optimising for minimal noise, it was necessary to focus on biases between 0.4 and 0.8 V. When using increased voltages, the speed at which analytes translocate increases, and therefore the resolution is reduced. It was therefore decided to focus on optimising noise parameters between 0.4 to 0.8 V. Noise also changes from experiment to experiment; therefore, repeated experiments were performed to determine overall effect of altering conditions. It is known that with an increase in pore size comes an increase in the noise experienced during experimentation, therefore where possible, the same pore, or a pore of similar size and taper were used (Kowalczyk *et al.*, 2011). To highlight the specific effect on data acquisition, example scans are shown zoomed in to between 3 and 4 seconds ($3x10^6$ and $4x10^6$ data points), and the RMS noise stated. All noise experiments were conducted to N = 3, however due to baseline noise variation between experiments, the figures shown are representative of the effect observed in each experiment.

4.3.1 Electrode preparation

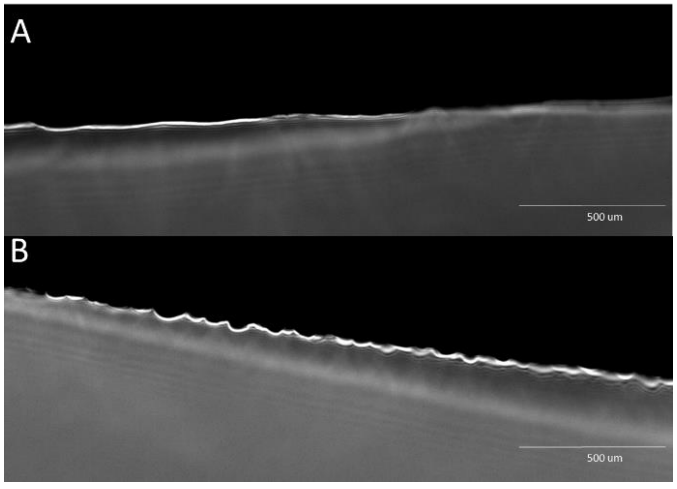

Figure 4.3. Electrode surface images at 40X magnification imaged using an optical microscope (Nikon, Type 105) produced by anodization (A) and chemical synthesis (B).

Electrodes imaged using optical microscopy show that when electrodes are prepared using anodization, then a smooth edge is generated. When chemical synthesis is used to produce electrodes, then a rougher surface, is produced, as illustrated in Figure 4.3.

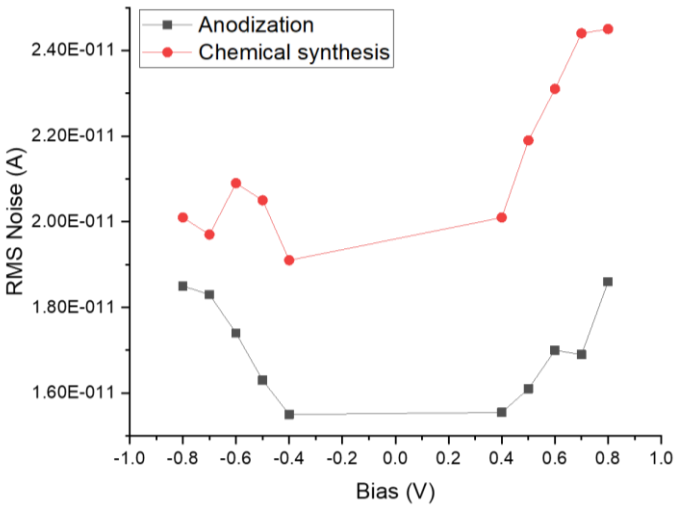

Figure 4.4. Noise profiles of anodized (grey) and chemically synthesised (red) Ag/AgCl electrodes.

When analysing the RMS noise profiles of both electrodes, it was observed that, as illustrated in Figure 4.4, when using a chemically synthesised electrode, the noise is much higher than in anodized electrodes. It can be suggested that the surface structure contributes to the increase in noise. Electrodes with a rough surface have a greater surface area, which can increase bubble formation, and trap bubbles close to the electrode surface. This reduces the effective surface area in contact with the electrolyte and can cause fluctuations in the current observed, increasing ambient noise (Taqieddin *et al.*, 2017). The size of the nanopore used for this experiment was ~14 nm.

From the results obtained, it was determined that anodized electrodes would be used for further experiments.

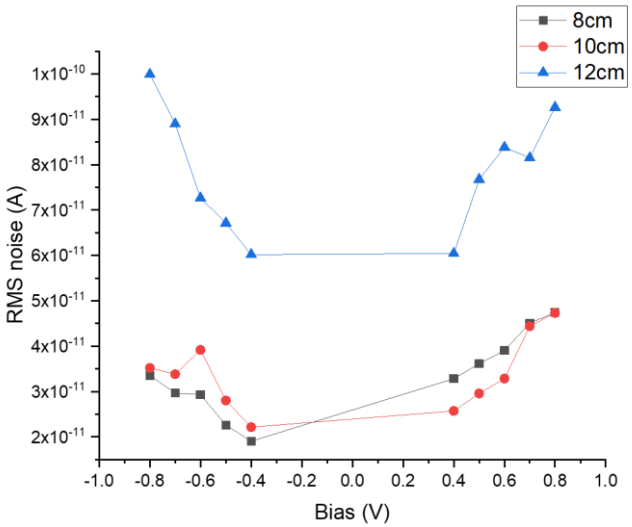

Figure 4.5. RMS noise profiles of electrodes at 8, 10, and 12 cm in length.

4.3.2 Electrode length

Differences in the overall length of the electrode show a difference in noise. As illustrated in Figure 4.5, when increasing electrode length to 12 cm, noise increases by an average ~45.0 pA compared to 8 cm electrodes, and ~44.6 pA

compared to 10 cm electrodes. The flexibility of the cables could contribute to this increase in noise. Longer electrodes can act as antennas, picking up noise which can distort the signal. They also have an increased capacitance which can increase the contribution to capacitive noise. The average noise difference between the 10 cm electrodes and the 8 cm electrodes is 0.4 pA. Whilst this average difference is small, the RMS noise was lower for 8 cm when using negative biases. For translocation of analytes such as DNA, it is the application of negative bias that is most interesting in our experimental set up. As the working electrode is outside the nanopipette, DNA will translocate across the pore when a negative bias is applied. It was therefore decided that 8 cm electrodes would be used for further experiments. The size of the nanopore used for this experiment was ~18 nm.

4.3.3 The effect of cable shielding on noise performance

The effect of changing the output cable between the oscilloscope and the PC had a significant effect on the RMS noise output. As is illustrated in Figure 4.6, it is possible to remove specific low frequency noise spikes from the data when changing from the standard USB B data transfer cable (4.6.A) to the silver shielded USB type B cable (Pearl, AudioQuest, 4.6.C). The shielded cable was able to remove frequencies at 50, 150 and 250 Hz. This is reflected in the corresponding scans (4.6.B and D respectively), with the standard USB B data transfer cable showing increased fluctuation in the baseline.

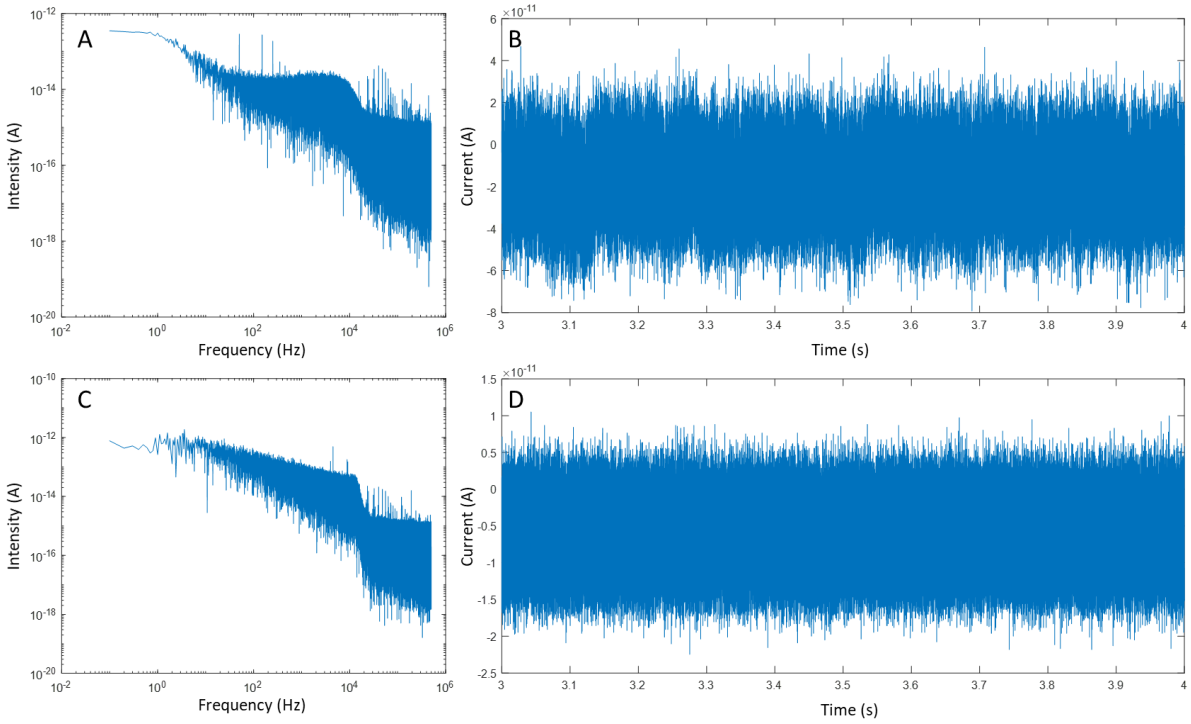

Figure 4.6. The difference in Fourier transform results and example scans between a standard printer cable (A and B respectively) and the silver shielded data transfer cable (C and D respectively).

Whilst removing the low frequency noise, it can be observed that frequencies at 4000, 4500, 8850, and 9150 Hz are induced. Alteration of the cable did not improve the induction of high frequency noise above 10 kHz. The size of the nanopore used for this experiment was ~10 nm.

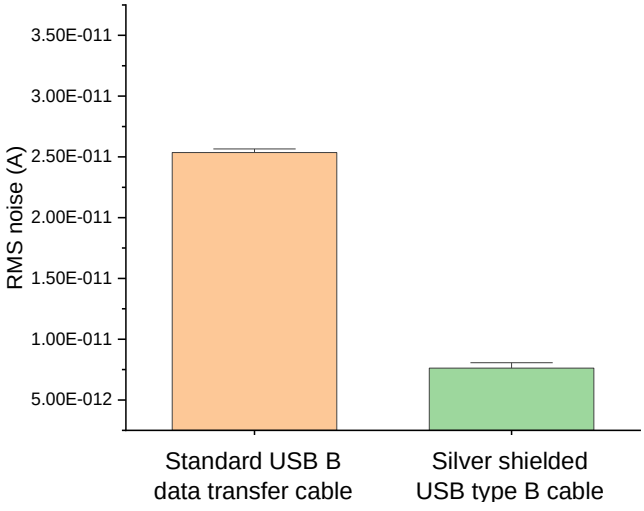

Figure 4.7. RMS noise observed between use of the standard printer cable and the silver shielded USB type B cable (Pearl cable, AudioQuest) for N=3.

The overall noise of altering the cables decreased when using the silver shielded USB type B cable (Pearl, AudioQuest) from an average of 25.3 ± 0.2 pA to 7.6 ± 0.2 pA ($p = <0.001$, One-Way ANOVA). This indicates an average decrease of 17.7 pA, as shown in Figure 4.7.

The size of the nanopore used for this experiment was ~7 nm. Due to the need to extract translocation events from the scans, a reduction in oscillation and overall noise improves the ability of the extraction code. It was therefore decided that the shielded cable would be used for further experiments.

4.3.4 Anti-vibration table (Granite)

Results were compared between the nanopore system placed on the benchtop and placed on a granite slab. As is highlighted in Figure 4.8, it can be seen that when running an experiment without the granite surface (4.8.A), there is an increased contribution to noise at 192 and 440 kHz. Using the granite surface (4.8.B), removed these contributions, however noise at 300 kHz was unaffected.

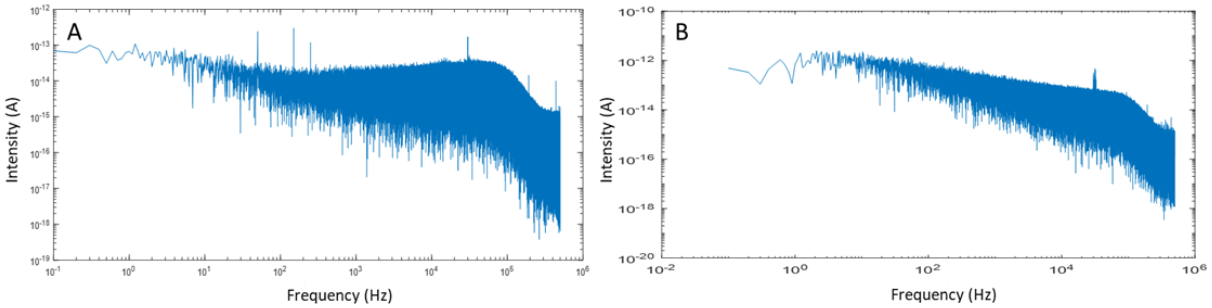

Figure 4.8. Fourier transform of data from experiments performed on the benchtop (A) and on a granite table (B).

This effect is thought to be due to the mass of the granite slab which reduces vibrations passing through the benchtop into the equipment (Hartlieb and Grafe, 2017; Li and Einstein, 2017). The overall noise for this experiment was 24.0 ± 0.4 pA for experiments on the granite slab and 22.2 ± 0.2 pA for the benchtop. The size of the nanopore used for this experiment was ~22 nm. From the results of the experiment, although there is a slight increase in overall noise, it was determined that further

experiments would be performed on top of the granite surface to limit contributions of high and low frequency noise.

4.3.5 Altering power supply and effect of grounding

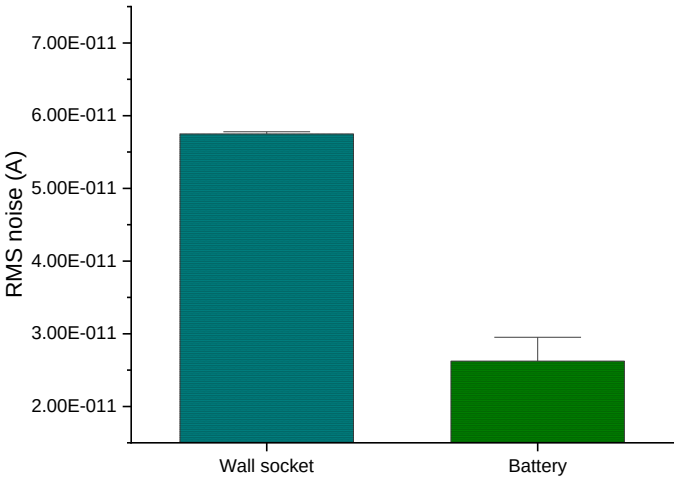

Figure 4.9. RMS noise in an experimental set up comparing using mains power against a battery pack for N=3.

As highlighted in Figure 4.9 it can be seen that by altering the power supply from mains power to a battery pack, the average noise is reduced by 31.2 pA from an average of 57.5 ± 0.2 pA to 26.2 ± 2.4 pA (p = <0.001, One-Way ANOVA). The example scans from this experiment, as illustrated in Figure 4.10, highlight that the mains power scan (4.10.A) shows increased oscillations compared to the battery power (4.10.B).

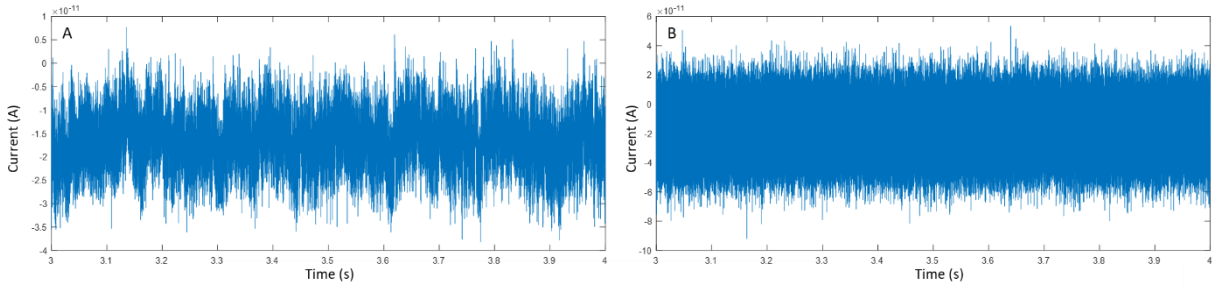

Figure 4.10. Example scans from experiments using mains power (A) and the battery pack (B).

When comparing the grounding source, it can be seen that grounding to the gas line provides the lowest overall noise at 13.4 ± 0.3 pA, while no grounding has the highest noise at 27.4 ± 0.3 pA. Grounding to the radiator had an average 18.9 ± 0.5 pA, the fuse box had 13.3 ± 0.3 pA and the plug socket 13.2 ± 0.1 pA, as illustrated in Figure 4.11.

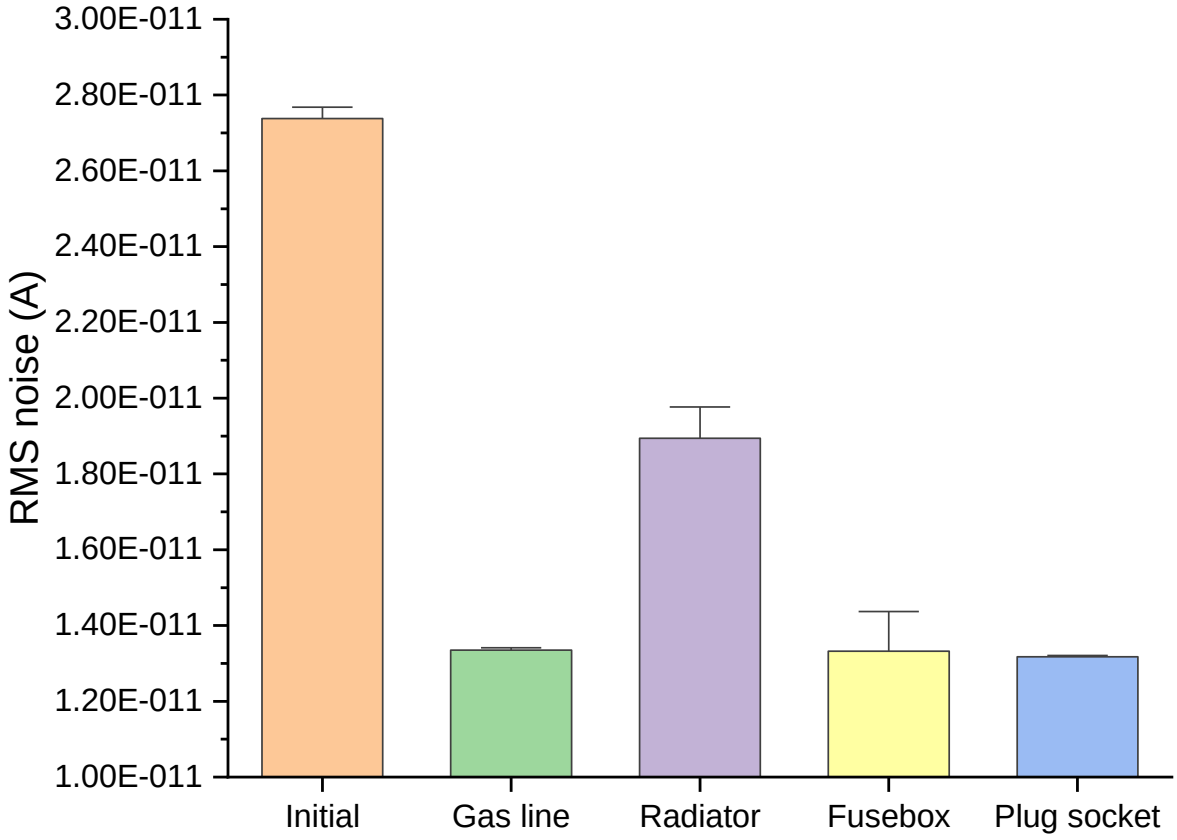

Figure 4.11. RMS noise comparison between different grounding sources for N=3.

The initial ungrounded experiment showed the induction of many unwanted frequencies of noise, in both the low and high frequency spectrum, which is illustrated in the scans produced (Figure 4.12.A and B respectively). Whilst the gas line provided the lowest overall noise, when comparing the scans generated and the resultant FT plots, it can be seen there is a significant difference between the groups. As is highlighted in Figure 4.12, the gas line grounding produces scans with increased oscillations in the baseline and increased high frequency noise peaks (Figure 4.12.C and D). This decreases the SNR and reduces potential resolution of translocation events. This is similarly present in the plug socket grounding experiments (Figure 4.12.I and J). Whilst the fuse box grounded scans had limited noise overall and a decrease in baseline oscillations (Figure 4.12.H), there is a drastic increase in the low frequency spectrum noise peaks (Figure 4.12.G). This is thought to be due to a

grounding loop issue. As it is necessary to power some of the equipment using mains power, grounding through the same circuit can induce low frequency noise feedback. Using the radiator as a grounding point provided a balance between unwanted noise frequency induction and stable scans (Figure 4.12.E and F). The low frequency noise is limited whilst the baseline has significantly reduced oscillations.

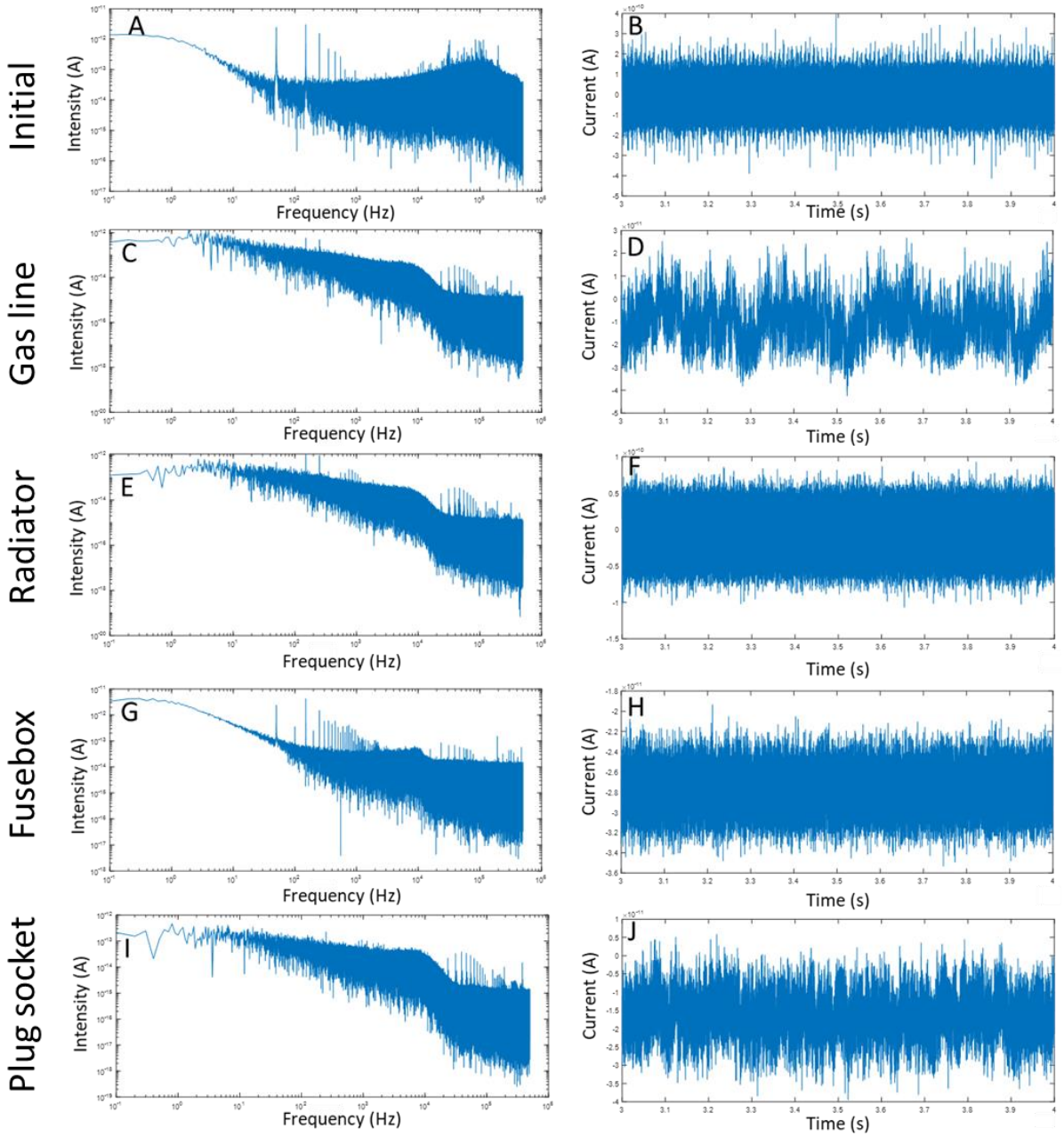

Figure 4.12. Fourier transform (A, C, E, G, and I) and example scans (B, D, F, H, and J) for different grounding sources including no grounding, gas line, radiator, fuse box, and plug socket respectively.

The size of the nanopore used for this experiment was ~16 nm for ungrounded, gas line, and plug socket samples, and ~17 nm for fuse box and radiator samples. Decreasing oscillations improves the ability to separate out translocation events from the noise. Therefore, it was decided that for high sensitivity experiments, the battery pack would be used. As a compromise between noise and event detection, the radiator was selected as the grounding point. The decrease in oscillation and noise frequency intensity is preferable over an overall noise decrease.

4.3.6 Filtering incoming electrical signal

Two methods of filtering the signal were used. Initially the analogue filter was used to determine the effect of altering the frequency on noise. An increase in the filter frequency, increases the resolution of the events observed and is necessary to detect and analyse subevents during translocation. The results, as illustrated in Figure 4.13, illustrates with an increase in filter frequency, there is an increase in noise and an increase in the standard deviation of the noise. This suggests an increasing influence of bias on noise at higher filter frequencies.

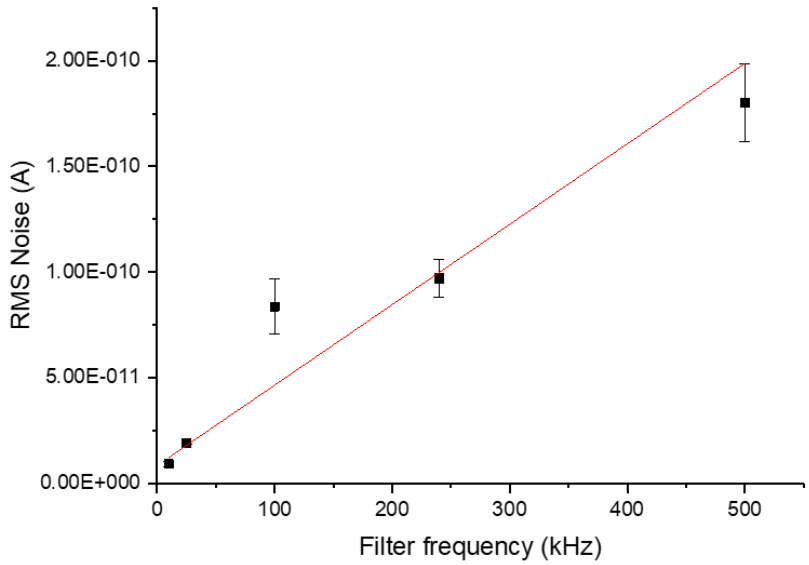

Figure 4.13. Effect of increasing filter frequency on RMS noise using an active filter with the error bars representing the standard deviation across the different experimental biases used (-0.8 to +0.8V), with a linear line of best fit to highlight the increase in noise.

In experiments conducted between 10 and 500 kHz filter frequencies, dielectric and capacitive noise are the contributing factor to noise. The results from this experiment follow a linear pattern and have an R^2 value of 0.93, indicating a relatively high degree of accuracy to the fit. As previous research suggest a linear increase to noise within the capacitive noise region, it can be suggested that these results concur with those previously determined (Edel and Albrecht, 2013; Hartel *et al.*, 2019; Fragasso *et al.*, 2020; Liang *et al.*, 2020). Further experimental results from lower filter frequencies, <10 kHz, are required to confirm the gradient change at 10 - 500 kHz.

It can be observed, as illustrated in Figure 4.14.A, when using a passive filter, noise increases similarly to that observed in the active filter. Noise increases respective to the filter frequency used, with the linear fit R^2 value of 0.99. This indicates a very high accuracy to the fit used. Figure 4.14.B highlights that the passive filters provide a lower overall noise profile compared to the analogue filter up to a frequency of 100 kHz.

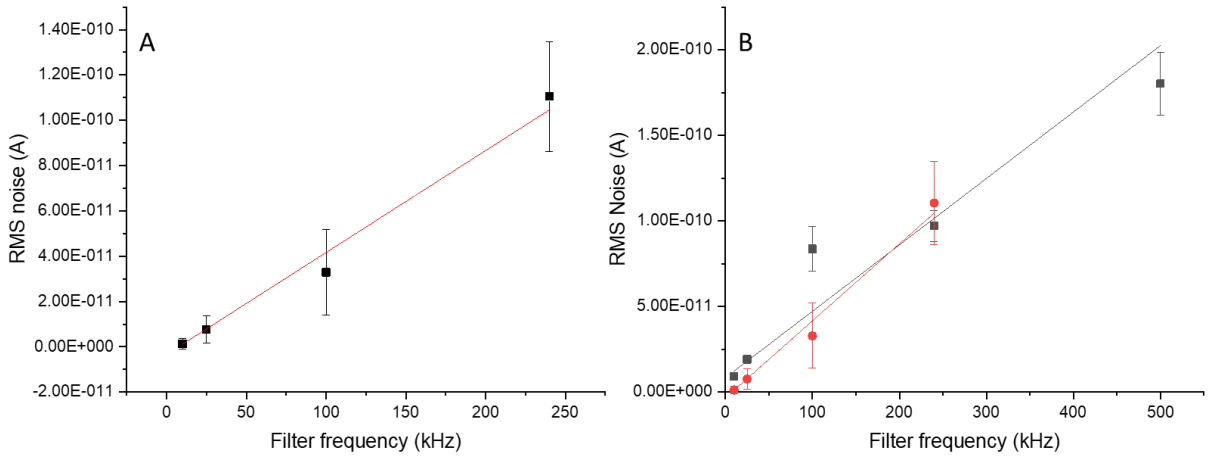

Figure 4.14. Effect of increasing filter frequency on RMS noise using passive filters (A) with the error bars representing the standard deviation across the different experimental biases used (-0.8 to +0.8 V). Comparisons between passive and active filters illustrates decreased noise in the passive filters up to 100 kHz, and increased noise in 240 kHz filtered results (B) for N=3.

Noise was reduced by 8.0 ± 2.4, 11.4 ± 5.1, and 50.8 ± 19.0 pA in the passive filters at 10, 25 and 100 kHz respectively when compared to the active filter. When using a 240 kHz filter, noise increased in the passive filter to 110.5 ± 24.3 pA, which is ~13.5 pA

above the analogue filter. Unfortunately, it was not possible to repeat the 500 kHz for the passive filter, as this filter is not owned by the group. Due to the reduced noise, and potential event resolution, when using the 100 kHz passive filter, it was decided that this filter would be used for translocation experiments.

Whilst the difference in filters provided a small improvement in noise, the most significant effect of changing from active to passive filters can be observed in the scans generated. As demonstrated in Figure 4.15, using a 10 kHz filter, but changing between the active (Figure 4.15.A) and passive (Figure 4.15.B) filters, decreases the baseline oscillation. This is also observed at 100 kHz (Figure 4.15.C and D respectively). As previously described in section 4.3.5, this change could be contributed to reducing the influence of mains power on the signal acquisition.

The size of the nanopore used for this experiment was ~15 nm. Decreasing oscillations and overall noise will allow for improved SNR. From the results obtained, it was decided that passive filters would be used for future experiments.

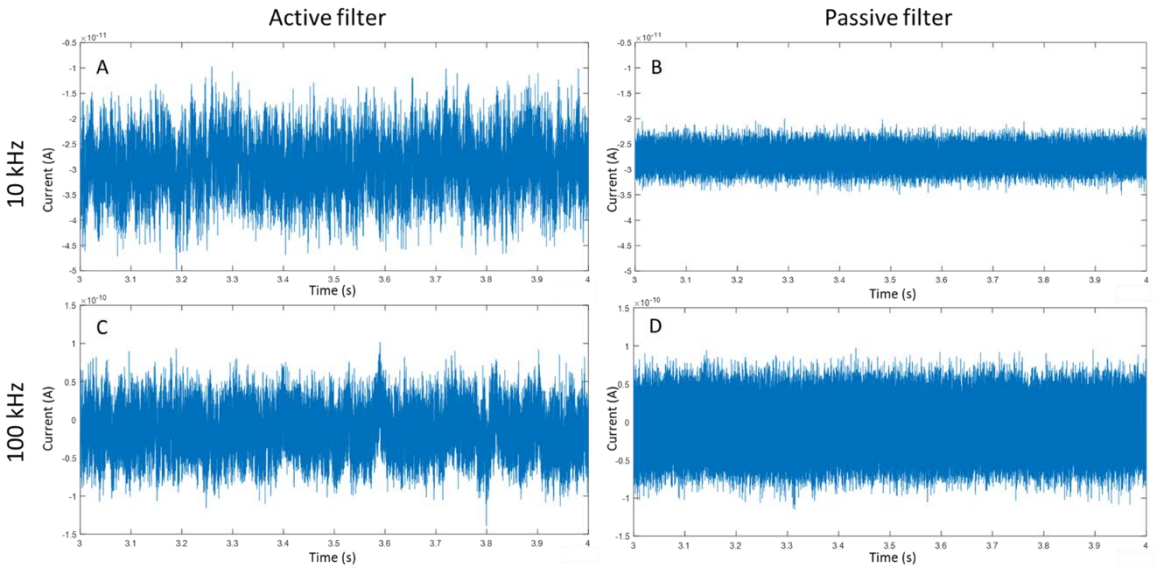

Figure 4.15. Example scans from 10 kHz filtered scans (A and B) from active (A) and passive (B) filters showing a decrease in baseline oscillation. Examples scans filtered at 100 kHz (C and D) for active (C) and passive (D) filters also show this decrease in baseline oscillation.

4.3.7 Noise profile applying all noise reduction methods

From applying all of the noise reduction techniques investigated during this research, it has been possible to reduce noise influence in scans obtained. As illustrated in Figure 4.16, it was possible to significantly reduce lower noise frequencies, with a very limited peak appearing at 50 Hz.

There is still some high frequency noise observable above 100 kHz, however the contribution of these frequencies to the scans obtained is significantly reduced, as can be observed in Figure 4.16.B. The oscillation in the baseline is minimal, and the overall noise was 24.0 ± 12.9 pA for the experiment. The size of the nanopore used for this experiment was ~20 nm.

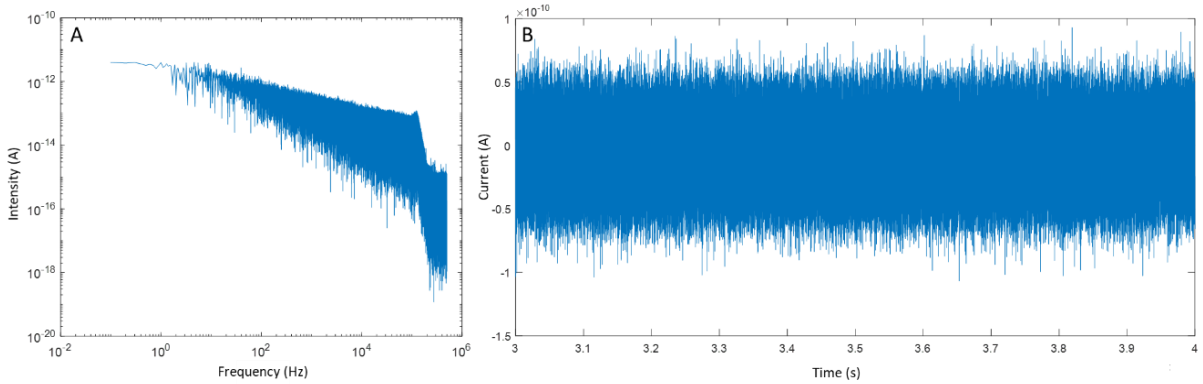

Figure 4.16. Example of 100 kHz filtered FT plot (A) and resultant scan (B) from applying all noise minimising techniques after optimisation.

4.4 Conclusion

From the results obtained it is possible to conclude that by preparing electrodes via anodization it was possible to decrease the noise experienced during data acquisition by a minimum of 7% to a maximum 31%, bias dependant. This could be due to the smooth surface created during anodization. From these results it was decided that anodised electrodes would be used for further experiments. Reducing the length of the electrodes from 12 to 10 cm showed a significant decrease in noise experienced.

Whilst decreasing the length further to 8 cm did not show a significant decrease in noise reduction, the shorter length prevents the potential for the electrode to make contact with the Faraday cage. Due to both lower noise acquisition and ease of experimental set-up, it was decided that the electrodes would be prepared at 8 cm in length. The use of a silver shielded USB type B cable (Pearl, AudioQuest) over a standard USB B data transfer cable allowed for the removal of some low frequency noise peaks. The cable was however unable to remove the high frequency noise peaks observed. Due to the average decrease of 17.8 pA in RMS noise, it was decided that the silver shielded cable would be used for further experimentation. Through the use of a granite surface to conduct sensing experiments, it was possible to limit the high frequency noise observed when performing the same experiment on the benchtop. It was also found that there was a decrease in the low frequency noise peaks. An increase of 1.9 pA RMS noise was observed when using the granite slab, however due to the decrease in the high frequency noise peaks it was decided that further experiments would be performed on the granite slab. When comparing the effects of changing between an active and passive filter, there is an overall decrease in the RMS noise values when using the passive filters below 240 kHz. At 240 kHz there is an observable increase in RMS noise. However due to the decrease in observable oscillation when using the passive filters, it was decided that these would be employed for DNA translocation experiments. This effect could be due to removing a component which relies on mains power to function. The decrease in oscillation allows for greater accuracy and reliability when identifying translocation events, improving the SNR, and decreasing false positive results during data extraction. To expand further on isolation of the sensing equipment from mains power, the use of a battery to power the amplifier

saw a significant reduction in RMS noise of 31.2 pA. This procedure also saw a significant decrease in oscillations when comparing the scans, further improving data extraction. It is possible to conclude that a major contributor to the noise observed during experiments is generated through the use of mains power. Therefore, where possible, elements of the experimental set-up were isolated from mains power and operated using battery power. When comparing the combination of all noise reduction techniques it was possible to significantly reduce the effect of extrinsic noise and reduce overall RMS noise. This will allow for greater resolution of DNA translocation events, and significantly improve SNR. Experimental procedure will involve using shorter, anodised electrodes, using a silver shielded data transfer cable, working on a granite slab, using a passive filter of 100 kHz or less, and working in an “off grid” approach using battery power where possible. Comparing the optimised experimental procedure to expected current blockage induced by DNA translocation, using a 100 kHz passive filter, a SNR of ~5 would be expected.

4.5 References

Albu, M., and Heydt, G.T. (2003) On the use of RMS values in power quality assessment. IEEE Trans Power Deliv 18: 1586–1587.

Boehme, T.K., and Bracewell, R. (1966) The Fourier Transform and its Applications. Am Math Mon 73: 685.

Cochrane, D., Chen, D.Y., and Boroyevic, D. (2003) Passive cancellation of common-mode noise in power electronic circuits. IEEE Trans Power Electron 18: 756–763.

Daintith, J. (2009) A dictionary of physics. 6th ed., Oxford University Press.

DATAFORTH (2002) R. 221–222 https://www.dataforth.com/protecting-signal-lines-against-electromagnetic-interference.aspx. Accessed October 24, 2021.

Davis, N. (2017) An Introduction to Filters - Technical Articles. allaboutcircuits.com https://www.allaboutcircuits.com/technical-articles/an-introduction-to-filters/. Accessed October 24, 2021.

Downie, J.D., Tomkos, I., Antoniades, N., and Boskovic, A. (2002) Effects of filter concatenation for directly modulated transmission lasers at 2.5 and 10 Gb/s. J Light Technol 20: 218–228.

Dutta, P., and Horn, P.M. (1981) Low-frequency fluctuations in solids: 1f noise. Rev Mod Phys 53: 497–516.

Edel, J.B., and Albrecht, T. (2013) Engineered Nanopores for Bioanalytical Applications: A Volume in Micro and Nano Technologies. Elsevier Inc., 1-172.

Fish, P.J. (2017) Electronic Noise and Low Noise Design. Macmillan International Higher Education, 1-30.

Fragasso, A., Schmid, S., and Dekker, C. (2020) Comparing Current Noise in Biological and Solid-State Nanopores. ACS Nano 14: 1338–1349.

Gamry (2021) Faraday Cage: What is a Faraday Cage-How Does a Faraday Cage Work? Gamry Instruments https://www.gamry.com/application-notes/instrumentation/faraday-cage/. Accessed October 24, 2021.

Hartel, A.J.W., Shekar, S., Ong, P., Schroeder, I., Thiel, G., and Shepard, K.L. (2019) High bandwidth approaches in nanopore and ion channel recordings - A tutorial review. Anal Chim Acta 1061: 13–27.

Hartlieb, P., and Grafe, B. (2017) Experimental Study on Microwave Assisted Hard Rock Cutting of Granite. BHM Berg- und Hüttenmännische Monatshefte 162: 77–81.

Hoogerheide, D.P., Garaj, S., and Golovchenko, J.A. (2009) Probing surface charge fluctuations with solid-state nanopores. Phys Rev Lett 102: 256804.

Jeon, S., Thundat, T., and Braiman, Y. (2006) Effect of normal vibration on friction in the atomic force microscopy experiment. Appl Phys Lett 88: 214102.

Khrais, N.N., Elrefaie, A.F., Wagner, R.E., and Ahmed, S. (1996) Performance of cascaded misaligned optical (De)multiplexers in multiwavelength optical networks. IEEE Photonics Technol Lett 8: 1073–1075.

Kowalczyk, S.W., Grosberg, A.Y., Rabin, Y., and Dekker, C. (2011) Modelling the conductance and DNA blockade of solid-state nanopores. Nanotechnology 22: 315101.

Li, B.Q., and Einstein, H.H. (2017) Comparison of Visual and Acoustic Emission Observations in a Four Point Bending Experiment on Barre Granite. Rock Mech Rock Eng 50: 2277–2296.

Liang, S., Xiang, F., Tang, Z., Nouri, R., He, X., Dong, M., and Guan, W. (2020) Noise in nanopore sensors: Sources, models, reduction, and benchmarking. Nami Jishu yu Jingmi Gongcheng/Nanotechnology Precis Eng 3: 9–17.

Motchenbacher, C., and Connelly, J. (1994) Low-noise electronic system design. Choice Rev Online 31: 2710.

Olson, H.F. (1956) Electronic Control of Noise, Vibration, and Reverberation. J Acoust Soc Am 28: 966–972.

Otani, T., Antoniades, N., Roudas, I., and Stern, T.E. (1999) Cascadability of passband-flattened arrayed waveguide-grating filters in WDM optical networks. IEEE Photonics Technol Lett 11: 1414–1416.

Ott, H.W. (1988) Noise Reduction Techniques in Electronic Systems. 2nd ed., John Wiley & Sons, 1-348.

Parkin, W.M., and Drndić, M. (2018) Signal and Noise in FET-Nanopore Devices. ACS Sensors 3: 313–319.

Rae, J.L., and Levis, R.A. (1992) A method for exceptionally low noise single channel recordings. Pflügers Arch Eur J Physiol 420: 618–620.

Shenoi, B.A. (2005) Introduction to Digital Signal Processing and Filter Design. 1st ed., Wiley-Interscience, 1-423.

Slichter, D.H., Naaman, O., and Siddiqi, I. (2009) Millikelvin thermal and electrical performance of lossy transmission line filters. Appl Phys Lett 94: 192508.

Smeets, R.M.M., Keyser, U.F., Dekker, N.H., and Dekker, C. (2008) Noise in solid-state nanopores. Proc Natl Acad Sci U S A 105: 417–421.

Srivastava, T., Virk, S., Kumar, A., and Ganguli, S. (2021) The role of electronic filters in biomedical applications. In Electronic Devices, Circuits, and Systems for Biomedical Applications. Academic Press, pp. 309–324.

Tabard-Cossa, V., Trivedi, D., Wiggin, M., Jetha, N.N., and Marziali, A. (2007) Noise analysis and reduction in solid-state nanopores. Nanotechnology 18: 305505.

Taqieddin, A., Nazari, R., Rajic, L., and Alshawabkeh, A. (2017) Review—Physicochemical Hydrodynamics of Gas Bubbles in Two Phase Electrochemical Systems. J Electrochem Soc 164: E448–E459.

Vasilescu, G. (2006) Electronic Noise and Interfering Signals. 1st ed., Springer-Verlag Berlin and Heidelberg GmbH & Co. K, 1-176.

Chapter 5

DNA translocation

Synopsis: In this chapter, DNA translocation will be examined more closely. A brief introduction to determining differently folded DNA structures is provided. One of the aims of this project was to observe and compare modified and unmodified DNA translocation. This chapter discusses the methods used to prepare, translocate, and analyse DNA translocation. This chapter also discusses the comparison between SCoNE and unmodified DNA.

5.1 Introduction

As initially discussed in Chapter 2, DNA translocation has been widely studied for a range of different purposes. DNA can translocate in a number of different conformations dependant on pore size. Due to the length and flexibility of DNA, these include linear, folded, or knotted conformations, as illustrated in Figure 5.1.A-C respectively (Steinbock *et al.*, 2010). In a linear conformation, the effect of the current induced, local electric field causes the DNA to unfurl and enter the pore, decreasing the current (5.1.A.2). As the DNA translocates through the pore, the structure continues to unfold allowing more of the DNA to pass through the pore until the entire fragment has translocated (Figure 5.1.A.3-5). Once translocated the current returns to baseline once again (Figure 5.1.A.6) (Edel and Albrecht, 2012).

DNA can fold along any point of the backbone; however, this often occurs at its terminal ends (Figure 5.1.B.2) (Edel and Albrecht, 2012, Kumar Sharma *et al.*, 2019). If the pore size is large enough to allow folded DNA to translocate, this will cause an additional decrease in current, in comparison to linear DNA (Figure 5.1.B.6). Once the folded section of DNA has translocated, the current returns to what would be expected for linear DNA (5.1.B.2-5). The translocation time is reduced comparable to the amount of DNA involved in the fold. The further the fold occurs along the DNA backbone, the shorter the event duration (Kumar Sharma *et al.*, 2019)

Knotting can also occur along the DNA backbone. This is where the DNA strands cross over one another in such a way that it cannot be unfolded during translocation. As the knot approaches the pore, it tightens until repulsion from the DNA backbone prevents further restriction. If the pore size allows for the knot to migrate through, it will translocate across the pore (Figure 5.1.C.3). Due to a fundamental limitation in time

resolution, it is not possible to detect all datapoints within the knot sub-event. Therefore, where folded DNA events appear square in shape, the time resolution limitation means that the sub-event appears triangular in shape. The other steps of the translocation follow the same pattern as linear DNA. Knots often take up less DNA length when created compared to DNA folding. Therefore, whilst they block more of the pore causing a decrease in current, these subevents are far shorter lived then their folded counterparts (Figure 5.1.C.6) (Kumar Sharma *et al.*, 2019).

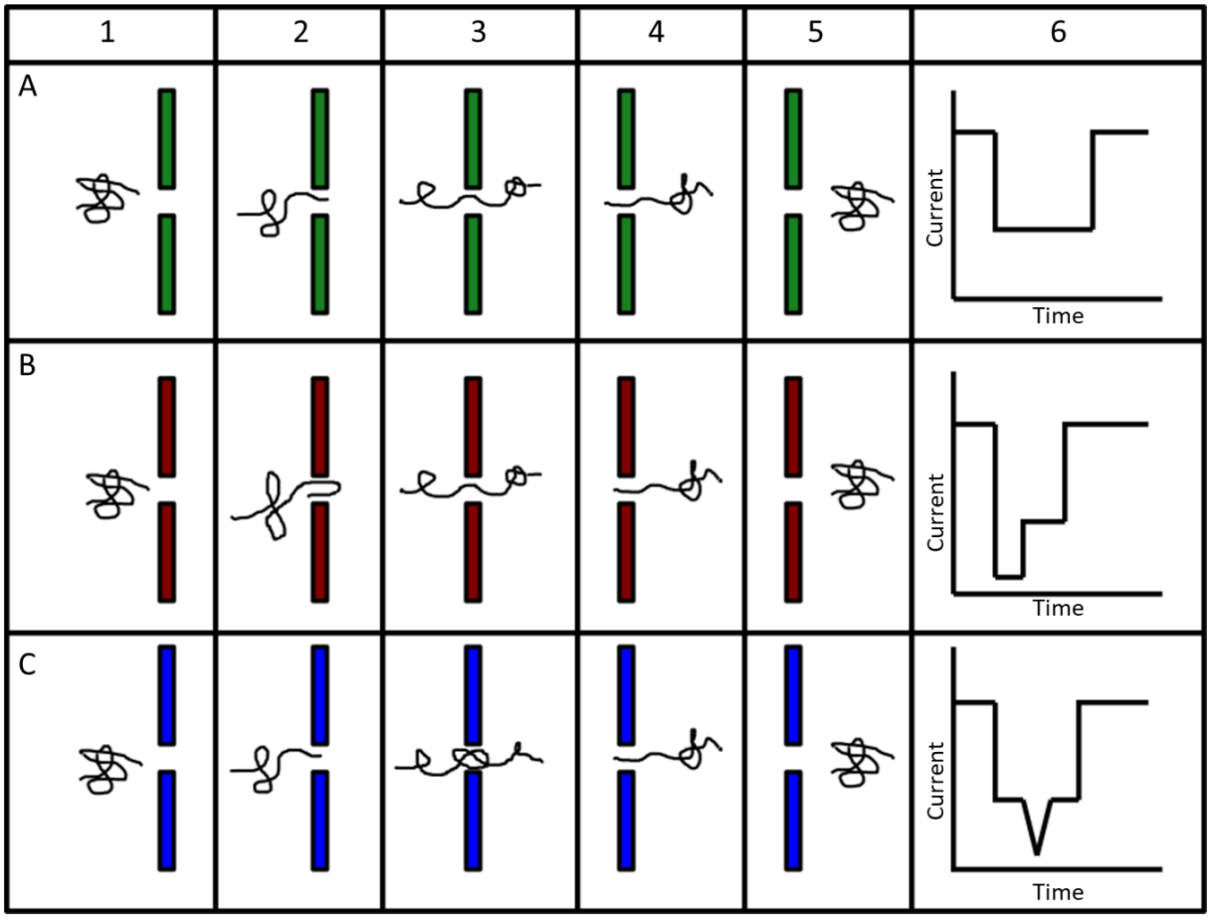

Figure 5.1. Translocation across a nanopore (1-5) for linear (A), folded (B), and knotted (C) DNA, with expected current event shapes (Edel and Albrecht, 2012).

In a study by Haque *et al.*, examining the folding pattern of 5 kbp through a phi29 gp10 DNA packaging protein channel, it was found that DNA translocated linearly 85% of the time, and folded 15% of the time (Haque *et al.*, 2015). When researching knotting

in lambda DNA (48.5 kbp), Sharma *et al.*, found that knotting occurred in 16% of the events observed (Kumar Sharma *et al.*, 2019).

The capability to study and analyse these subevents becomes significantly relevant when analysing modified DNA structures. In research conducted by Bell and Keyser, it was possible to detect DNA origami structures incorporated into the DNA backbone. It was found that by using a series of DNA dumbbell structures, it was possible to encode a 3-bit barcode along the DNA backbone and detect this using a nanopipette (Bell and Keyser, 2016). In further research conducted by Loh *et al.*, it was shown that DNA could be constructed to contain 88 nucleotide single stranded DNA overhangs. It was possible to use the single stranded overhangs to bind complementary strands. The difference between the bound and unbound struts could then be differentiated by nanopipette sensing (Loh *et al.*, 2018). Whilst these techniques offer structural changes to the DNA backbone allowing for detection using nanopore sensing, Chen *et al.,* used M.TaqI to enzymatically modify a dsDNA backbone to contain an 11-nucleotide overhang motif. Plasmid DNA with known enzymatic sites was used to create the structure, and the relative locations detected using nanopore sensing (Chen *et al.*, 2021).

Although each DNA modification method for nanopore sensing presents some advantages, the restriction of overhang location, and the preparation of the structures, limits their potential for use with biomolecule detection.

5.2 Methods

5.2.1 General preparations

Electrolytes was prepared as stated previously in section 3.2.1. Ag/AgCl electrodes were prepared as stated previously in section 3.2.2. Nanopipettes were prepared and fabricated as stated previously in section 3.2.3, and section 3.2.5 respectively, and pore size determined as stated previously in section 3.2.6.

The low noise amplifier and surrounding set up were as described in section 4.2.1, using anodized, 8 cm electrodes, section 4.2.2, silver shielded data cable where possible, section 4.2.4, on a granite surface, section 4.2.5, using battery power where possible, section 4.2.6, and passive filters where possible, section 4.2.7. Liquid cells were prepared as stated previously in section 3.2.6, nanopipettes prepared as stated in section 3.2.5, and pore size measured as stated in section 3.2.6. Some experiments were performed prior to acquisition of some of the items, and where applicable, this is noted in the results. During experiments several different bias voltages were used from -0.4 to -0.8 V, and 0.4 to 0.8 V. The control code was adapted due to work hour restrictions during the COVID-19 pandemic to include an automation script. Each full experiment can take between 10-12 hours to complete, and the university opening hours did not allow for this to be conducted in person. The adaptation contained code to interface with the user interface (UI) and automatically alter the bias applied at specific time intervals. The bias was changed every 20 minutes, equivalent to 104 files. Measurements were taken at varying filter frequencies for event detection as described in each results section.

Blank experiments were performed as described, without the addition of DNA into the liquid cell.

5.2.2 DNA preparation

DNA fragments (NoLimits™, ThermoFisher) of 4 kbp or 10 kbp (0.5 µg/µL) were pipetted into the liquid cell containing 2 mL of 4M LiCl 10% TE buffer to achieve a final concentration of 300 pM. Initial DNA concentration was calculated using equation 5.1.

$$[DNA(\mu M)] = \frac{\mu g/\mu l}{Length(bp) \times 660 g/mol} \quad (5.1)$$

5.2.3 Data Analysis

Custom code (written by author) was written to analyse the trends and characteristics of events across the range of biases used from the generated .ana file, as described in Chapter 4, and is available in Appendix 5. The code includes a visual representation of the events, k-means algorithm for data clustering, analysis of event duration dependant on bias, analysis of event current dependant on bias, and inverse bias, exact event finding and visualisation tool, average event shape visualisation, a subevent search tool (adapted from code courtesy of Tim Albrecht), and a subevent characteristics and analysis tool. Subevents were determined through the application of a threshold value, as determined by the user, and applied to the baseline of the DNA translocation event. Those peaks whose max current value crossed the threshold value were analysed further.

Initially the code contained, a visual representation of the events, analysis of event duration dependant on bias, and an analysis of event current dependant on bias. During the first second ($1x10^6$ datapoints) the amplifier ramps up to the voltage set by the user. Therefore, a threshold was used to remove data acquired before this point. Due to noise experienced during the experiments, as described in Chapter 4, the code was further adapted to include a current threshold to remove low intensity noise picked

up during event analysis. This is user defined, so each data set was treated individually. A k-means clustering algorithm, inbuild into MATLAB, was then applied to isolate DNA events, and their characteristics determined.

The ability to locate, isolate and visualise single events was then developed. Code was written to provide a user interface with a data cursor to select the event of interest. The event was then located within the mat.ana file and its starting and ending data point used as identifiers. These identifiers were then used to search through the mat.res files until a perfect match was found. The file number was then exported to the MATLAB workspace. Using this number, the entire scan with the single event highlighted was plotted and visual inspection performed. To visualise an average event, the clustered values for each bias were used and the durations normalised to the longest event. The current values for each data point were then averaged and a plot generated.

From the results of the adapted subevent search, further code was developed to analyse characteristics containing: the number of subevents per event; max current of the subevent; subevent duration; distance between subevents in data points and a conversion to number of DNA bases between subevents; starting point of the subevent in relation to the whole event in data points and base pairs; and a normalisation of the subevent location.

5.3 Results and discussion

Results were split into blank and DNA translocation experiments. Blank translocation experiments were carried out to determine baseline noise events picked up during data extraction from true DNA translocation results. The noise event intensity and duration results were then utilised during DNA translocation experiments, allowing for improved separation of the two event types. Example DNA translocation experiments are shown for discussion purposes.

5.3.1 Blank system

Experiments were performed using a silver shielded cable, using battery power and a 100 kHz passive filter. The pore size of the nanopipette used was ~18 nm.

Increase in noise intensity with an increase in bias is highlighted in Chapter 4. Illustrated in Figure 5.2 is an example of a blank translocation experiment conducted at -0.5 V. The current time plot generated shows a number of noise events picked up during event extraction. Whilst in section 4.3.7 it was possible to achieve an RMS noise of 24.0 ± 12.9 pA, noise remained a significant issue faced throughout all translocation experiments.

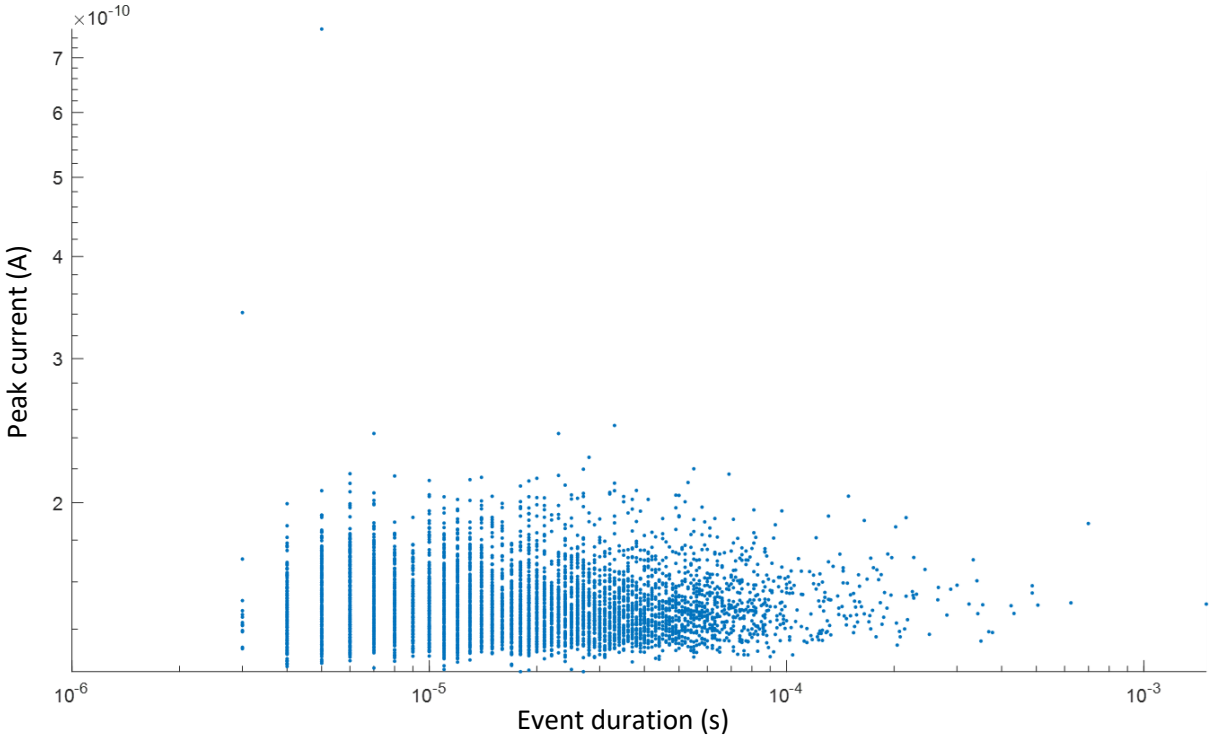

Figure 5.2. A blank translocation experiment highlighting noise events.

It is known that with increasing bias potential, there is an increase in noise. The results obtained agree with previous findings, but also show that there is an increase in singular noise peak occurrence. Without further data clustering, this could pose difficulty in data separation. Translocation events of both noise and analytes follow that as the bias voltage increases, the current decrease observed by these should increase also.

From the results of these experiments, it was possible to characterise and categorise noise peaks to develop clustering and analysis tools to exclude them from analyte translocation experiments. For analyte translocation experiments, manual sampling was performed to corroborate the results from the blank translocation results.

5.3.2 Translocation of 10kbp dsDNA

5.3.2.1 dsDNA translocation of 10 kbp at 10, 50, 100, 250 and 500 kHz filter frequencies

Initially a selection of filter frequencies was explored to determine the potential for event extraction at higher filter frequencies. The results from these translocation experiments shows that as filter frequency increases the noise decreases in duration, but increases in intensity, as illustrated in Figure 5.3. With increasing baseline noise, it becomes more difficult to detect DNA events, resulting in a loss of DNA translocation data points. From the analysis performed, filter frequency should be kept below 250 kHz due to data loss. While the higher frequencies allow for greater event resolution, it increases the difficulty in analysis and requires significantly increased acquisition time for the same number of events. With the research target being a reduction in diagnosis time, high filter frequencies will impede this. It was therefore decided that 100 kHz would be used for further experimentation as this provided an acceptable compromise between baseline noise, event detection, and event-baseline resolution.

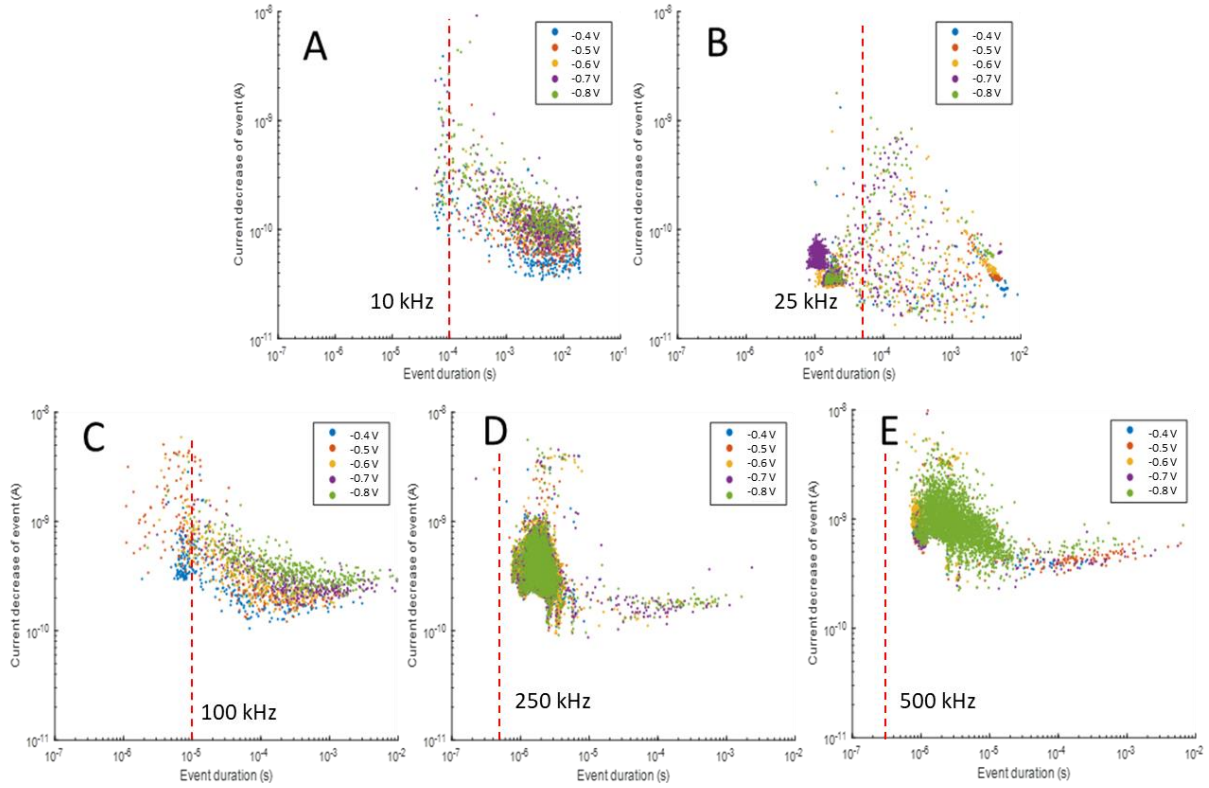

Figure 5.3. Scatter plot results of DNA translocation at 10 (A), 25 (B), 100 (C), 250 (D) and 500kHz (E), with 101 scans per bias (10 kbp [DNA] = 300pM in 2 mL 4M LiCl 10% TE buffer). The red dotted line represents the filter frequency cut off in the time domain.

Experiments were performed using a standard printer cable, using mainline power and the analogue filter. The size of the nanopipettes used were ~15, ~22, ~19, ~12, and ~17 nm respective to increasing filter frequency. The most promising results from this exploration were the results from the 10 and 100 kHz filtered translocation experiments. These results were then analysed further.

5.3.2.2 dsDNA translocation of 10 kbp using a 10 kHz filter

Experiments were performed using a standard USB B data transfer cable, using mainline power and the active filter. The size of the nanopipette used was ~15 nm. DNA translocation experiments were performed to corroborate previous research performed in the group, whilst optimising noise. The difference in event magnitude, as illustrated in Figure 5.4, demonstrates the overall decrease in event duration with bias.

Although optimized as far as possible, noise continued to be an issue. Further extraction tools were used to isolate DNA translocation events. The difference in event duration was used as one of the extraction methods, converting the event durations into a logarithmic value and utilizing histograms for each bias. A Gaussian fit was applied to the frequency counts data for the log time to extract the most likely DNA translocation time, as illustrated in Figure 5.5. The maxima of the Gaussian fit peaks were used as the most probable translocation time. These values were converted back to seconds and a plot generated showing the change in both intensity and duration, as illustrated in Figure 5.6.A and B respectively.

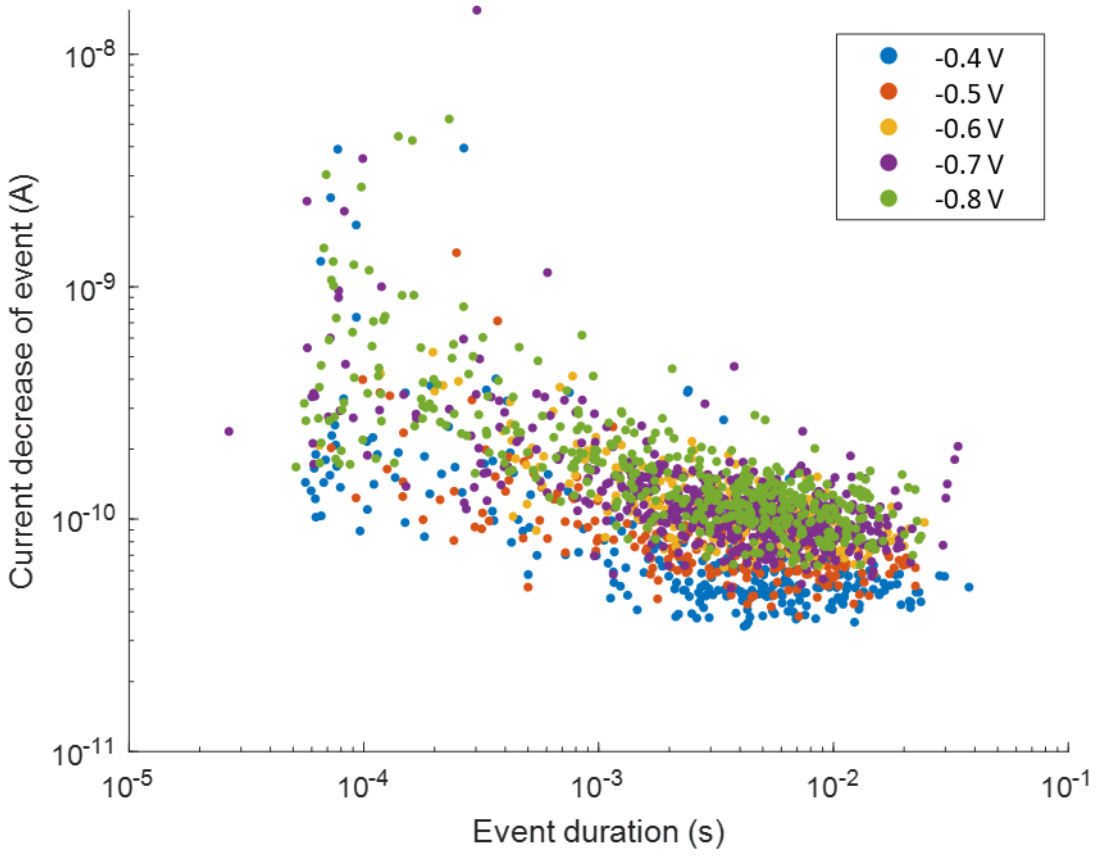

Figure 5.4. Event extraction results from 505 10s-scans (101 per bias) illustrating the dependence of event duration and current decrease by bias (filter frequency: 10 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

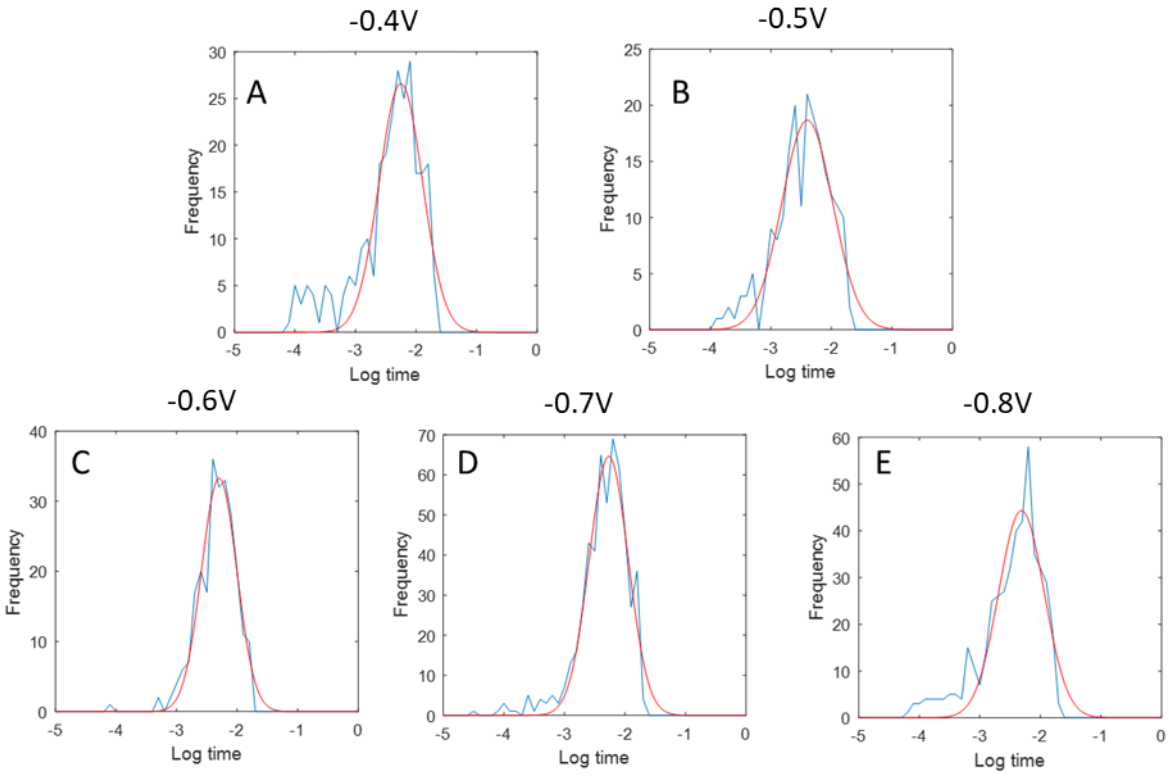

Figure 5.5. One dimensional histogram plots of the extracted event data showing results for biases from -0.4 to -0.8V (A-E respectively) with associated R^2 values of 0.9197, 0.8836, 0.9477, 0.9657, and 0.9095 respectively.

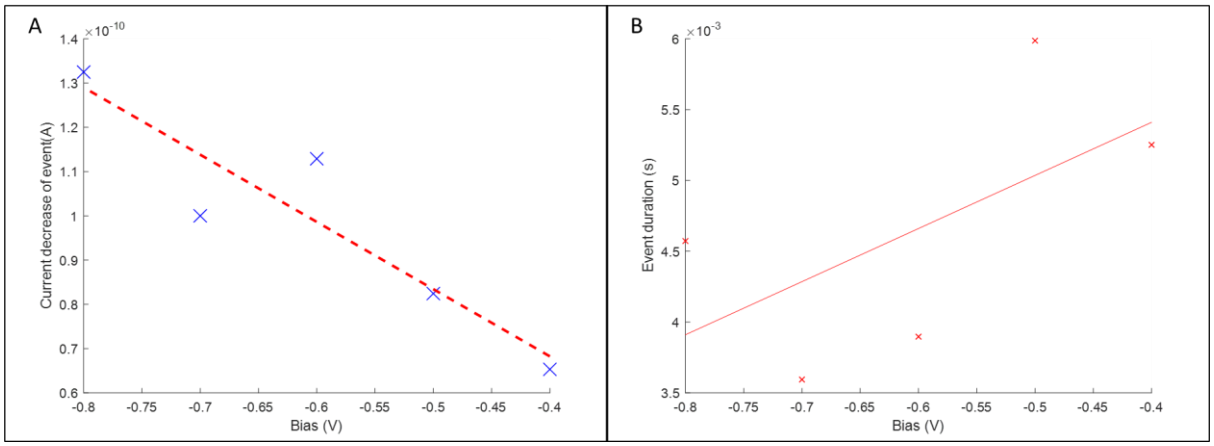

Figure 5.6. The results from double gaussian fitting results illustrating the increase in intensity of translocation peaks with bias (A) and the decrease in event duration with an increase in bias (B), (filter frequency: 10 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

The number of events per bias was also extracted from the Gaussian fit results to determine the translocation frequency. Figure 5.7 shows the increase in translocation frequency with an increase in bias. Due to the low translocation frequency experienced with biases less than -0.5 V, the Gaussian fitting can include noise results within the fit, accounting for the increased translocation frequency observed. When accounting for the difference in conductivity between the electrolyte used for these experiments and those used in literature, the increase in translocation frequency with bias are equivocal to literature results, with the achieved translocation frequency of ~1 per second per nM at -0.8 V (Fraccari *et al.*, 2016; Bell and Keyser, 2016; Loh *et al.*, 2018).

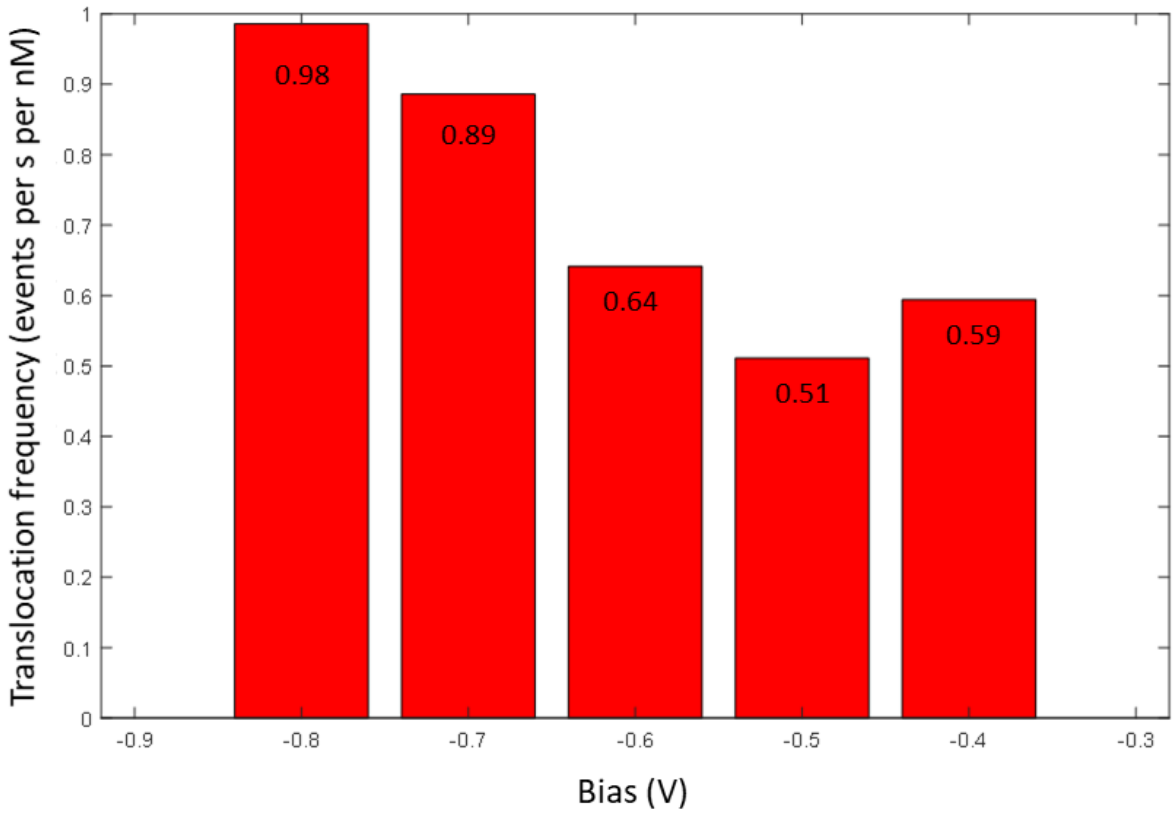

Figure 5.7. Translocation frequency results obtained from applying a single Gaussian fitting results highlighting an increase in translocation frequency with bias, (filter frequency: 10 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

Examples of the DNA translocation events are highlighted in Figure 5.8. These are suggested from manual inspection of singular translocation events. From these examples, it is possible to see DNA translocating in several different conformations. Figure 5.8.A illustrated DNA translocating in a folded conformation, where 5.8.B demonstrates DNA translocating in a linear fashion. Both C and D show examples of knots in the DNA. Figure 5.8.C highlights knots appearing near the beginning and end of the DNA strand, where D shows these knots forming at the beginning, and middle of the strand. The colour of the translocation events highlighted is used as a tool during analysis to separate DNA events from the backbone. It is not representative of any significance.

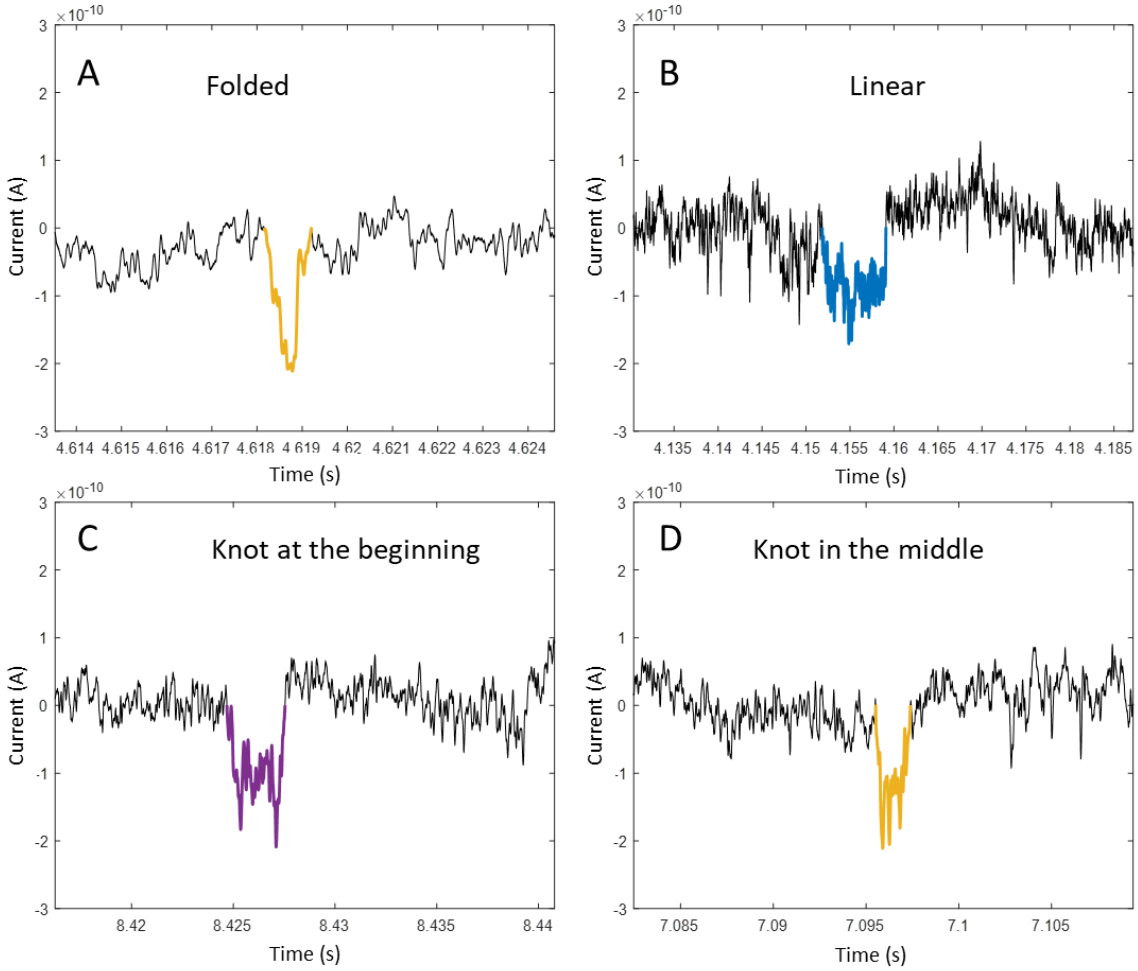

Figure 5.8. 10kbp DNA translocation examples showing folded (A), linear (B), and potential knotted events (C, D) (filter frequency: 10 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

5.3.2.3 dsDNA translocation of 10 kbp using a 100 kHz filter

Experiments were performed using a standard USB B data transfer cable, using mainline power and an active filter. The size of the nanopipette used was ~19 nm. Increasing the filter frequency, increases the resolution of the DNA translocation events, however, this also increases the noise observed during experimentation. To determine the effect on observable DNA translocation, the filter frequency was increased to 100 kHz. From the results of the 100 kHz filtered experiments, the same trend of increased translocation peak current intensity and a decrease in event duration with an increase in bias, is observed as in the 10 kHz filtered experiment, Figure 5.9. Individual histogram analysis and peak maxima extraction confirm the observation, as illustrated in Figure 5.10, and 5.11 respectively. It can also be noted that the histogram distributions also show an increase in noise peaks appearing in the extracted data. This shows that with increasing filter frequency, the intensity of the noise peaks increases similarly, and increases the difficulty in DNA translocation event extraction. The translocation frequency results are also comparable to those noted in both the 10 kHz filtered experiments and previous published work, Figure 5.12.

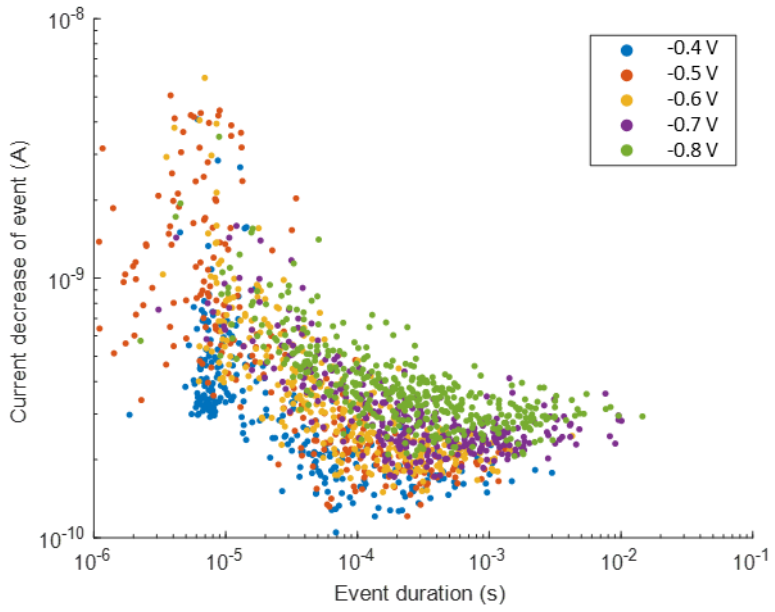

Figure 5.9. Event extraction results from 505 10s-scans (101 per bias) illustrating the dependence of event duration and current decrease by bias (filter frequency: 100 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

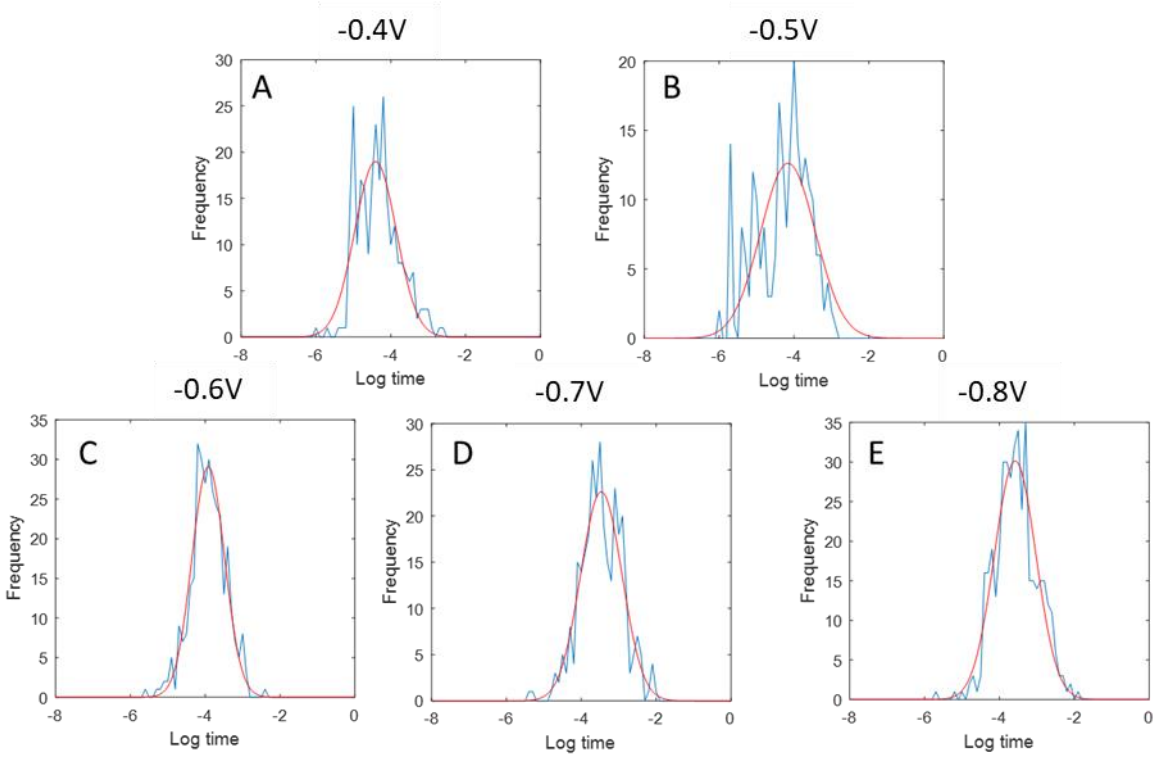

Figure 5.10. One dimensional histogram plots of the extracted event data showing results for biases from -0.4 to -0.8V (A-E respectively) with associated R2 values of 0.9535, 0.8035, 0.8816, 0.8216, and 0.9239 respectively.

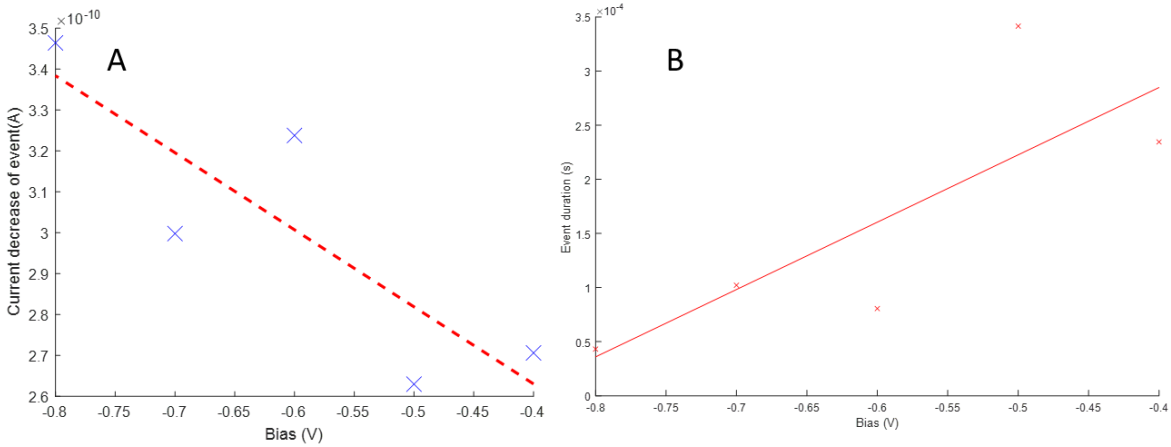

Figure 5.11. The results from Gaussian fitting results illustrating the increase in intensity of translocation peaks with bias (A) and the decrease in event duration with an increase in bias (B) (filter frequency: 100 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

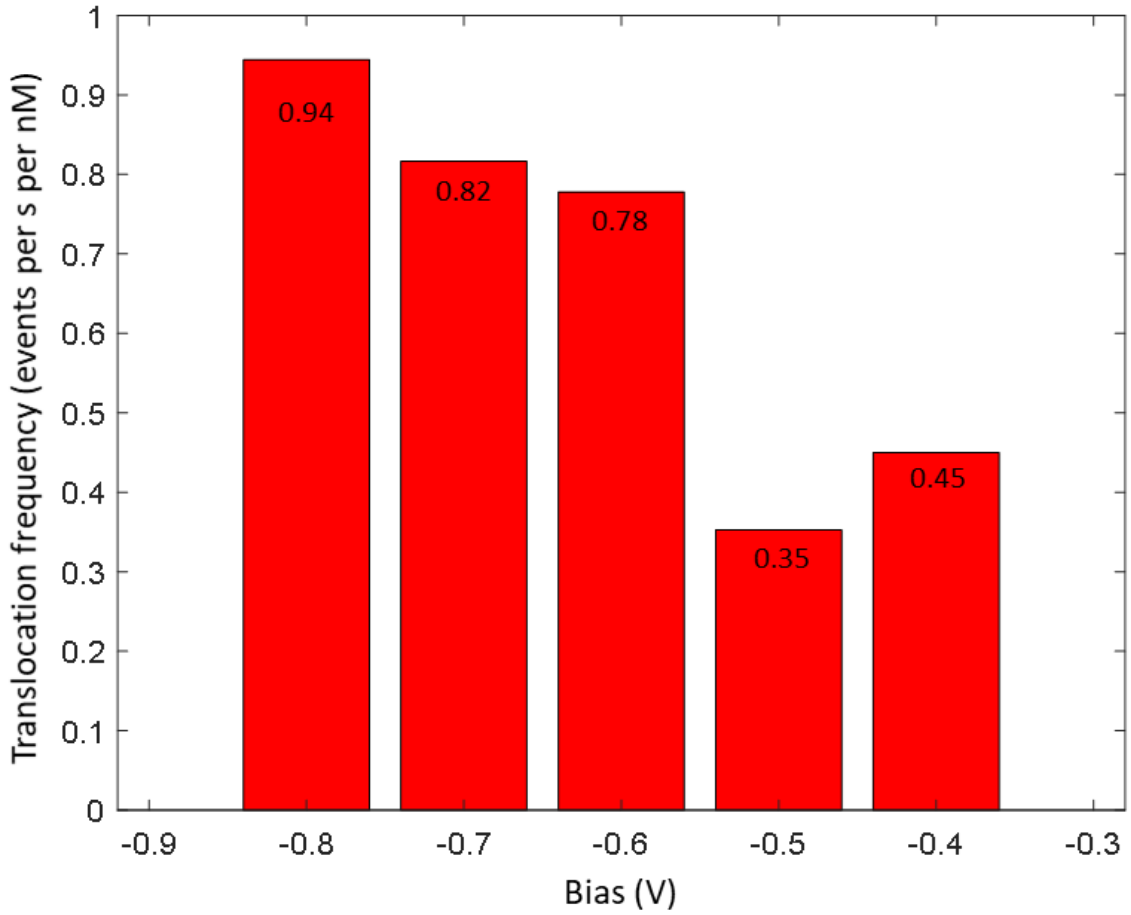

Figure 5.12. Translocation frequency results obtained from a single Gaussian fitting results highlighting an increase in translocation frequency with bias (filter frequency: 100 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

Examples of the DNA translocation events are highlighted in Figure 5.13. From these examples, it is possible to see DNA translocating in several different conformations. Figure 5.13.A and B demonstrates DNA translocating in a linear fashion, D shows an example of folded DNA translocation, and C shows and example of double folded DNA translocation.

The results from the 100 kHz filtered experiment show similar trends in increasing event intensity and decrease in event duration with increasing bias. The intensity of the 100 kHz is much greater (average across all biases 3.128X) than that acquired in the 10 kHz experiments. This indicates that as the filter frequency is increased, the signal acquired is also increased and prevents amplitude degradation, observed at low

filter frequencies (Shekar *et al.*, 2016). The event duration at 100 kHz is also much less than the 10 kHz experiments (average across all bias's 4.078X less). This indicates that it is possible to resolve DNA translocation results that were previously obscured due to the filter applied (Shekar *et al.*, 2016). However, these results also highlight the inclusion of more noise due to the increased bandwidth. The similarity between translocation frequencies indicates that, although translocation events can be more easily resolved, the same number of DNA fragments are passing through the pore per second in both experiments. This highlights the reliability of the analysis and reproducibility of the experimental set up.

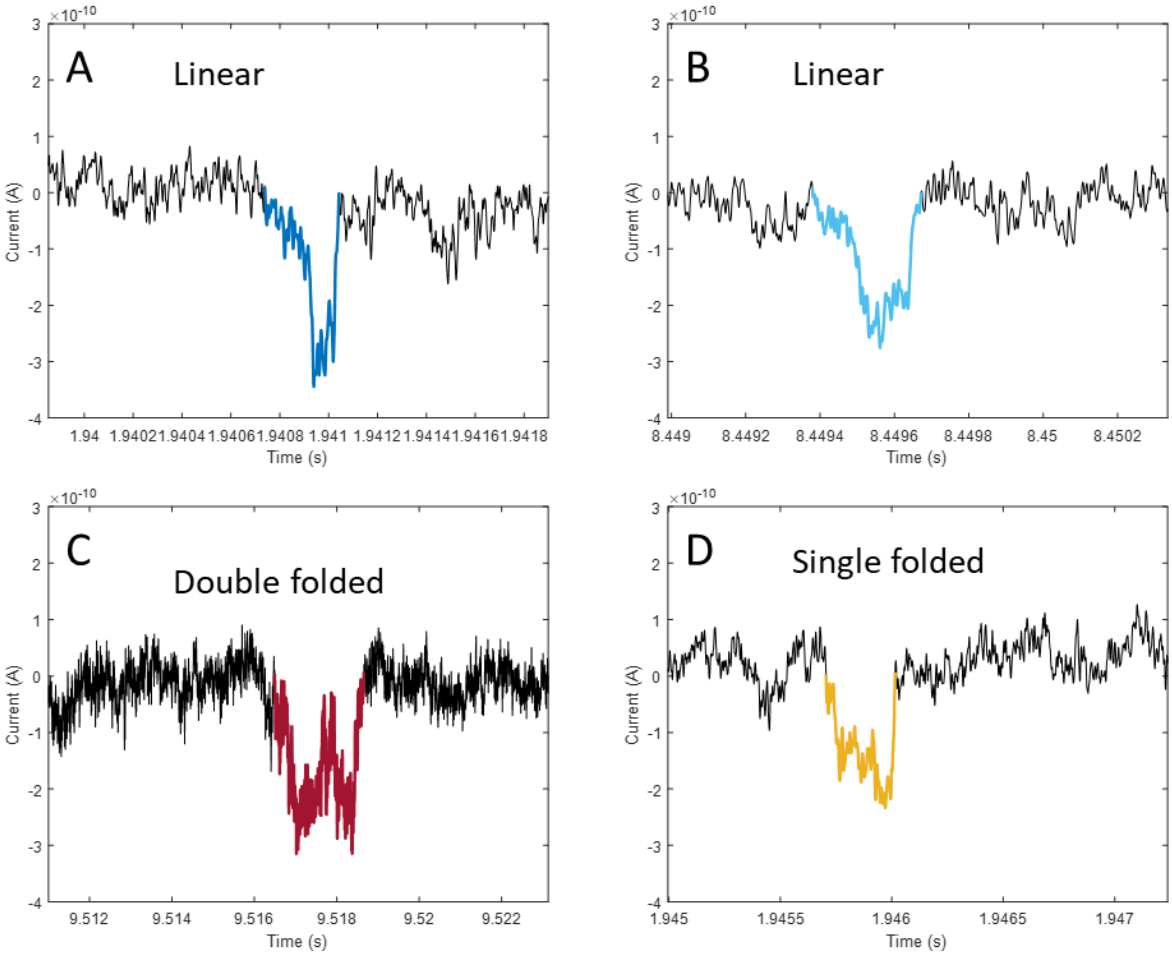

Figure 5.13. 10 kbp DNA translocation examples showing linear (A and B), double folded (C), and single folded events (D) (filter frequency: 100 kHz; 10 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

5.3.3 Translocation of 4kbp dsDNA

Experiments were performed using a silver shielded USB type B cable (Pearl, AudioQuest), using battery power and a passive 100 kHz filter. The size of the nanopipette was ~21 nm. Translocation experiments were performed to corroborate previous findings published by members of the group. In the experiment conducted, additional noise features were observed in the high intensity region, but with a more typical event duration, as illustrated in Figure 5.14. Unfortunately, the noise experienced in the -0.4 V bias prohibited the acquisition of useable data points and has therefore been excluded from the analysis. This noise cluster is also observable in the -0.7 V bias results; however, this did not prohibit extraction of DNA translocation events. From this, it is possible to suggest that the -0.4 bias had limited DNA translocation results for extraction, as is consistent with other experiments.

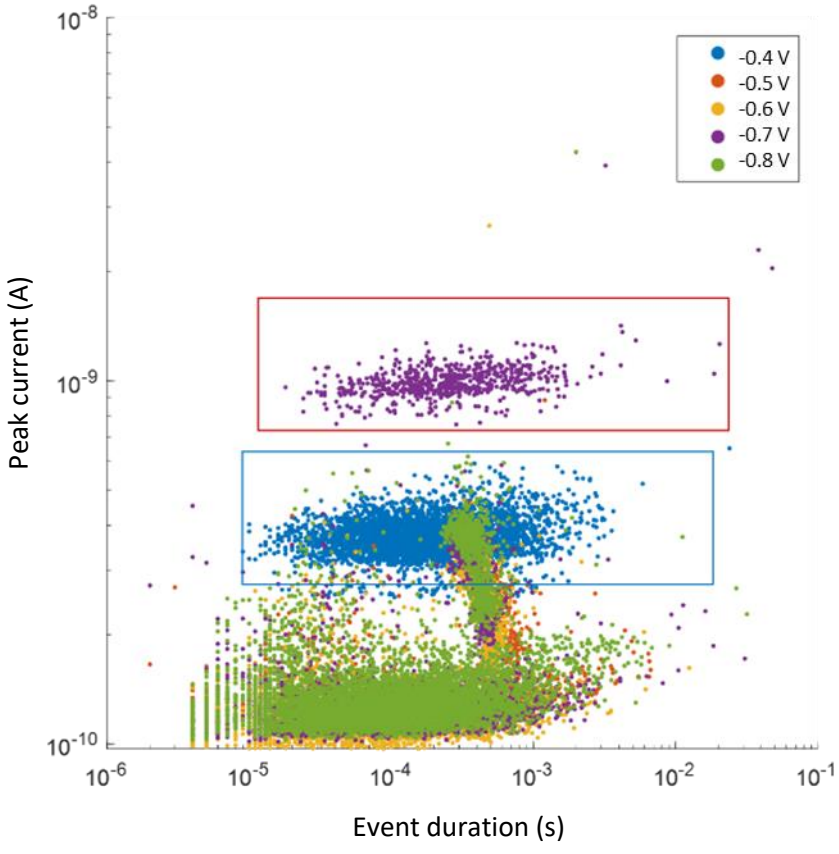

Figure 5.14. Event extraction results from 505 scans (101 per bias) illustrating the dependence of event duration and current decrease by bias. Additional noise features for -0.4V (blue box) and -0.7V (red box) are highlighted (filter frequency: 100 kHz; 4 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

Translocation results from 4kbp dsDNA, show characteristic patterns for both intensity and event duration as illustrated in Figure 5.15.B, C and D.

As the bias increases the event duration decreases, and event intensity increases. Due to the similarity between DNA event translocation and noise events, it was not possible to apply a Gaussian fit approach for data separation as was previously done. Thresholding was applied with an upper and lower limit to isolate DNA events from the surrounding data, as illustrated in Figure 5.15.A. This proved to be a highly effective method for data separation and was incorporated into future analysis.

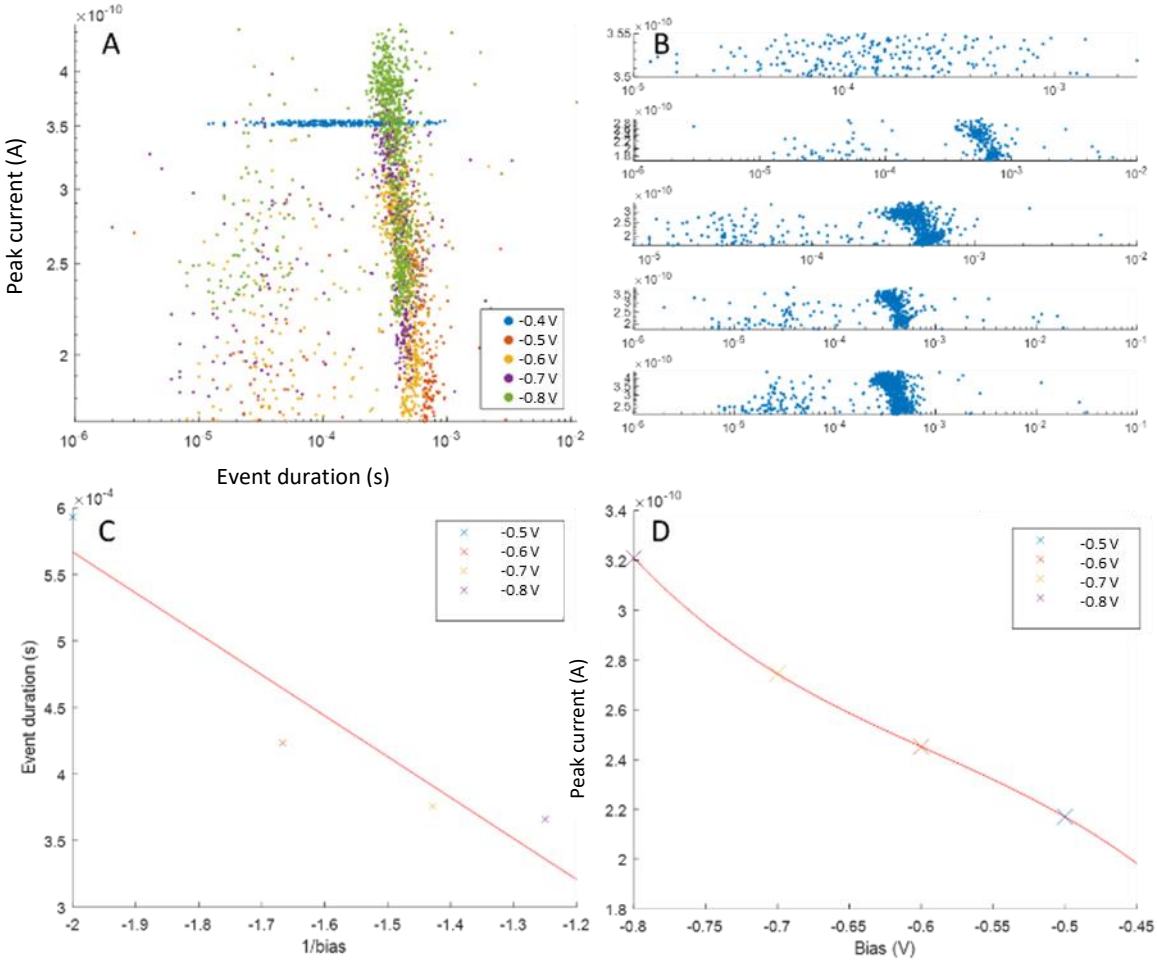

Figure 5.15. DNA translocation isolation results (A) from 505 scans (101 per bias), with the bias event profiles shown for clarity (B) with DNA events extracted highlighted (green box), illustrating the dependence of event duration (C) and current decrease (D) by bias, (filter frequency: 100 kHz; 4 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

Single event extraction and analysis highlights both similarities and differences between translocation experiments. Figure 5.16 illustrates five examples of 4 kbp DNA translocation events. Ten events from each bias were selected at random for single event analysis and displayed a significantly higher proportion of linear translocation events in comparison to 10 kbp DNA. It is possible to suggest that due to the increased length of the 10 kbp DNA, the increased flexibility of the DNA allowed for more folding to occur when entering the nanopore. This is further enhanced by the pore for the 4 kbp DNA translocation being larger than for the 10 kbp DNA translocations at 10 and 100 kHz. A smaller pore should encourage more linear DNA translocation as opposed to folded DNA configurations.

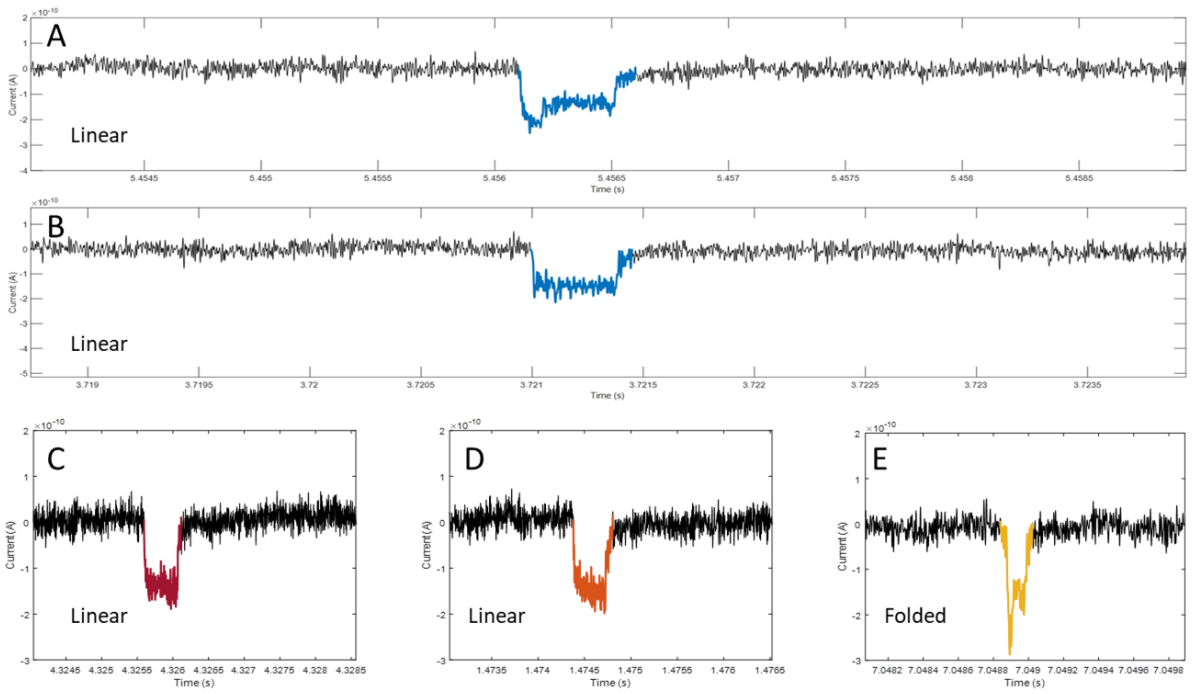

Figure 5.16. Examples of DNA translocation events with linear (A-D) and folded (E) conformations (filter frequency: 100 kHz; 4 kbp [DNA] = 300 pM in 2 mL 4 M LiCl 10% TE buffer).

5.3.4 SCoNE DNA translocation

During translocation experiments, two separate SCoNE structures were investigated, a tetramer and a decamer.

Tetramer translocation and analysis was performed after the development of the streptavidin bead extraction method. The structure of the tetramer used was $4S2P_{3,4}2I_{1,2}$, as highlighted in Figure 5.17, contained biotin isolation groups at positions 1 and 2, with a bound IL-6 aptamer at position 3 and an unbound procalcitonin aptamer at position 4.

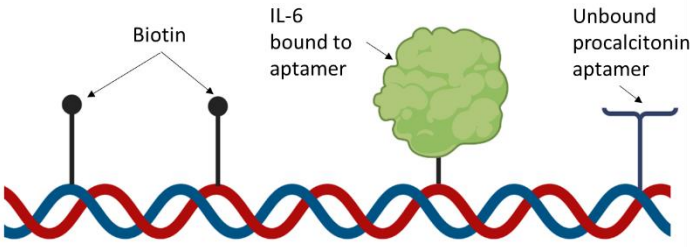

Figure 5.17. Illustration of the SCoNE tetramer structure with biotin groups at position 1 and 2 (ball ended struts), IL-6 aptamer bound at position 3 (green protein shape), and unbound procalcitonin at position 4 (Y shaped strut).

From the results of the translocation experiment, it was possible to observe SCoNE translocation events, as illustrated in Figure 5.18.A. From the results, it was then possible to apply a threshold allowing separation of the SCoNE data cluster from the noise cluster, Figure 5.18.B. These clustered results were then analysed based on bias dependence and compared against bare 4 kbp DNA translocation.

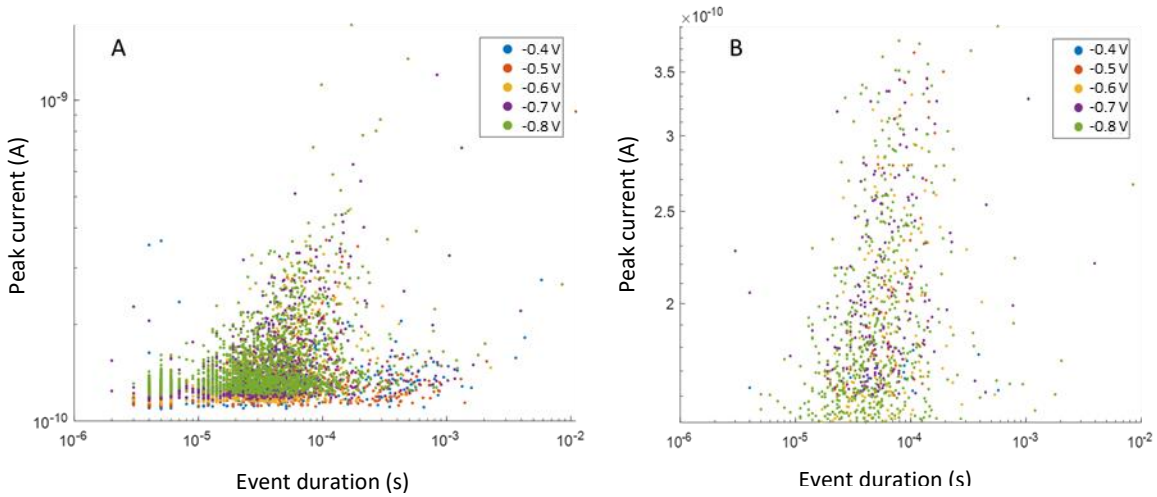

Figure 5.18. Translocation experiment of a tetramer SCoNE structure (A) with extraction of the SCoNE events from noise events (B) where event duration is compared against maximum intensity.

From the analysis of bias dependence, it is possible to conclude that SCoNE structures behave as expected with event duration decreasing with increasing bias. The presence of the probe strand, or the bound protein does not affect the translocation properties of the DNA, as illustrated in Figure 5.19.

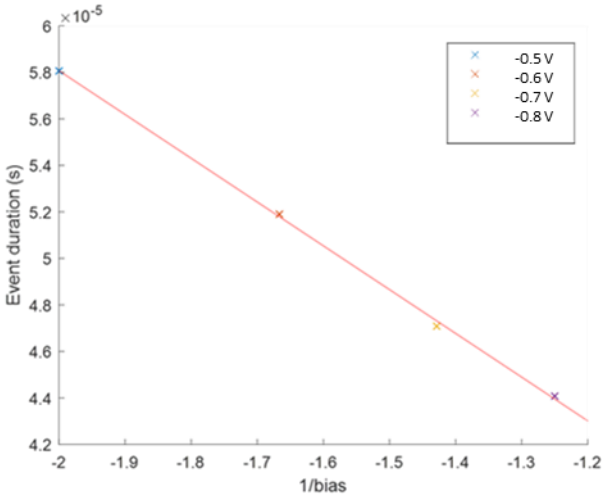

Figure 5.19. Inverse relationship between bias and event duration for tetramer SCoNE structures

Subevent analysis provides a further insight into the substructure of the events as highlighted in Figure 5.20. A subevent threshold between 25 - 175 pA was applied to remove non-specific or noise-based interference (5.20.B). DNA has the capability to translocate both forwards and backwards through the nanopore. In Figure 5.20, the length of the backbone was normalised to values 0-1. In forward translocation, expectations would be to observe subevent peaks at or near relative positions 0, 0.25, 0.5 and 0.75, relating to the 2 iDNA probes, followed by the bound IL-6 aptamer then finally the unbound procalcitonin aptamer. In a backwards translocation, expectations would be to observe sub peaks at or near relative positions 0.25, 0.5, 0.75 and near 1, reversing the order of the probes. As can be seen from the plot (5.20.A), there are several peaks which emerge from the events, with a 25 pA applied threshold. Notably the subevent peaks in the -0.5 V bias data had a significantly reduced number of translocations, and therefore the peaks visualised may not represent the overall substructure. It was determined that the -0.6 V bias data would be used for further comparison as this had the most events to analyse, providing a more representable dataset. In lower thresholds (25 - 75 pA, 5.20.B) peaks can be observed at near 0,

near 0.25, 0.5, near 0.75 and some emergence near 1. Due to the size of the biotin binding groups and the unbound aptamer, it is possible that during data acquisition some of the resolution near the beginning and end of the event is lost. This indicates that a loss of groups at positions 0 and 1 could be likely. When a subevent threshold between 100 - 150 pA is applied, it is possible to observe a loss of these initial, and ending peaks, however this allowed for a greater resolution of subevents at positions 0.25 and 0.5. This is indicated by a narrowing of the subevent peaks. The increased frequency of occurrence at position 0.5 coincides with the position of the bound IL-6 aptamer. Due to the decrease in current created when translocating larger analytes, this provides a good indication that the bound probe is present, generating a subevent peak which stands out from the baseline. Above a threshold of 150 pA only the peak at 0.25 remains, indicating that potential subevents have been removed. Using the 25 – 75 pA thresholds applied to both tetramer and 4kbp DNA allowed for the comparison of subevent features. It is possible to observe that the bare DNA creates a smooth shape across the event, lacking any defined subevent peaks (5.20.C). The presence of the subevent peaks at approximately the expected positions further supports the success of the assembly method. Due to differences in starting concentration, frequency counts have been normalised against total counts for direct comparison.

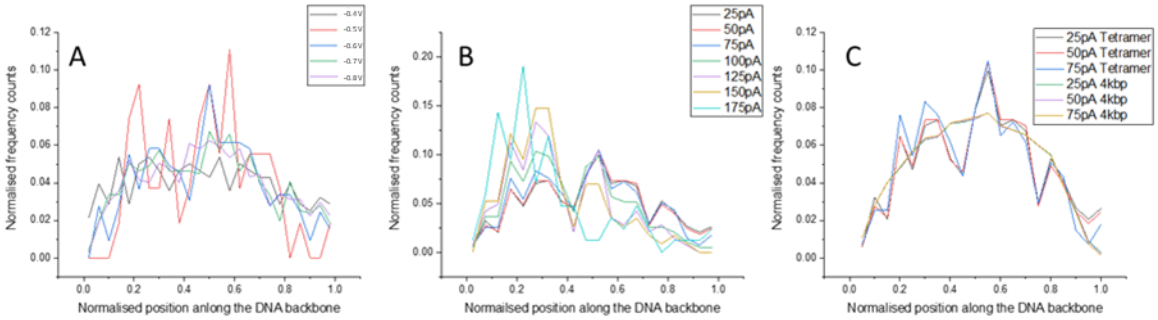

Figure 5.20. Comparison of average subevents analysis conducted between tetramer SCoNE DNA (N = 806) and bare 4 kbp DNA (N = 1244) fragments. The subevent position along the DNA backbone at different applied biases (A). A comparison between subevent threshold (B) shows that increasing threshold value decreases the number of peaks observable (25 – 175 pA thresholds respectively). A comparison between tetramer SCoNE structures and bare 4 kbp DNA fragments shows a lack of additional subevent peaks in the bare DNA samples represented in 2D for direct analysis (C).

When comparing singular events, it is possible to see the trends observed in the subevent analysis. Figure 5.21 highlights example events (excluding A, E, and I) observed in the initial scans (A, E, and I) across three of the biases used during experiments (-0.5, -0.6 and -0.7 V respectively). These show a clear definition of subevents appearing from the baseline as opposed to noise contribution. Whilst the overall shape of the events remains square, the short subevent peaks project from the base of the event at a higher amplitude then the noise. This is further emphasised when comparing events to bare 4 kbp DNA translocation events (Figure 5.22, excluding A, E, and I). Experiments were conducted using nanopores of similar size (19 and 21 nm respectively), with the respective DNA differences. It can also be noted that translocation frequency is much higher in bare 4 kbp DNA, this is due to a higher starting concentration. Short subevent peaks in these events mimic the noise amplitude observed at the scan baseline.

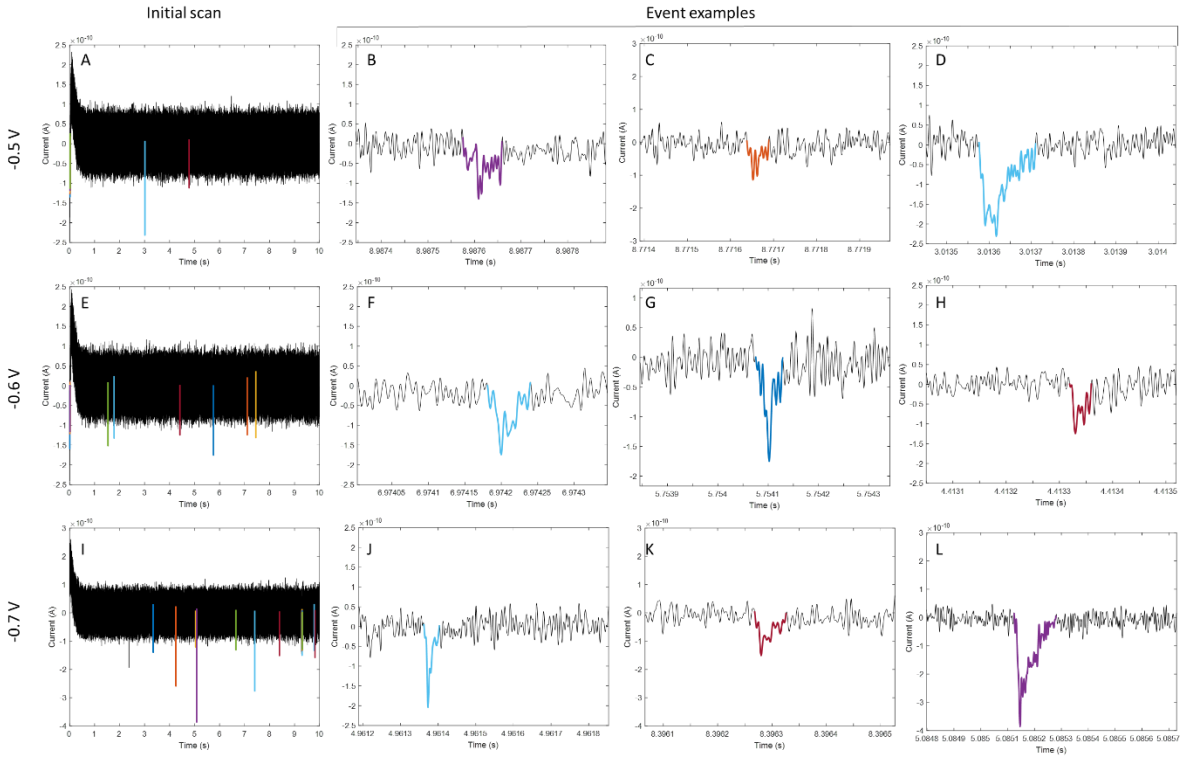

Figure 5.21. Examples of tetramer SCoNE DNA current-time traces (A, E and I) for -0.5, -0.6, and -0.7 respectively, with single events highlighted for each bias (B-D, F-H, and J-L).

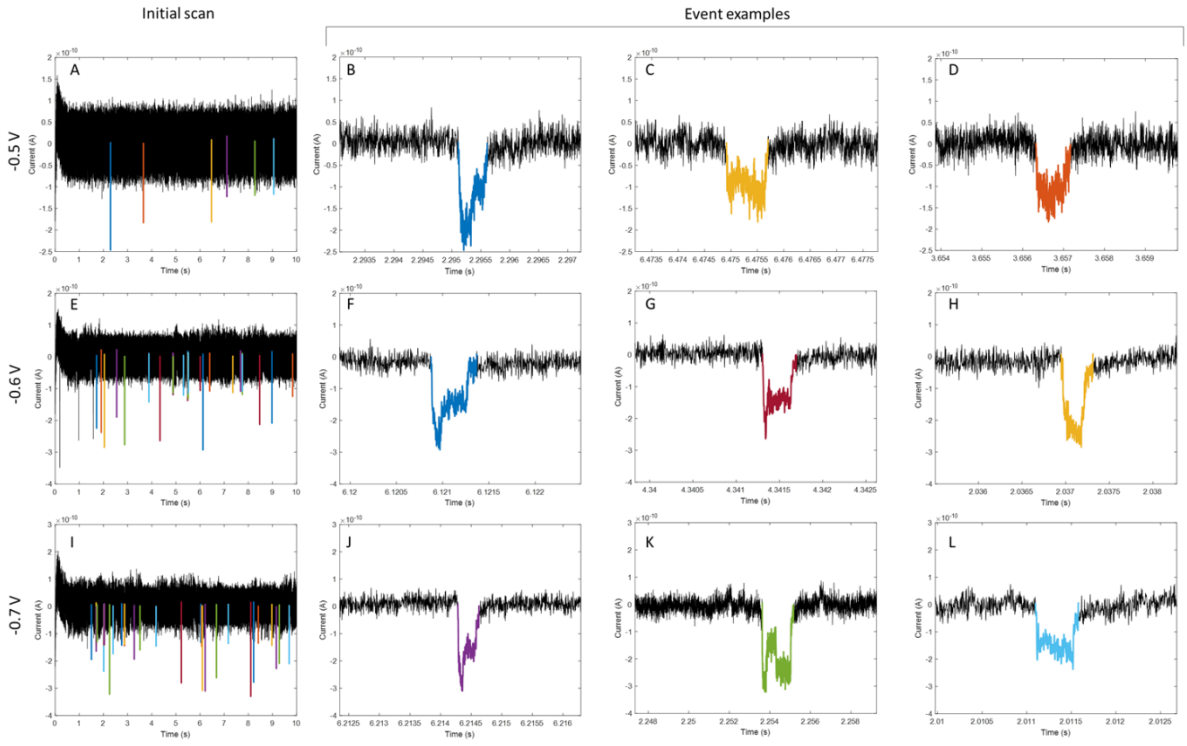

Figure 5.22. Examples of bare 4kbp DNA current-time traces (A, E and I) for -0.5, -0.6, and -0.7 respectively, with single events highlighted for each bias (B-D, F-H, and J-L).

The decamer translocation experiment was performed prior to the streptavidin bead extraction method developed, therefore gel extraction was used to obtain the SCoNE structure, as described in section 2.2.4.1. The structure of the decamer, as illustrated in Figure 5.23, used alternating IL-6 (black structures) and procalcitonin aptamers (red structures). The experiment was performed such that all probes were unbound, to determine if the aptamer substructures could be detected.

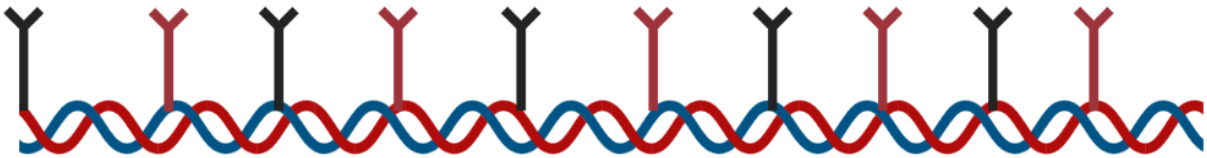

Figure 5.23. Illustration of the SCoNE decamer structure with IL-6 (black) and procalcitonin (red) aptamers.

From the results of the translocation experiment, it is possible to observe SCoNE decamer translocation results, as is illustrated by Figure 5.24.A. The emergence of 2 clear clusters can be observed, most notably in the -0.4 and -0.8 V biases, showing clear separation between SCoNE structures and noise peaks. From the translocation results, it was then possible to apply a threshold to the results to isolate the SCoNE structure data points, as is illustrated by Figure 5.24.B. These results were then analysed and compared against bare 10 kbp DNA translocation results.

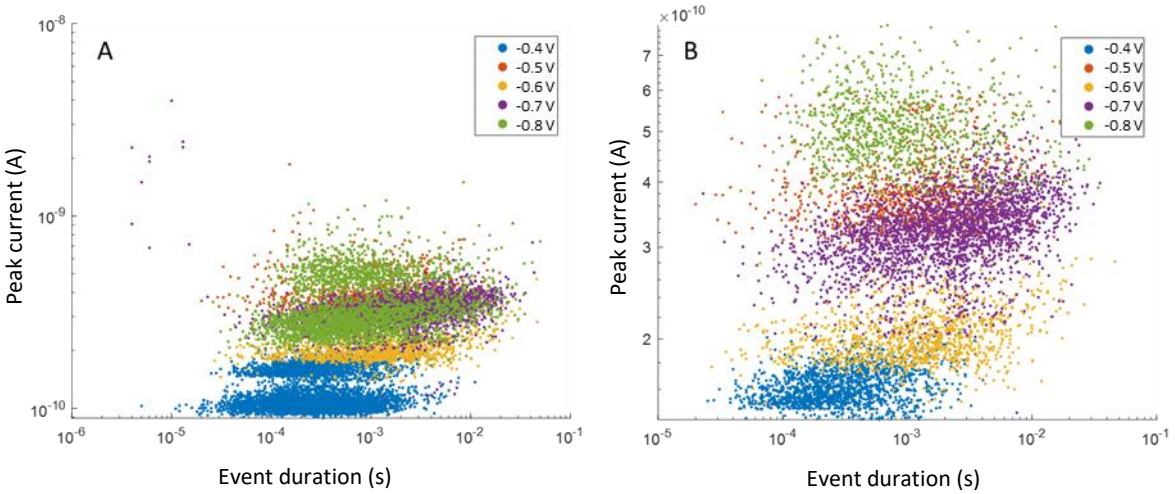

Figure 5.24. Translocation experiment of a decamer SCoNE structure (A) with extraction of the SCoNE events from noise events (B) where event duration is compared against maximum intensity.

From the analysis of inverse bias dependence, it is possible to see that there is an overall decrease in event duration with increased bias, as illustrated in Figure 5.25. It is unknown why, at -0.7 V, increased event duration is observed. Due to this, it was not possible to apply a linear line of best fit. The presence of the probe group attachment could be influencing the translocation of the SCoNE structure.

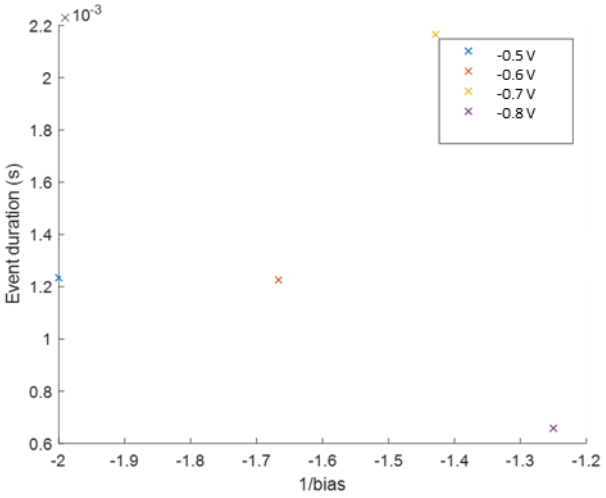

Figure 5.25. Inverse relationship between bias and event duration for decamer SCoNE structures

Due to the gel extraction method used to prepare the SCoNE DNA, it is possible to suggest that the gel solubilisation solution also caused degradation of the folded probe structures, or the bonding between the probe and the DNA backbone. This would limit the potential to detect substructures and render the structure more similar to bare 10kbp DNA.

Analysis of the subevents from the decamer SCoNE structures provided further insight into previous findings. Subevents were analysed at a 25 pA threshold for decamer SCoNE structures and 10 kbp fragments, Figure 5.26.A and B, respectively. In Figure 5.26, the length of the backbone was normalised to values 0 - 1 with the relative positions of sub structures falling between these values. In a forward translocation, expectations would be to observe events at, or near, values between 0 and 0.9, at increasing 0.1 values. In a backwards translocation, expectations would be to observe subevent peaks at, or near, relative positions between 0.1 and 1, at increasing values of 0.1. As can be seen in the SCoNE subevent analysis plot (5.26.A), there are several peaks which emerge from the events generated. The peaks observed in the -0.4 V bias have increased normalised frequency counts with all backwards translocation peaks noticeable (0.1 to 1). This dataset had fewer translocation events then the other biases, contributing to the difference in normalised frequency counts. When a 25 pA threshold was applied to the 10 kbp DNA, subevents can also be observed, however, the locations are more erratic than those observed in the decamer data (5.26.B). From the generated 2D line plot, Figure 5.26.C, it can be observed that similar profiles are generated for decamer SCoNE translocations at -0.6, -0.7 and -0.8 V biases, with significant overlap between the -0.7 and -0.8 V lines. A comparison of subevent profiles between decamer SCoNE structures and 10 kbp bare DNA at -0.6 V, Figure 5.26.D, highlights that there is little difference between the two DNA types. Whilst it is possible to suggest that there are probe groups present on the SCoNE DNA structure from individual analysis, there is not a significant difference between the profiles, making separation of the two groups difficult. This observation could be caused due to degradation of the probe groups during extraction. Further experimentation would be

required, incorporating iDNA fragments into the assembly of a decamer and repeating the translocation.

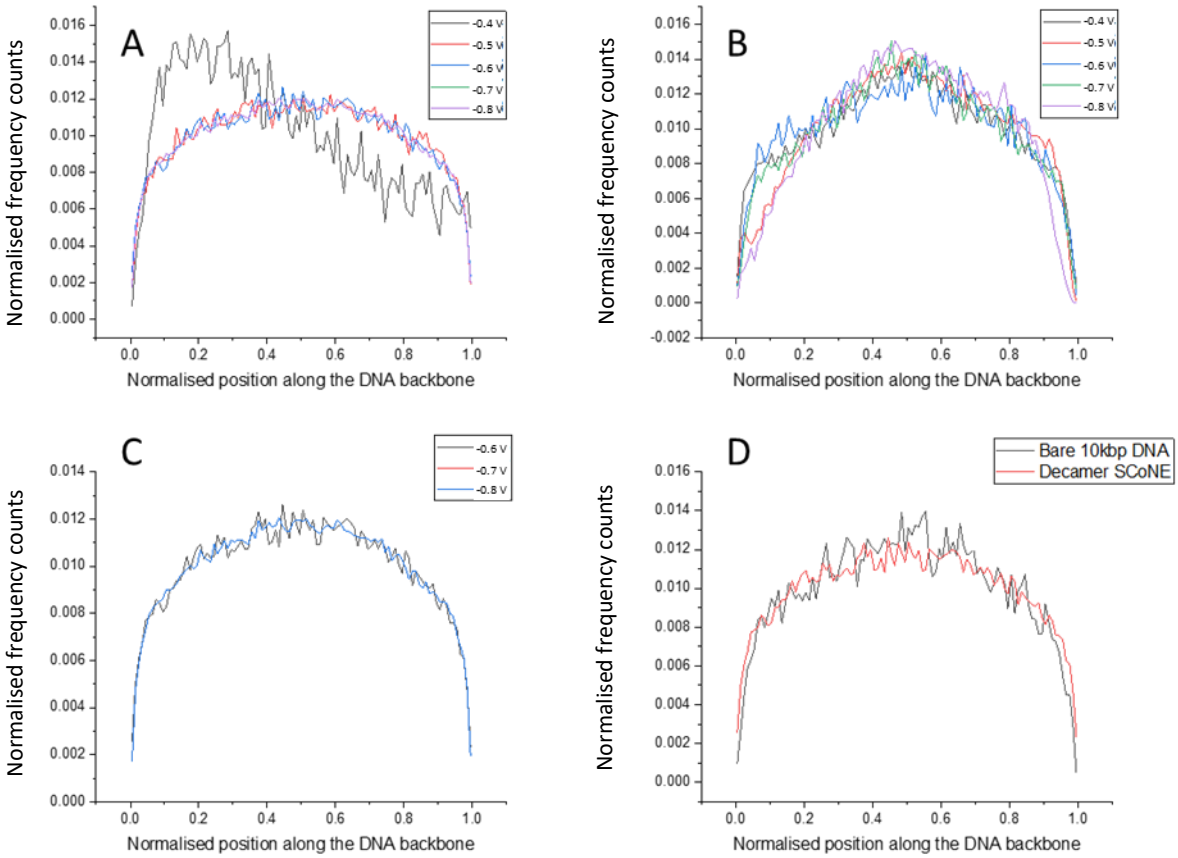

Figure 5.26. Comparison of subevents analysis conducted between 4mer SCoNE DNA and bare 4kbp DNA fragments. A 2D comparison between subevent position and bias using a 25pA threshold for a decamer SCoNE structure (A) shows that there are a number of subevent peaks observable. A 2D illustration of subevent position along the bare 10kbp DNA backbone across a range of biases shows similar positioning of the subevents (B). The 2D representation highlights some subevents present for the decamer SCoNE structure (C). Comparison between decamer SCoNE DNA and bare 10kbp DNA fragments at -0.6V bias shows similar profiles (D), indicating difficulty separating the groups.

When comparing singular events from decamer translocation experiments, it is possible to see expected substructure in the subevent analysis. Figure 5.27 highlights example events (excluding A, E, and I) observed in the initial scans (A, E, and I) across three of the biases used during experiments (-0.5, -0.6 and -0.7 V respectively). Subevents can be observed appearing from the baseline, with noise oscillations also present. The subevent appearance becomes pronounced when comparing events to bare 10 kbp DNA translocation events (Figure 5.28, excluding A, E, and I). Experiments were conducted using nanopores of similar size (15 and 19 nm respectively), with the respective DNA differences.

It was from the results of the experiments conducted using decamer SCoNE DNA that a new method for extraction and isolation was developed to prevent degradation of the probe structures. By using a technique which relied on creating and maintaining the integrity of the SCoNE DNA structures for isolation, it could be ensured that the final isolate would be the expected structure. The use of biotin as a tag, and its isolation using streptavidin beads was selected and incorporated into the overall SCoNE assembly method due to the low degradative nature of the technique.

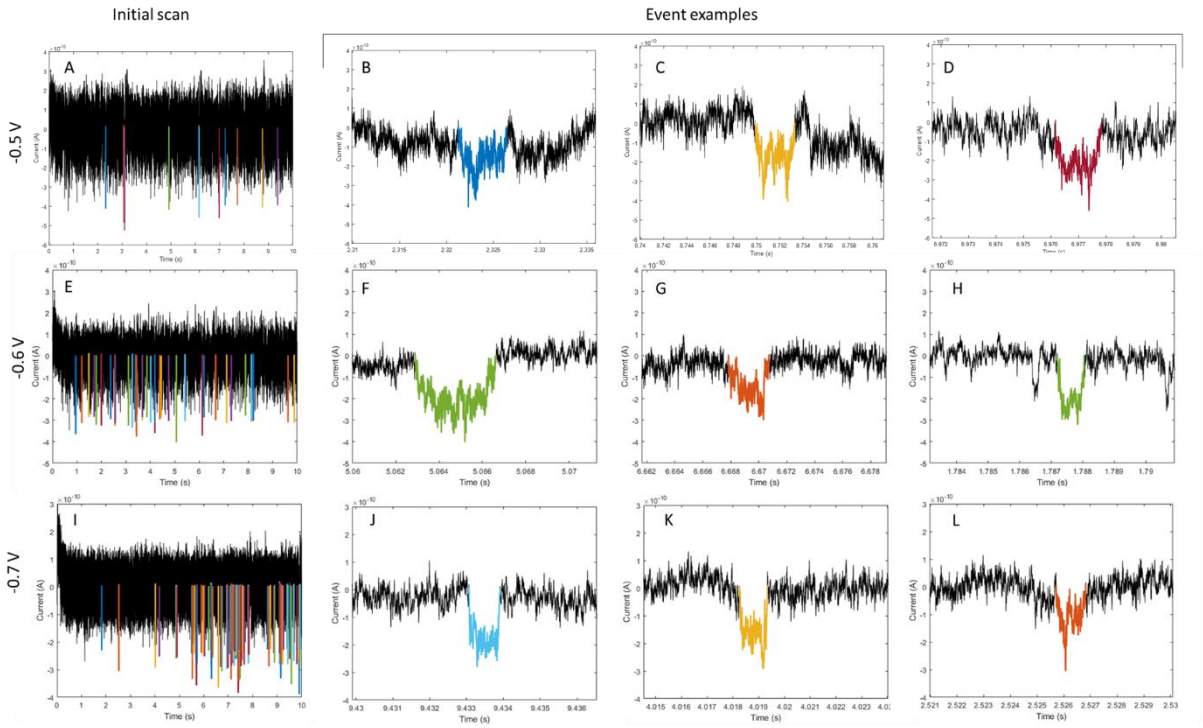

Figure 5.27. Examples of decamer SCoNE DNA current-time traces (A, E and I) for -0.5, -0.6, and -0.7 respectively, with single events highlighted for each bias (B-D, F-H, and J-L).

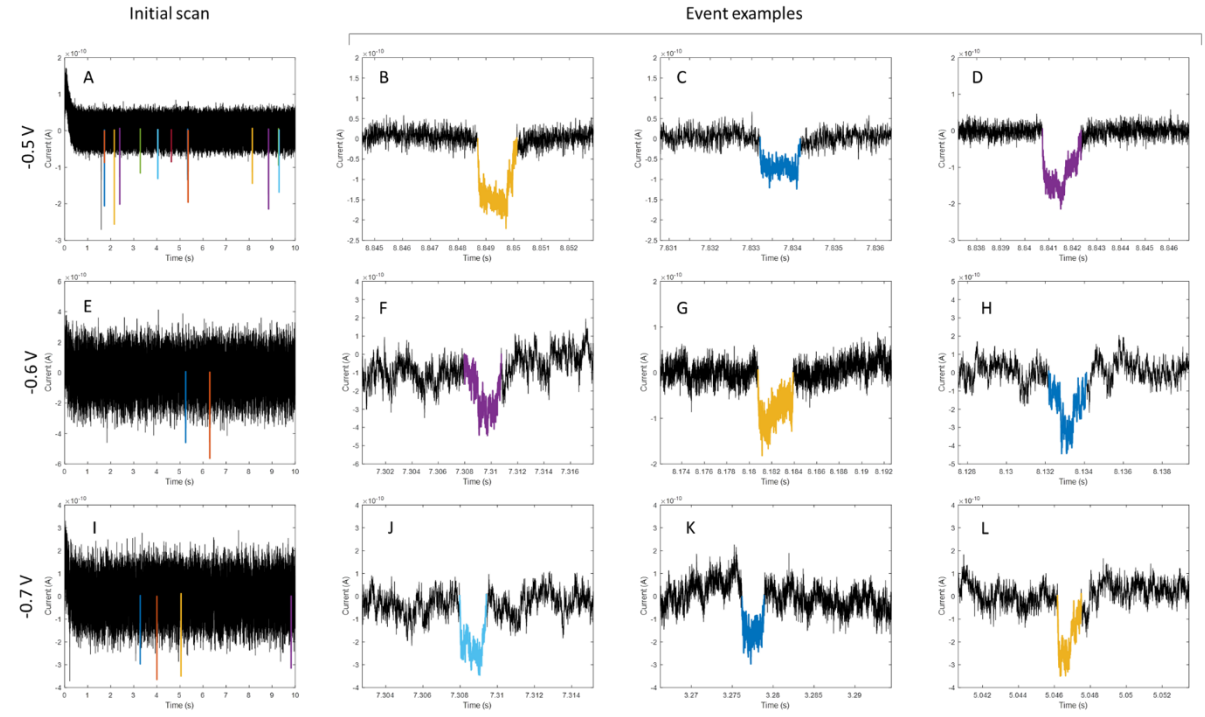

Figure 5.28. Examples of bare 10kbp DNA current-time traces (A, E and I) for -0.5, -0.6, and -0.7 respectively, with single events highlighted for each bias (B-D, F-H, and J-L).

5.4 Conclusion

From the results obtained, it can be concluded that it was possible to translocate, isolate and analyse DNA translocation through a nanopore under a variety of different conditions. It has also been possible to categorise noise events allowing for greater separation of noise and DNA events during further experiments. The results from the DNA translocation experiments highlights that the translocation frequency of ~1 per second per nM is similar to literature results. The expected profile, in terms of both intensity increasing and duration decreasing with an increase in applied bias, is similar to that obtained in the literature. From a comparison between 4 kbp and 10 kbp DNA translocation experiments, it has also been shown that there is a 3X increase in event duration when increasing the size of the DNA fragment, consistent with results from the research conducted by Fraccari *et al.* Through the analysis of single events, it has also been possible to highlight the potential ways in which DNA translocation can occur. Evidence for the formation of folded and knotted DNA has also been found, however while several of the events show these potential structures, it is not possible to absolutely confirm their presence. It is also possible to conclude that the development and adaptation of the code to include a variety of different separation techniques through analysis of the experimental data has improved the capability to detect and isolate DNA translocation events, whilst removing significant noise features.

The results of the translocation experiments highlight the ability of SCoNE DNA to be differentiated from unmodified DNA. Most notably in the tetramer results, the subevent peaks are clearly defined at positions near 0, 0.25, 0.5, and near 0.75 along the DNA backbone when applying a variety of thresholds to the data. Results from the decamer structure, whilst producing subevent peaks, was not significantly different from the

results obtained from bare 10 kbp DNA translocation. This could be due to the gel extraction method used to obtain the structure, causing degradation of the probe structures. The results from the translocation experiments also highlights that the tetramer structure behaves similarly to bare DNA. Additional repeats of these experiments with differently constructed structures would allow for further confirmation of the structures shape, ability, and limitations. The current limitation of the technique is the limited concentration of the product that can be produced, due to enzyme capabilities. Further development of an in-house enzyme mix could allow for a greater concentration to be constructed, allowing for more experiments to be conducted using the same samples.

5.5 References

Bell, N.A.W., and Keyser, U.F. (2016) Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. Nat Nanotechnol 11: 645–651.

Chen, K., Gularek, F., Liu, B., Weinhold, E., and Keyser, U.F. (2021) Electrical DNA Sequence Mapping Using Oligodeoxynucleotide Labels and Nanopores. ACS Nano 15: 2679–2685.

Edel, J., and Albrecht, T. (2012) Nanopores for Bioanalytical Applications Proceedings of the International Conference. 24-31.

Fraccari, R.L., Carminati, M., Piantanida, G., Leontidou, T., Ferrari, G., and Albrecht, T. (2016) High-bandwidth detection of short DNA in nanopipettes. Faraday Discuss 193: 459–470.

Haque, F., Wang, S., Stites, C., Chen, L., Wang, C., and Guo, P. (2015) Single pore translocation of folded, double-stranded, and tetra-stranded DNA through channel of bacteriophage phi29 DNA packaging motor. Biomaterials 53: 744–752.

Kumar Sharma, R., Agrawal, I., Dai, L., Doyle, P.S., and Garaj, S. (2019) Complex DNA knots detected with a nanopore sensor. Nat Commun 10: 4473.

Loh, A.Y.Y., Burgess, C.H., Tanase, D.A., Ferrari, G., McLachlan, M.A., Cass, A.E.G., and Albrecht, T. (2018) Electric Single-Molecule Hybridization Detector for Short DNA Fragments. Anal Chem 90: 14063–14071.

Shekar, S., Niedzwiecki, D.J., Chien, C.C., Ong, P., Fleischer, D.A., Lin, J., et al. (2016) Measurement of DNA translocation dynamics in a solid-state nanopore at 100 ns temporal resolution. Nano Lett 16: 4483–4489.

Steinbock, L.J., Otto, O., Chimerel, C., Gornall, J., and Keyser, U.F. (2010) Detecting DNA folding with nanocapillaries. Nano Lett 10: 2493–2497.

Chapter 6 – Final conclusion

Through the development of the SCoNE methodology, it was possible to overcome the limitations of Gibson assembly and incorporate 60 bp fragments into DNA assembly. The most interesting finding was the ability to replace the probe attached without affecting the overall protocol. SCoNE structures can be created and stored at -20°C without affecting its functionality, meaning that structures can be quickly assembled from stored reagents. The addition of biotin probes allowed for the development of an improved isolation technique and for post capture functionalisation, such as the binding of the ABC complex in ELISA experiments. The ELISA results also highlight that for trimer structures approximately 70% of the expected SCoNE structures in solution were measured with captured proteins. This decreases to 45% when measuring tetramer structures with IL-6 secondary antibodies and 28% for the procalcitonin secondary antibodies. Issues faced when using the procalcitonin secondary antibody could account for the decreased yield. It is also possible to suggest that with the addition of more DNA fragments the shape of the DNA could inhibit secondary antibody binding, leading to the decreased yield observed. It can be suggested that further optimisation of the distance between probes could improve the capture yield.

From optimisation of the pipette pulling parameters, it was possible to create a protocol for generating nanopipettes with pores sizing; greater than 30 nm, less than 30 nm, and less than 10 nm. The results suggest that alteration of the primary pull parameter significantly reduced the size of the nanopore generated, however this also increased the length of the taper. In contrast to this, decreasing the secondary pull increased the pore size, whilst decreasing taper length. When optimising the temperature parameter, it is noted that nanopores of less than 5 nm could be produced when decreasing

primary and increasing secondary pull heat. This had little effect on the taper length observed. When examining the primary pull force, reducing the pull power allowed for a pore size of less than 10 nm and decreased the overall taper length. Whilst changing laboratory conditions posed an issue for reproducibility, altering the pull force dependant on ambient temperature allowed for consistent pipette production. These parameters were then used as the standard for pipette production for use in future experiments.

The results from noise optimisation allowed for the establishment of a standard experimental procedure for all translocation experiments. It is possible to conclude that all the potential noise sources explored had a significant effect on the noise experienced during experiments. Whilst not possible to eliminate all of the noise from the generated signal, the overall effect of noise on data acquisition was reduced significantly. By anodizing electrodes, it was possible to see a reduction of 7 - 31%, dependant on the bias used. It was hypothesised that the smoother surface observed, in comparison to the chemically prepared electrodes, could reduce the production of nanobubbles and improve current flow. However, this requires further experimentation to confirm. When reducing the length of the electrodes, a significant decrease in noise was observed between 12 cm and 10 cm electrodes. This effect was not observed when decreasing the electrode length further, however the shorter electrodes reduced potential interaction with the surrounding Faraday cage and was therefore used for further experimentation. Analysing the potential for noise introduction during data transfer, a silver shielded cable was used to explore this. The shielded cable was capable of removing some of the low frequency noise peaks, but unable to remove the high frequency noise peaks. An overall decrease of 17.7 pA in RMS noise was also

observed when using the silver shielded cable. Due to this reduction and the removal of low frequency noise, it was decided that this cable would be incorporated into the experimental procedure. Notably the use of a granite anti-vibration table had a significant effect on reducing high and low frequency noise. A slight increase of 1.8 pA was observed in RMS noise when working on the granite surface, however due to the reduction of high and low frequency noise spikes, further experiments would incorporate this feature into the experimental procedure. Altering between active and passive filters showed an overall decrease in RMS noise when using passive filters below 240 kHz. At 240 kHz there is an increase in RMS noise when using the passive filters. Due to the decrease of RMS noise and signal oscillation, it was decided that passive filters would be incorporated into the experimental procedure. It was hypothesised that this effect could be due to the removal of mains power. This hypothesis was then tested by changing the power supply from mains to battery power. When using battery power, a decrease of 31.2 pA RMS noise and a significant decrease in signal oscillation over mains power was observed. It is possible to conclude from these results that using mains power is a major contributor to the noise observed during experiments. When comparing the combination of all noise reduction techniques it was possible to significantly reduce the effect of extrinsic noise and reduce overall RMS noise. The decrease in oscillation allowed for greater accuracy and reliability when identifying translocation events, improving the SNR, and decreasing false positive results during data extraction. Translocation experimental procedure would involve using shorter, anodised electrodes, a silver shielded data transfer cable, working on a granite slab, using a passive filter of 100 kHz or less, and using battery power where possible. Comparing the optimised experimental procedure

to expected current blockage induced by DNA translocation, using a 100 kHz passive filter, a SNR of ~5 would be expected.

From the results of the translocation experiment, it can be concluded that it was possible to translocate, isolate and analyse DNA translocation under several different conditions. It was also possible to categorise noise events and allow for greater separation between noise and DNA events. The DNA translocation frequency of ~1 per second per nM, event intensity and event duration are similar to previous literature results. Comparing results between 4 kbp and 10 kbp DNA translocation experiments, there is a 3X increase in event duration when increasing the size of the DNA fragment, consistent with results from the research conducted by Fraccari *et al.* Through the analysis of single events, it has also been possible to highlight evidence for the formation of folded and knotted DNA translocation events. While several of the events show these potential structures, it is not possible to absolutely confirm their presence. The development and adaptation of the code to include different separation techniques through analysis of the experimental data, has improved the capability to detect and isolate DNA translocation events, whilst removing significant noise features.

Analysis and comparison of the translocation experiments highlights the ability of SCoNE DNA to be differentiated from unmodified DNA. The results from the tetramer translocation analysis highlights the appearance of subevent peaks clearly defined at positions near 0, 0.25, 0.5, and near 0.75 along the DNA backbone. Results from the decamer structure, whilst producing subevent peaks, was not significantly different from the results obtained from bare 10 kbp DNA translocation. This could be due to the gel extraction method used to obtain the structure, causing degradation of the probe structures, or the inability of the nanopore to resolve the unbound probes.

Additional repeats of these experiments with differently constructed structures would allow for further confirmation of the structures shape, ability, and limitations. Current limitations of the technique are concentration of the product, due to enzyme capabilities. Further development of an in-house enzyme mix could allow for a greater concentration to be constructed, allowing for more experiments to be conducted using the same samples.

From contributions made from side projects explored during this research, as discussed in Appendix 1, it is possible to conclude that there are some promising results, however further optimisation would be required. Development of a simulated translocation code produced events with substructure as designed by the user. Whilst the overall results do not reflect a typical translocation experiment, the 10 second scans are visually representative of translocation scans and events. Further optimisation is required to create more realistic overall results. A proof-of-concept experiment for DNA isolation was designed and performed to be integrated with SCoNE DNA translocation. The primary goal of this research was to develop a method for extracting bound SCoNE DNA from patient samples. It was found that in 2.5 hours, 74.3 pM (24.77%) of the starting DNA concentration was recovered. Whilst this method is not fully optimised, it is possible to conclude that this technique could isolate DNA. The final side project was to image SCoNE DNA using AFM. Although the results are not conclusive, it is possible to conclude that DNA was captured and imaged using this technique. It is also possible to suggest that the results provide some complementary evidence that the DNA contained modifications to the backbone.

To conclude the overall arc of this research, the development of a new DNA assembly technique with functional biomolecule capture probes was successful, and the conjugated probes can be detected using nanopore sensing.

Chapter 7 – Plan for future work

There are several sections of this work which would merit further exploration and development. Firstly, the further optimisation of the SCoNE DNA assembly protocol to include additional capture probes should be explored. Work is currently being undertaken by the Albrecht group to explore antibody conjugation in place of aptamers. Preliminary research has shown some success in performing this conjugation using gel retardation assays. Work is currently being undertaken to explore the binding capacity of SCoNE structures using ELISAs. Due to time restrictions, it was not possible to fully explore the in-house enzyme mixture. Further exploration of this could potentially allow for greater concentrations of SCoNE DNA to be assembled in a single reaction vial. Experiments using complex mixtures of different proteins also merits further research, to ensure specificity and selectivity when using patient samples. Further translocation experiments should also be conducted to corroborate previous findings.

Further optimisation of nanopipettes to produce 10 nm pores with minimal taper length should also be considered to improve sensitivity of nanopore sensing. As is well known, noise is a constant issue faced with nanopore sensing. While it is not possible to entirely remove noise from a system, further optimisation for the removal of extrinsic noise should be explored.

Finally, the other projects explored during this research require further exploration and optimisation. Further development of the simulated translocation code to perform similarly to real translocation experiments could allow for improved subevent determination and allow for the inclusion of a neural network to identify bound or unbound probes on a SCoNE structure. Whilst the development of a DNA isolation

technique showed promising initial results, the low yield and incompatibility with nanopore sensing requires further exploration. Initial AFM imaging experiments showed the capability to capture DNA on a mica surface, however the coagulation of DNA constructs prevented conformation of the structures created. Further optimisation and imaging experiments using higher SCoNE structure concentration could allow for improved resolution and provide key evidence to confirm the exact shape of the SCoNE DNA.

Whilst all of these projects merit further research, it is the opinion of the author that the development of SCoNE assembly should be the primary focus. This should expand on including other probe molecules and increase the complexity of the structures generated. Current SCoNE structures should also be translocated using a commercially available nanopore system to determine its potential use in the wider nanopore community. Unfortunately, due to time restrictions, it was not possible to explore this final point.

Chapter 8 – Complete bibliography

Abe, R., Hirasawa, H., Oda, S., Sadahiro, T., Nakamura, M., Watanabe, E., et al. (2008) Up-Regulation of Interleukin-10 mRNA Expression in Peripheral Leukocytes Predicts Poor Outcome and Diminished Human Leukocyte Antigen-DR Expression on Monocytes in Septic Patients. J Surg Res 147: 1–8.

Albrecht, T. (2019) Single-Molecule Analysis with Solid-State Nanopores. Annu Rev Anal Chem 12: 371–387.

Albu, M., and Heydt, G.T. (2003) On the use of RMS values in power quality assessment. IEEE Trans Power Deliv 18: 1586–1587.

Alsaif, M., Guest, P.C., Schwarz, E., Reif, A., Kittel-Schneider, S., Spain, M., et al. (2012) Analysis of serum and plasma identifies differences in molecular coverage, measurement variability, and candidate biomarker selection. Proteomics - Clin Appl 6: 297–303.

Battaglia, F., Meucci, V., Tognetti, R., Bonelli, F., Sgorbini, M., Lubas, G., et al. (2020) Procalcitonin detection in veterinary species: Investigation of commercial elisa kits. Animals 10: 1–12.

Bell, N.A.W., and Keyser, U.F. (2015) Specific protein detection using designed DNA carriers and nanopores. J Am Chem Soc 137: 2035–2041.

Bell, N.A.W., and Keyser, U.F. (2016) Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. Nat Nanotechnol 11: 645–651.

Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002) Biochemistry. 9th ed., W H Freeman, 227-260.

Bestor, T.H. (2000) The DNA methyltransferases of mammals. Hum Mol Genet 9: 2395–2402.

Binnig, G. (1988) Atomic force microscope and method for imaging surfaces with atomic resolution. US Pat 4,724,318 1–8.

Boehme, T.K., and Bracewell, R. (1966) The Fourier Transform and its Applications. Am Math Mon 73: 685.

Brock, T.D., and Freeze, H. (1969) Thermus aquaticus gen. n. and sp. n., a nonsporulating extreme thermophile. J Bacteriol 98: 289–297.

Broza, Y.Y., Zhou, X., Yuan, M., Qu, D., Zheng, Y., Vishinkin, R., et al. (2019) Disease Detection with Molecular Biomarkers: From Chemistry of Body Fluids to Nature-Inspired Chemical Sensors. Chem Rev 119: 11761–11817.

BSGM (2021) List of Genetic Clinics - The British Society for Genetic Medicine. https://bsgm.org.uk/healthcare-professionals/list-of-genetic-clinics/. Accessed May 31, 2022.

Bulbul, G., Chaves, G., Olivier, J., Ozel, R., and Pourmand, N. (2018) Nanopipettes as Monitoring Probes for the Single Living Cell: State of the Art and Future Directions in Molecular Biology. Cells 7: 55.

Butler, T.Z., Pavlenok, M., Derrington, I.M., Niederweis, M., and Gundlach, J.H. (2008) Single-molecule DNA detection with an engineered MspA protein nanopore. Proc Natl Acad Sci U S A 105: 20647–20652.

Cao, Y., Yu, Y.J., Jiang, Q.Y., Sun, Y., Mao, Z., Wang, J., and Chen, F. (2021) The Advance of Plasmonic-Electric Nanopipette Sensing in Single Cells. Curr Pharmacol Reports 7: 55–66.

Centers for Disease Control and Prevention (2018) What is sepsis? | Sepsis | CDC. Centers Dis Control Prev.

Chalmeau, J., Monina, N., Shin, J., Vieu, C., and Noireaux, V. (2011) α-Hemolysin pore formation into a supported phospholipid bilayer using cell-free expression. Biochim Biophys Acta - Biomembr 1808: 271–278.

Chehrehasa, F., Meedeniya, A.C.B., Dwyer, P., Abrahamsen, G., and Mackay-Sim, A. (2009) EdU, a new thymidine analogue for labelling proliferating cells in the nervous system. J Neurosci Methods 177: 122–130.

Chen, K., Gularek, F., Liu, B., Weinhold, E., and Keyser, U.F. (2021) Electrical DNA Sequence Mapping Using Oligodeoxynucleotide Labels and Nanopores. ACS Nano 15: 2679–2685.

Chen, P., Gu, J., Brandin, E., Kim, Y.R., Wang, Q., and Branton, D. (2004) Probing single DNA molecule transport using fabricated nanopores. Nano Lett 4: 2293–2298.

Cochrane, D., Chen, D.Y., and Boroyevic, D. (2003) Passive cancellation of common-mode noise in power electronic circuits. IEEE Trans Power Electron 18: 756–763.

Couldrey, C., Brauning, R., Henderson, H. V., and McEwan, J.C. (2015) Genome-wide DNA methylation analysis: No evidence for stable hemimethylation in the sheep muscle genome. Anim Genet 46: 185–189.

Coulter, W.H. (1953) Means for counting particles suspended in a fluid. US Patent, 2656508. United States Pat Off Patentiert am 20: 1953.

Daintith, J. (2009) A dictionary of physics. 6th ed., Oxford University Press.

Dangi, A.K., Sinha, R., Dwivedi, S., Gupta, S.K., and Shukla, P. (2018) Cell line techniques and gene editing tools for antibody production: A review. Front Pharmacol 9: 630.

DATAFORTH (2002) R. 221–222 https://www.dataforth.com/protecting-signal-lines-against-electromagnetic-interference.aspx. Accessed October 24, 2021.

Davis, N. (2017) An Introduction to Filters - Technical Articles. allaboutcircuits.com https://www.allaboutcircuits.com/technical-articles/an-introduction-to-filters/. Accessed October 24, 2021.

Deamer, D., Akeson, M., and Branton, D. (2016) Three decades of nanopore sequencing. Nat Biotechnol 34: 518–524.

DeBlois, R.W., and Bean, C.P. (1970) Counting and sizing of submicron particles by the resistive pulse technique. Rev Sci Instrum 41: 909–916.

Dellinger, R.P., Levy, M.M., Rhodes, A., Annane, D., Gerlach, H., Opal, S.M., et al. (2013) Surviving Sepsis Campaign. Crit Care Med 41: 580–637.

Derrington, I.M., Butler, T.Z., Collins, M.D., Manrao, E., Pavlenok, M., Niederweis, M., and Gundlach, J.H. (2010) Nanopore DNA sequencing with MspA. Proc Natl Acad Sci U S A 107: 16060–16065.

Deutschman, C.S., and Tracey, K.J. (2014) Sepsis: Current dogma and new perspectives. Immunity 40: 463–475.

Devaraj, N.K., and Finn, M.G. (2021) Introduction: Click Chemistry. Chem Rev 121: 6697–6698.

Ding, S., Gao, C., and Gu, L.-Q. (2009) Capturing Single Molecules of Immunoglobulin and Ricin with an Aptamer-Encoded Glass Nanopore. Anal Chem 81: 6649–6655.

Dorsett, M., Westlund, B., and Schedl, T. (2009) METT-10, a putative methyltransferase, inhibits germ cell proliferative fate in Caenorhabditis elegans. Genetics 183: 233–247.

Downie, J.D., Tomkos, I., Antoniades, N., and Boskovic, A. (2002) Effects of filter concatenation for directly modulated transmission lasers at 2.5 and 10 Gb/s. J Light Technol 20: 218–228.

Dunne, P.W., Roberts, D.L., Quinones, M.P., Velligan, D.I., Paredes, M., and Walss-Bass, C. (2017) Immune markers of social cognitive bias in schizophrenia. Psychiatry Res 251: 319–324.

Dupont, C., Armant, D.R., and Brenner, C.A. (2009) Epigenetics: Definition, mechanisms and clinical perspective. Semin Reprod Med 27: 351–357.

Dutta, P., and Horn, P.M. (1981) Low-frequency fluctuations in solids: 1f noise. Rev Mod Phys 53: 497–516.

Edel, J., and Albrecht, T. (2012) Nanopores for Bioanalytical Applications Proceedings of the International Conference. 24–31.

Edel, J., and Albrecht, T. (2013) Engineered Nanopores for Bioanalytical Applications: A Volume in Micro and Nano Technologies. Elsevier Inc., 1–172.

Erlich, H.A. (2015) *PCR technology: Principles and applications for DNA amplification*. 1st ed., Palgrave Macmillan, 7–22.

Faqi, A.S. (2017) *A Comprehensive Guide to Toxicology in Nonclinical Drug Development*. Elsevier, 326–496.

Farrelly, L.A., Thompson, R.E., Zhao, S., Lepack, A.E., Lyu, Y., Bhanu, N. V., et al. (2019) Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3. Nature 567: 535–539.

FDA-NIH Biomarker Working Group (2016) BEST (Biomarkers, EndpointS, and other Tools) Resource. Updat Sept 25 55 https://www.ncbi.nlm.nih.gov/books/NBK326791/. Accessed July 26, 2021.

Feng, W.N., He, J., Li, S.Y., Yang, D., Li, H.N., Yang, G.C., et al. (2020) Detection of Secretion of Exosomes from Individual Cell in Real-Time by Multifunctional Nanoelectrode-Nanopore Nanopipettes. Chinese J Anal Chem 48: e20061–e20068.

Firnkes, M., Pedone, D., Knezevic, J., Döblinger, M., and Rant, U. (2010) Electrically facilitated translocations of proteins through silicon nitride nanopores: Conjoint and competitive action of diffusion, electrophoresis, and electroosmosis. Nano Lett 10: 2162–2167.

Fish, P.J. (2017) Electronic Noise and Low Noise Design. Macmillan International Higher Education, 1–30.

Flanagan, R.J., and Jones, A.L. (2004) Fab antibody fragments: Some applications in clinical toxicology. Drug Saf 27: 1115–1133.

Fologea, D., Gershow, M., Ledden, B., McNabb, D.S., Golovchenko, J.A., and Li, J. (2005) Detecting single stranded DNA with a solid state nanopore. Nano Lett 5: 1905–1909.

Fraccari, R.L., Carminati, M., Piantanida, G., Leontidou, T., Ferrari, G., and Albrecht, T. (2016) High-bandwidth detection of short DNA in nanopipettes. Faraday Discuss 193: 459–470.

Fragasso, A., Schmid, S., and Dekker, C. (2020) Comparing Current Noise in Biological and Solid-State Nanopores. ACS Nano 14: 1338–1349.

Frament, C.M., and Dwyer, J.R. (2012) Conductance-based determination of solid-state nanopore size and shape: An exploration of performance limits. J Phys Chem C 116: 23315–23321.

Fu, Y., Tokuhisa, H., and Baker, L.A. (2009) Nanopore DNA sensors based on dendrimer-modified nanopipettes. Chem Commun (Camb) 4877–4879.

Gamry (2021) Faraday Cage: What is a Faraday Cage-How Does a Faraday Cage Work? Gamry Instruments https://www.gamry.com/application-notes/instrumentation/faraday-cage/. Accessed October 24, 2021.

Gao, F., Gao, Y., Zhang, S.J., Zhe, X., Meng, F.L., Qian, H., et al. (2017) Alteration of plasma cytokines in patients with active epilepsy. Acta Neurol Scand 135: 663–669.

Gardner, A., Carpenter, G., and So, P.-W. (2020) Salivary Metabolomics: From Diagnostic Biomarker Discovery to Investigating Biological Function. Metab 2020, Vol 10, 10–47.

Gazquez, C., Ribal, M.J., Marín-Aguilera, M., Kayed, H., Fernández, P.L., Mengual, L., and Alcaraz, A. (2012) Biomarkers vs conventional histological analysis to detect lymph node micrometastases in bladder cancer: A real improvement? BJU Int 110: 1310–1316.

Gibson, D.G., and Russello, S. (2010) How Gibson Assembly® is Changing Synthetic Biology | NEB. New Engl BioLabs https://www.neb.com/tools-and-resources/feature-articles/gibson-assembly-building-a-synthetic-biology-toolset. Accessed February 17, 2022.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6: 343–345.

Gilboa, T., Garden, P.M., and Cohen, L. (2020) Single-molecule analysis of nucleic acid biomarkers – A review. Anal Chim Acta 1115: 61–85.

Golla, J.P., Zhao, J., Mann, I.K., Sayeed, S.K., Mandal, A., Rose, R.B., and Vinson, C. (2014) Carboxylation of cytosine (5caC) in the CG dinucleotide in the E-box motif (CGCAG|GTG) increases binding of the Tcf3|Ascl1 helix-loop-helix heterodimer 10-fold. Biochem Biophys Res Commun 449: 248–255.

Gomez-Marquez, J., and Hamad-Schifferli, K. (2021) Local development of nanotechnology-based diagnostics. Nat Nanotechnol 16: 484–486.

Gong, H., Holcomb, I., Ooi, A., Wang, X., Majonis, D., Unger, M.A., and Ramakrishnan, R. (2016) Simple Method to Prepare Oligonucleotide-Conjugated Antibodies and Its Application in Multiplex Protein Detection in Single Cells. Bioconjug Chem 27: 217–225.

Gu, R., Lamas, J., Rastogi, S.K., Li, X., Brittain, W., and Zauscher, S. (2015) Photocontrolled micellar aggregation of amphiphilic DNA-azobenzene conjugates. Colloids Surfaces B Biointerfaces 135: 126–132.

Gubbiotti, A., Baldelli, M., Muccio, G. Di, Malgaretti, P., Marbach, S., and Chinappi, M. (2022) Electroosmosis in nanopores: computational methods and technological applications. Adv Phys X 7: 2036638

Hack, C.E., Hart, M., Schijndel, R.J.M.S. Van, Eerenberg, A.J.M., Nuijens, J.H., Thijs, L.G., and Aarden, L.A. (1992) Interleukin-8 in sepsis: Relation to shock and inflammatory mediators. Infect Immun 60: 2835–2842.

Hanly, W.C., Artwohl, J.E., and Bennett, B.T. (1995) Review of Polyclonal Antibody Production Procedures in Mammals and Poultry. ILAR J 37: 93–118.

Hansson, O., Finnebraaten, M., Heitmann, I.K., Ramse, M., and Bouzga, M. (2009) Trace DNA collection-Performance of minitape and three different swabs. Forensic Sci Int Genet Suppl Ser 2: 189–190.

Haque, F., Wang, S., Stites, C., Chen, L., Wang, C., and Guo, P. (2015) Single pore translocation of folded, double-stranded, and tetra-stranded DNA through channel of bacteriophage phi29 DNA packaging motor. Biomaterials 53: 744–752.

Harmel, R., and Fiedler, D. (2018) Features and regulation of non-enzymatic post-translational modifications. Nat Chem Biol 14: 244–252.

Hartel, A.J.W., Shekar, S., Ong, P., Schroeder, I., Thiel, G., and Shepard, K.L. (2019) High bandwidth approaches in nanopore and ion channel recordings - A tutorial review. Anal Chim Acta 1061: 13–27.

Hartlieb, P., and Grafe, B. (2017) Experimental Study on Microwave Assisted Hard Rock Cutting of Granite. BHM Berg- und Hüttenmännische Monatshefte 162: 77–81.

Haschek, W.M., Rousseaux, C.G., Wallig, M.A., Bolon, B., and Ochoa, R. (2013) Haschek and Rousseaux's Handbook of Toxicologic Pathology. 1st ed., Elsevier Inc., 9–20.

Hirsch, M.S., and Watkins, J. (2020) A comprehensive review of biomarker use in the gynecologic tract including differential diagnoses and diagnostic pitfalls. Adv Anat Pathol 27: 164–192.

Hoogerheide, D.P., Garaj, S., and Golovchenko, J.A. (2009) Probing surface charge fluctuations with solid-state nanopores. Phys Rev Lett 102: 256804.

Hottenstein, C., Szapacs, M., Fuller, K., and Evans, C. (2017) Platforms and techniques used for biomarker assays: where are we now? https://doi.org/104155/bio-2017-0107 9: 1029–1031.

Hsu, W.Y., Chen, C.J., Huang, Y.C., Tsai, F.J., Jeng, L. Bin, and Lai, C.C. (2013) Urinary nucleosides as biomarkers of breast, colon, lung, and gastric cancer in Taiwanese e81701. PLoS One 8: e81701.

Huo, M.Z., Li, M.Y., Ying, Y.L., and Long, Y.T. (2021) Is the volume exclusion model practicable for nanopore protein sequencing? Anal Chem 93: 11364–11369.

Iqbal, S.M., Akin, D., and Bashir, R. (2007) Solid-state nanopore channels with DNA selectivity. Nat Nanotechnol 2: 243–248.

Ivanov, Y.D., Bukharina, N.S., Pleshakova, T.O., Frantsuzov, P.A., Andreeva, E.Y., Kaysheva, A.L., et al. (2014) Atomic force microscopy fishing and mass spectrometry identification of gp120 on immobilized aptamers. Int J Nanomedicine 9: 4659–4670.

Janeway, C., Travers, P., Walport, M., and Capra, J. (2005) *Immunobiology: the immune system in health and disease*. 5th ed., New York: Garland Science, 1–62.

Jeon, S., Thundat, T., and Braiman, Y. (2006) Effect of normal vibration on friction in the atomic force microscopy experiment. Appl Phys Lett 88: 214102.

Jeong, J.A. (2014) ELISA technical guide and protocols_ thermo. Appl Mech Mater 597: 421–424.

Jia, H., Guo, Y., Zhao, W., and Wang, K. (2014) Long-range PCR in next-generation sequencing: Comparison of six enzymes and evaluation on the MiSeq sequencer. Sci Rep 4: 1–8.

Kasianowicz, J.J., Brandin, E., Branton, D., and Deamer, D.W. (1996) Characterization of individual polynucleotide molecules using a membrane channel. In Proceedings of the National Academy of Sciences of the United States of America. Proc Natl Acad Sci U S A, 13770–13773.

Kasteren, P.B. van, Veer, B. van der, Brink, S. van den, Wijsman, L., Jonge, J. de, Brandt, A. van den, et al. (2020) Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. J Clin Virol 128: 104412.

Kemmler, M., Sauer, U., Schleicher, E., Preininger, C., and Brandenburg, A. (2014) Biochip point-of-care device for sepsis diagnostics. Sensors Actuators, B Chem 192: 205–215.

Khan, T.K. (2016) *Clinical Diagnosis of Alzheimer's Disease BT - Biomarkers in Alzheimer's Disease*. 1st ed., Academic Press, 27–48.

Kharas, M.G., Janes, M.R., Scarfone, V.M., Lilly, M.B., Knight, Z.A., Shokat, K.M., and Fruman, D.A. (2008) Ablation of PI3K blocks BCR-ABL leukemogenesis in mice, and a dual PI3K/mTOR inhibitor prevents expansion of human BCR-ABL+ leukemia cells. J Clin Invest 118: 3038–3050.

Khrais, N.N., Elrefaie, A.F., Wagner, R.E., and Ahmed, S. (1996) Performance of cascaded misaligned optical (De)multiplexers in multiwavelength optical networks. IEEE Photonics Technol Lett 8: 1073–1075.

Kilty, C.G., Keenan, J., and Shaw, M. (2007) Histologically defined biomarkers in toxicology. Expert Opin Drug Saf 6: 207–215.

Kim, S.C., Kannam, S.K., Harrer, S., Downton, M.T., Moore, S., and Wagner, J.M. (2014) Geometric dependence of the conductance drop in a nanopore due to a particle. Phys Rev E - Stat Nonlinear, Soft Matter Phys 89: 42702.

Klungland, A., and Robertson, A.B. (2017) Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine. Free Radic Biol Med 107: 62–68.

Kolb, H.C., Finn, M.G., and Sharpless, K.B. (2001) Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew Chemie - Int Ed 40: 2004–2021.

Konstantinova, I.M., Tsimokha, A.S., and Mittenberg, A.G. (2008) Role of Proteasomes in Cellular Regulation. Int Rev Cell Mol Biol 267: 59–124.

Koopmann, J., Fedarko, N.S., Jain, A., Maitra, A., Iacobuzio-Donahue, C., Rahman, A., et al. (2004) Evaluation of osteopontin as biomarker for pancreatic adenocarcinoma. Cancer Epidemiol Biomarkers Prev 13: 487–91.

Korman, N.J. (2020) Management of psoriasis as a systemic disease: what is the evidence? Br J Dermatol 182: 840–848.

Kotiw, M., Johnson, M., Pandey, M., Fry, S., Hazell, S.L., Netter, H.J., et al. (2012) Immunological response to parenteral vaccination with recombinant hepatitis B virus surface antigen virus-like particles expressing Helicobacter pylori KatA epitopes in a murine H. pylori challenge model. Clin Vaccine Immunol 19: 268–276.

Kowalczyk, S.W., Grosberg, A.Y., Rabin, Y., and Dekker, C. (2011) Modelling the conductance and DNA blockade of solid-state nanopores. Nanotechnology 22: 315101.

Kowalczyk, S.W., Wells, D.B., Aksimentiev, A., and Dekker, C. (2012) Slowing down DNA translocation through a nanopore in lithium chloride. Nano Lett 12: 1038–1044.

Kreplak, L. (2016) Introduction to atomic force microscopy (AFM) in biology. Curr Protoc Protein Sci 2016: 17.7.1-17.7.21.

Kruspe, S., and Giangrande, P.H. (2017) Aptamer-siRNA chimeras: Discovery, progress, and future prospects. Biomedicines 5: 5–45.

Kuhn, M., Sühs, K.W., Akmatov, M.K., Klawonn, F., Wang, J., Skripuletz, T., et al. (2018) Mass-spectrometric profiling of cerebrospinal fluid reveals metabolite biomarkers for CNS involvement in varicella zoster virus reactivation. J Neuroinflammation 15: 1–15.

Kumar Sharma, R., Agrawal, I., Dai, L., Doyle, P.S., and Garaj, S. (2019) Complex DNA knots detected with a nanopore sensor. Nat Commun 10: 4473.

Kumar, S., Chinnusamy, V., and Mohapatra, T. (2018) Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. Front Genet 9: 640.

Kumar, Y., Qunibi, M., Neal, T.J., and Yoxall, C.W. (2001) Time to positivity of neonatal blood cultures. Arch Dis Child - Fetal Neonatal Ed 85: 182F – 186.

Laterza, O.F., Hendrickson, R.C., and Wagner, J.A. (2007) Molecular biomarkers. Drug Inf J 41: 573–585.

Levine, R.L. (2002) Carbonyl modified proteins in cellular regulation, aging, and disease. Free Radic Biol Med 32: 790–796.

Levis, R.A., and Rae, J.L. (1993) The use of quartz patch pipettes for low noise single channel recording. Biophys J 65: 1666–1677.

Li, B.Q., and Einstein, H.H. (2017) Comparison of Visual and Acoustic Emission Observations in a Four Point Bending Experiment on Barre Granite. Rock Mech Rock Eng 50: 2277–2296.

Li, J., Gershow, M., Stein, D., Brandin, E., and Golovchenko, J.A. (2003) DNA molecules and configurations in a solid-state nanopore microscope. Nat Mater 2: 611–615.

Li, J., Stein, D., McMullan, C., Branton, D., Aziz, M.J., and Golovchenko, J.A. (2001) Ion-beam sculpting at nanometre length scales. Nature 412: 166–169.

Li, X., Xu, Z., Pang, X., Huang, Y., Yang, B., Yang, Y., et al. (2017) Interleukin-10/lymphocyte ratio predicts mortality in severe septic patients. PLoS One 12: e0179050.

Liang, S., Xiang, F., Tang, Z., Nouri, R., He, X., Dong, M., and Guan, W. (2020) Noise in nanopore sensors: Sources, models, reduction, and benchmarking. Nami Jishu yu Jingmi Gongcheng/Nanotechnology Precis Eng 3: 9–17.

Liebes, Y., Drozdov, M., Avital, Y.Y., Kauffmann, Y., Rapaport, H., Kaplan, W.D., and Ashkenasy, N. (2010) Reconstructing solid state nanopore shape from electrical measurements. Appl Phys Lett 97: 223105.

Lin, G.L., McGinley, J.P., Drysdale, S.B., and Pollard, A.J. (2018) Epidemiology and Immune Pathogenesis of Viral Sepsis. Front Immunol 9: 2147.

Liszewski, M.K., Farries, T.C., Lublin, D.M., Rooney, I.A., and Atkinson, J.P. (1996) Control of the complement system. Adv Immunol 61: 201–283.

Litman, G.W., Rast, J.P., Shamblott, M.J., Haire, R.N., Hulst, M., Roess, W., et al. (1993) Phylogenetic diversification of immunoglobulin genes and the antibody repertoire. Mol Biol Evol 10: 60–72.

Liyanage, V.R.B., Jarmasz, J.S., Murugeshan, N., Bigio, M.R.D., Rastegar, M., and Davie, J.R. (2014) DNA Modifications: Function and Applications in Normal and Disease States. Biology (Basel) 3: 670–723.

Loh, A. (2017) A DNA Assay On a String – Fabrication and Sensing - Thesis.

Loh, A.Y.Y., Burgess, C.H., Tanase, D.A., Ferrari, G., McLachlan, M.A., Cass, A.E.G., and Albrecht, T. (2018) Electric Single-Molecule Hybridization Detector for Short DNA Fragments. Anal Chem 90: 14063–14071.

Maglia, G., Restrepo, M.R., Mikhailova, E., and Bayley, H. (2008) Enhanced translocation of single DNA molecules through α-hemolysin nanopores by manipulation of internal charge. Proc Natl Acad Sci 105: 19720–19725.

Mallikaratchy, P. (2017) Evolution of complex target SELEX to identify aptamers against mammalian cell-surface antigens. Molecules 22: 22–215.

Mateus-Pinheiro, A., Pinto, L., and Sousa, N. (2011) Epigenetic (de)regulation of adult hippocampal neurogenesis: implications for depression. Clin Epigenetics 3: 5.

Matsuda, F. (2016) Technical Challenges in Mass Spectrometry-Based Metabolomics. Mass Spectrom 5: S0052.

McAllister, A.B., Menchaca, C.X., Freeman, J., Chen, R., and Shen, M. (2018) A Nanopipette Platform for Delivering Nanoliter Volumes of Acetylcholine. J Electrochem Soc 165: G3093–G3098.

Medicine, I. of (2004) Physiological Biomarkers for Predicting Performance. In Monitoring Metabolic Status: Predicting Decrements in Physiological and Cognitive Performance. National Academies Press (US), 103.

Medzhitov, R., and Janeway, C.A. (1997) Innate immunity: impact on the adaptive immune response. Curr Opin Immunol 9: 4–9.

Moore, K.W., Waal Malefyt, R. De, Coffman, R.L., and O'Garra, A. (2001) Interleukin-10 and the interleukin-10 receptor. Annu Rev Immunol 19: 683–765.

Moore, L.D., Le, T., and Fan, G. (2013) DNA methylation and its basic function. Neuropsychopharmacology 38: 23–38.

Morris, C.A., Friedman, A.K., and Baker, L.A. (2010) Applications of nanopipettes in the analytical sciences. Analyst 135: 2190–2202.

Motchenbacher, C., and Connelly, J. (1994) Low-noise electronic system design. Choice Rev Online 31: 31-2710-31–2710.

Muthukumar, M. (2007) Mechanism of DNA transport through pores. Annu Rev Biophys Biomol Struct 36: 435–450.

Nelson, M., and McClelland, M. (1987) The effect of site-specific methylation on restriction-modification enzymes. Nucleic Acids Res 15 Suppl: r219.

NHS (2020) Blood tests - NHS. NHS https://www.nhs.uk/conditions/blood-tests/. Accessed July 26, 2021.

NHS (2021a) Interleukin - 6 (IL-6) - North West London Pathology. https://www.nwlpathology.nhs.uk/tests-database/interleukin-6-il-6/. Accessed June 1, 2022.

NHS (2021b) C-reactive protein (CRP) - North West London Pathology. https://www.nwlpathology.nhs.uk/tests-database/c-reactive-protein-crp/. Accessed June 1, 2022.

NHS (2021c) Procalcitonin - North West London Pathology. https://www.nwlpathology.nhs.uk/tests-database/procalcitonin/. Accessed June 1, 2022.

Nordström, A., and Lewensohn, R. (2009) Metabolomics: Moving to the Clinic. J Neuroimmune Pharmacol 2009 51 5: 4–17.

Nowak, S.J., and Corces, V.G. (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet 20: 214–220.

Olson, H.F. (1956) Electronic Control of Noise, Vibration, and Reverberation. J Acoust Soc Am 28: 966–972.

Olszewski, N.E., West, C.M., Sassi, S.O., and Hartweck, L.M. (2010) O-GlcNAc protein modification in plants: Evolution and function. Biochim Biophys Acta - Gen Subj 1800: 49–56.

Otani, T., Antoniades, N., Roudas, I., and Stern, T.E. (1999) Cascadability of passband-flattened arrayed waveguide-grating filters in WDM optical networks. IEEE Photonics Technol Lett 11: 1414–1416.

Ott, H.W. (1988) Noise Reduction Techniques in Electronic Systems. 2nd ed., John Wiley & Sons, 1–348.

Parkin, W.M., and Drndić, M. (2018) Signal and Noise in FET-Nanopore Devices. ACS Sensors 3: 313–319.

Parlato, M., and Cavaillon, J.M. (2014) Host response biomarkers in the diagnosis of sepsis: A general overview. Methods Mol Biol 1237: 149–211.

Peeling, R.W., Murtagh, M., and Olliaro, P.L. (2019) Epidemic preparedness: why is there a need to accelerate the development of diagnostics? Lancet Infect Dis 19: e172–e178.

Pellitero, M.A., Shaver, A., and Arroyo-Currás, N. (2020) Critical Review—Approaches for the Electrochemical Interrogation of DNA-Based Sensors: A Critical Review. J Electrochem Soc 167: 037529.

Plesa, C., Ruitenberg, J.W., Witteveen, M.J., and Dekker, C. (2015) Detection of individual proteins bound along DNA using solid-state nanopores. Nano Lett 15: 3153–3158.

Pujol, P., Barberis, M., Beer, P., Friedman, E., Piulats, J.M., Capoluongo, E.D., et al. (2021) Clinical practice guidelines for BRCA1 and BRCA2 genetic testing. Eur J Cancer 146: 30–47.

Rae, J.L., and Levis, R.A. (1992) A method for exceptionally low noise single channel recordings. Pflügers Arch Eur J Physiol 420: 618–620.

Ramazzotti, D. (2014) An Observational Study: The Effect of Diuretics Administration on Outcomes of Mortality and Mean Duration of I.C.U. Stay, Thesis, MIT, 1-75.

Raymond, J.J., Oorschot, R.A.H. van, Gunn, P.R., Walsh, S.J., and Roux, C. (2009) Trace evidence characteristics of DNA: A preliminary investigation of the persistence of DNA at crime scenes. Forensic Sci Int Genet 4: 26–33.

Reth, M. (2013) Matching cellular dimensions with molecular sizes. Nat Immunol 14: 765–767.

Rodgers, P.J., Amemiya, S., Wang, Y., and Mirkin, M. V. (2010) Nanopipette Voltammetry of Common Ions across the Liquid−Liquid Interface. Theory and Limitations in Kinetic Analysis of Nanoelectrode Voltammograms. Anal Chem 82: 84–90.

Rothery, A.M., Gorelik, J., Bruckbauer, A., Yu, W., Korchev, Y.E., and Klenerman, D. (2003) A novel light source for SICM-SNOM of living cells. J Microsc 209: 94–101.

Russell, C., Ward, A.C., Vezza, V., Hoskisson, P., Alcorn, D., Steenson, D.P., and Corrigan, D.K. (2019) Development of a needle shaped microelectrode for electrochemical detection of the sepsis biomarker interleukin-6 (IL-6) in real time. Biosens Bioelectron 126: 806–814.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science (80-) 230: 1350–1354.

Schrlau, M.G., Dun, N.J., and Bau, H.H. (2009) Cell Electrophysiology with Carbon Nanopipettes. ACS Nano 3: 563–568.

Schvartzman, J.M., Thompson, C.B., and Finley, L.W.S. (2018) Metabolic regulation of chromatin modifications and gene expression. J Cell Biol 217: 2247–2259.

Selleck, M.J., Senthil, M., and Wall, N.R. (2017) Making Meaningful Clinical Use of Biomarkers. Biomark Insights 12: 1–7.

Shekar, S., Niedzwiecki, D.J., Chien, C.C., Ong, P., Fleischer, D.A., Lin, J., et al. (2016) Measurement of DNA translocation dynamics in a solid-state nanopore at 100 ns temporal resolution. Nano Lett 16: 4483–4489.

Shenoi, B.A. (2005) *Introduction to Digital Signal Processing and Filter Design*. 1st ed., Wiley-Interscience, 1–423.

Shevchuk, A.I., Gorelik, J., Harding, S.E., Lab, M.J., Klenerman, D., and Korchev, Y.E. (2001) Simultaneous measurement of Ca2+ and cellular dynamics: combined scanning ion conductance and optical microscopy to study contracting cardiac myocytes. Biophys J 81: 1759–64.

Shi, W., Friedman, A.K., and Baker, L.A. (2017) Nanopore Sensing. Anal Chem 89: 157–188.

Si, W., and Aksimentiev, A. (2017) Nanopore Sensing of Protein Folding. ACS Nano 11: 7091–7100.

Sigma-Aldrich (2018) DNA Oligonucleotide Synthesis. https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/genomics/pcr/dna-oligonucleotide-synthesis. Accessed July 26, 2021.

Sinclair, K., and Dudley, E. (2019) Metabolomics and Biomarker Discovery. Adv Exp Med Biol 1140: 613–633.

Slichter, D.H., Naaman, O., and Siddiqi, I. (2009) Millikelvin thermal and electrical performance of lossy transmission line filters. Appl Phys Lett 94: 192508.

Smeets, R.M.M., Keyser, U.F., Dekker, N.H., and Dekker, C. (2008) Noise in solid-state nanopores. Proc Natl Acad Sci U S A 105: 417–421.

Smith, J.J., Sorensen, A.G., and Thrall, J.H. (2003) Biomarkers in imaging: Realizing radiology's future. Radiology 227: 633–638.

Solier, C., and Langen, H. (2014) Antibody-based proteomics and biomarker research-current status and limitations. Proteomics 14: 774–783.

Song, L., Hobaugh, M.R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J.E. (1996) Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. Science (80-) 274: 1859–1866.

Srinivas, P.R., Kramer, B.S., and Srivastava, S. (2001) Trends in biomarker research for cancer detection. Lancet Oncol 2: 698–704.

Srivastava, T., Virk, S., Kumar, A., and Ganguli, S. (2021) The role of electronic filters in biomedical applications. In Electronic Devices, Circuits, and Systems for Biomedical Applications. Academic Press, pp. 309–324.

Stein, U., Meller, A., Rozevsky, Y., Gilboa, T., Kooten, X.F. Van, Kobelt, D., and Huttner, D. (2020) Quantification of mRNA expression using single-molecule nanopore sensing. ACS Nano 14: 13964–13974.

Steinbock, L.J., Lucas, A., Otto, O., and Keyser, U.F. (2012) Voltage-driven transport of ions and DNA through nanocapillaries. Electrophoresis 33: 3480–3487.

Steinbock, L.J., Otto, O., Chimerel, C., Gornall, J., and Keyser, U.F. (2010) Detecting DNA folding with nanocapillaries. Nano Lett 10: 2493–2497.

Stenman, U.H. (2016) Biomarker development, from bench to bedside. Crit Rev Clin Lab Sci 53: 69–86.

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

Strimbu, K., and Tavel, J.A. (2010) What are Biomarkers? Curr Opin HIV AIDS 5: 463.

Suganuma, M., Yamaguchi, K., Ono, Y., Matsumoto, H., Hayashi, T., Ogawa, T., et al. (2008) TNF-α-inducing protein, a carcinogenic factor secreted fromH. pylori, enters gastric cancer cells. Int J Cancer 123: 117–122.

Suryavanshi, A.P., and Yu, M.-F. (2007) Electrochemical fountain pen nanofabrication of vertically grown platinum nanowires. Nanotechnology 18: 105305.

Sze, J.Y.Y., Ivanov, A.P., Cass, A.E.G., and Edel, J.B. (2017) Single molecule multiplexed nanopore protein screening in human serum using aptamer modified DNA carriers. Nat Commun 8: 1–10.

Tabard-Cossa, V., Trivedi, D., Wiggin, M., Jetha, N.N., and Marziali, A. (2007) Noise analysis and reduction in solid-state nanopores. Nanotechnology 18: 305505.

Taeb, A.M., Hooper, M.H., and Marik, P.E. (2017) Sepsis: Current definition, pathophysiology, diagnosis, and management. Nutr Clin Pract 32: 296–308.

Talaga, D.S., and Li, J. (2009) Single-molecule protein unfolding in solid state nanopores. J Am Chem Soc 131: 9287–9297.

Taqieddin, A., Nazari, R., Rajic, L., and Alshawabkeh, A. (2017) Review—Physicochemical Hydrodynamics of Gas Bubbles in Two Phase Electrochemical Systems. J Electrochem Soc 164: E448–E459.

The Editors of Encyclopedia Britannica (2020) antibody | Definition, Structure, Function, & Types | Britannica. https://www.britannica.com/science/antibody. Accessed August 2, 2021.

The UK Sepsis Trust (2019) The UK SEPSIS Trust FAQs.

Tiede, C., Bedford, R., Heseltine, S.J., Smith, G., Wijetunga, I., Ross, R., et al. (2017) Affimer proteins are versatile and renewable affinity reagents. Elife 6: 24903.

Tiwari, A., and Gong, S. (2009) Electrochemical detection of a breast cancer susceptible gene using cDNA immobilized chitosan-co-polyaniline electrode. Talanta 77: 1217–1222.

Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science (80-) 249: 505–510.

U.S. Food & Drug Administration (2019) About Biomarkers and Qualification. fda.gov https://www.fda.gov/drugs/biomarker-qualification-program/about-biomarkers-and-qualification. Accessed July 26, 2021.

Umehara, S., Karhanek, M., Davis, R.W., and Pourmand, N. (2009) Label-free biosensing with functionalized nanopipette probes. Proc Natl Acad Sci U S A 106: 4611–4616.

Umehara, S., Pourmand, N., Webb, C.D., Davis, R.W., Yasuda, K., and Karhanek, M. (2006) Current rectification with poly-L-lysine-coated quartz nanopipettes. Nano Lett 6: 2486–2492.

Vasilescu, G. (2006) Electronic Noise and Interfering Signals. 1st ed., Springer-Verlag Berlin and Heidelberg GmbH & Co. K, 1–176.

Vliet, E.A. van, Puhakka, N., Mills, J.D., Srivastava, P.K., Johnson, M.R., Roncon, P., et al. (2017) Standardization procedure for plasma biomarker analysis in rat models of epileptogenesis: Focus on circulating microRNAs. Epilepsia 58: 2013–2024.

Vuković, P., Peccatori, F.A., Massarotti, C., Miralles, M.S., Beketić-Orešković, L., and Lambertini, M. (2021) Preimplantation genetic testing for carriers of BRCA1/2 pathogenic variants. Crit Rev Oncol Hematol 157: 103201.

Wadhawan, J.D., and Compton, R.G. (2013) Electrochemistry. 12th ed., Royal Society of Chemistry, 198–234 .

Wand, N.O., Smith, D.A., Wilkinson, A.A., Rushton, A.E., Busby, S.J.W., Styles, I.B., and Neely, R.K. (2019) DNA barcodes for rapid, whole genome, single-molecule analyses. Nucleic Acids Res 47: e68–e68.

Wang, Y., Montana, V., Grubišić, V., Stout, R.F., Parpura, V., and Gu, L.Q. (2015) Nanopore sensing of botulinum toxin type B by discriminating an enzymatically cleaved peptide from a synaptic protein synaptobrevin 2 derivative. ACS Appl Mater Interfaces 7: 184–192.

Wang, Z., Liu, Y., Yu, L., Li, Y., Qian, G., and Chang, S. (2019) Nanopipettes: A potential tool for DNA detection. Analyst 144: 5037–5047.

Wanunu, M. (2012) Nanopores: A journey towards DNA sequencing. Phys Life Rev 9: 125–158.

Wanunu, M., Morrison, W., Rabin, Y., Grosberg, A.Y., and Meller, A. (2010) Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. Nat Nanotechnol 5: 160–165.

Welsh, L., Maleszka, R., and Foret, S. (2017) Detecting rare asymmetrically methylated cytosines and decoding methylation patterns in the honeybee genome. R Soc Open Sci 4: 170248.

Wen, L., and Tang, F. (2014) Genomic distribution and possible functions of DNA hydroxymethylation in the brain. Genomics 104: 341–346.

Wild, D., John, R., and Sheehan, C. (2013) *The Immunoassay Handbook Theory and applications of ligand binding , Edited by Associate Editors*. 4th ed., Elsevier Ltd, 29–242.

Wilkinson, A.A., Jagu, E., Ubych, K., Coulthard, S., Rushton, A.E., Kennefick, J., et al. (2020) Site-Selective and Rewritable Labeling of DNA through Enzymatic, Reversible, and Click Chemistries. ACS Cent Sci 6: 525–534.

Wilson, J., Sarthak, K., Si, W., Gao, L., and Aksimentiev, A. (2019) Rapid and Accurate Determination of Nanopore Ionic Current Using a Steric Exclusion Model. ACS Sensors 4: 634–644.

Woodcock, D.M., Lawler, C.B., Linsenmeyer, M.E., Doherty, J.P., and Warren, W.D. (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. J Biol Chem 272: 7810–7816.

Woof, J.M., and Burton, D.R. (2004) Human antibody–Fc receptor interactions illuminated by crystal structures. Nat Rev Immunol 2004 42 4: 89–99.

Xu, X., Clément, P., Eklöf-Österberg, J., Kelley-Loughnane, N., Moth-Poulsen, K., Chávez, J.L., and Palma, M. (2018) Reconfigurable Carbon Nanotube Multiplexed Sensing Devices. Nano Lett 18: 4130–4135.

Xue, L., Yamazaki, H., Ren, R., Wanunu, M., Ivanov, A.P., and Edel, J.B. (2020) Solid-state nanopore sensors. Nat Rev Mater 5: 931–951.

Yang, L., and Yamamoto, T. (2016) Quantification of virus particles using nanopore-based resistive-pulse sensing techniques. Front Microbiol 7: 1500.

Yaul, M., Bhatti, R., and Lawrence, S. (2008) Evaluating the process of polishing borosilicate glass capillaries used for fabrication of in-vitro fertilization (iVF) micro-pipettes. Biomed Microdevices 10: 123–128.

Yoshio Hirao, Tsugiyasu Kanda, Yoshimasa Aso, Masato Mitsuhashi, and Isao Kobayashi (2000) Interleukin-8—An Early Marker for Bacterial Infection. Lab Med 31: 39–44.

Yu, R.-J., Ying, Y.-L., Gao, R., and Long, Y.-T. (2019) Confined Nanopipette Sensing: From Single Molecules, Single Nanoparticles, to Single Cells. Angew Chemie Int Ed 58: 3706–3714.

Yue, L., Amelia, Y., Cass, T., and Albrecht, T. (2017) A DNA Assay On a String – Fabrication and Sensing. Thesis.

Zhang, H., Chon, C.H., Pan, X., and Li, D. (2009) Methods for counting particles in microfluidic applications. Microfluid Nanofluidics 7: 739–749.

Zhang, S., Li, M., Su, B., and Shao, Y. (2018) Fabrication and Use of Nanopipettes in Chemical Analysis. Annu Rev Anal Chem 11: 265–286.

Zinellu, A., Arru, F., Vito, A. De, Sassu, A., Valdes, G., Scano, V., et al. (2021) The De Ritis ratio as prognostic biomarker of in-hospital mortality in COVID-19 patients. Eur J Clin Invest 51: e13427.

Appendix 1 – Other explored projects

In addition to the main project, 3 further projects were explored. These were designed to integrate with the larger body of work, but due to time constraints it was not possible to finalize the programs and techniques described. As a part of SCoNE development, a simulation script was written to create artificial translocation events to assist with the development of extraction and identification of experimental data. During the development of SCoNE extraction methods, isolation of SCoNE structures from more complex patient samples was investigated. This work involved applying a bias differential across a filter membrane to remove larger contaminants and move only the DNA to the site of sampling. Finally, the imaging of SCoNE DNA using atomic force microscopy (AFM) imaging was explored to confirm the structure of the constructs.

A1.1 Simulated data

MATLAB code was designed to generate data to assist with the development of SCoNE substructure identification. The development consisted of 4 main parts: development of a graphic user interface for designing the experiment, creation of a background signal generator with interference, event design and creation for possible DNA conformations, and comparable results to experimental data optimisation. The current code is in development and the analysis results are described here.

A1.1.1 Construction of the graphical user interface

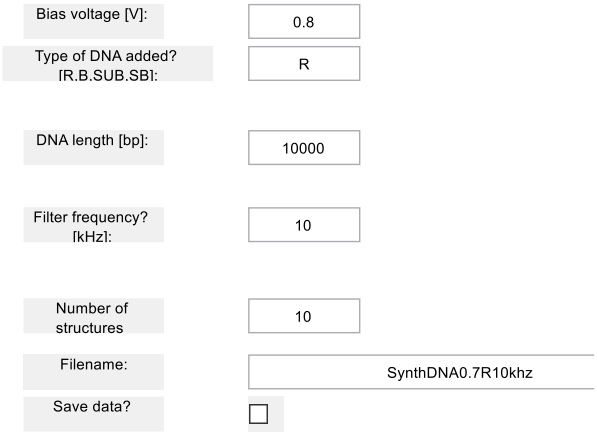

Figure A1.1 Example of the graphic user interface

As illustrated in Figure A1.1, the graphical user interface is capable of accepting and interpreting information provided by the user including bias voltage, DNA event desired, length of DNA, filter frequency desired, number of substructures expected, and a file name creator. Where dsDNA only is expected, the code will ignore the number of structures input.

A1.1.2 Background signal generator

During the design of the code, it was necessary to create a baseline noise generator comparable to that experienced during experimental work. This was achieved through RMS analysis of experimentally acquired data. The values were then applied as positive and negative limits to the noise signal generator with a random function used to create values in-between these points, as illustrated in Figure A1.2.A. During experiments, there are additional noise spikes observed whose generation cannot be readily identified. Therefore, to contribute to the realism of the data, additional noise spikes were added mimicking a random distribution of short duration spikes illustrated in Figure A1.2.B. The intensity and duration for these spikes were obtained from experimental results.

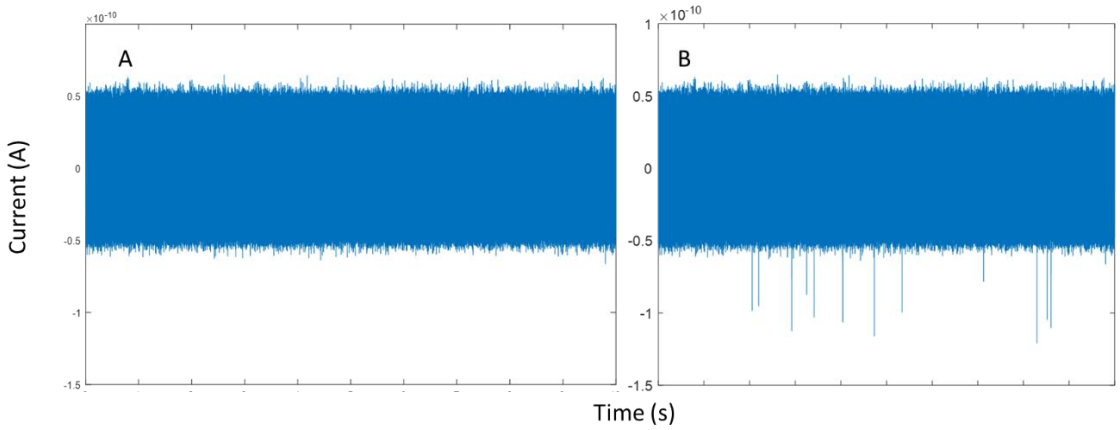

Figure A1.2. Example of simulated data with no artificial spikes (A) and artificial spikes (B) introduced using a 10 kHz filtered data.

A1.1.3 Translocation event simulation

From optimisation of the code and both manual analysis and data extraction from experimental results, it was possible to create a synthetic data generator that could simulate any of the 4 main translocation event types as illustrated in Figure A1.3.

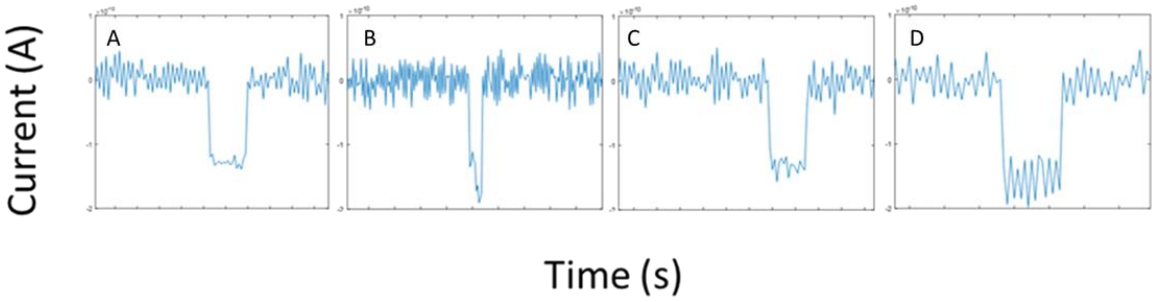

Figure A1.3. Illustration of the 4 main types of DNA translocation events, bare dsDNA (A), folded dsDNA (B), SCoNE DNA with unbound probes attached (C), and SCoNE DNA with bound probes (D) all at a bias of -0.8 V and a simulated filter frequency of 10 kHz.

A1.1.4 In silico translocation results

Data acquired from the authors research was used as a base with data previously acquired within the group used for corroboration (Loh *et al.*, 2018) to create *in silico* translocation data. In Figure A1.4, it can be observed that the translocation and noise data synthesised is too compartmentalised when compared to experimental results. While the data generation script requires further development to match more realistic

results, the DNA translocations follow the same trends observed in experimental data, as illustrated in Figure A1.5.

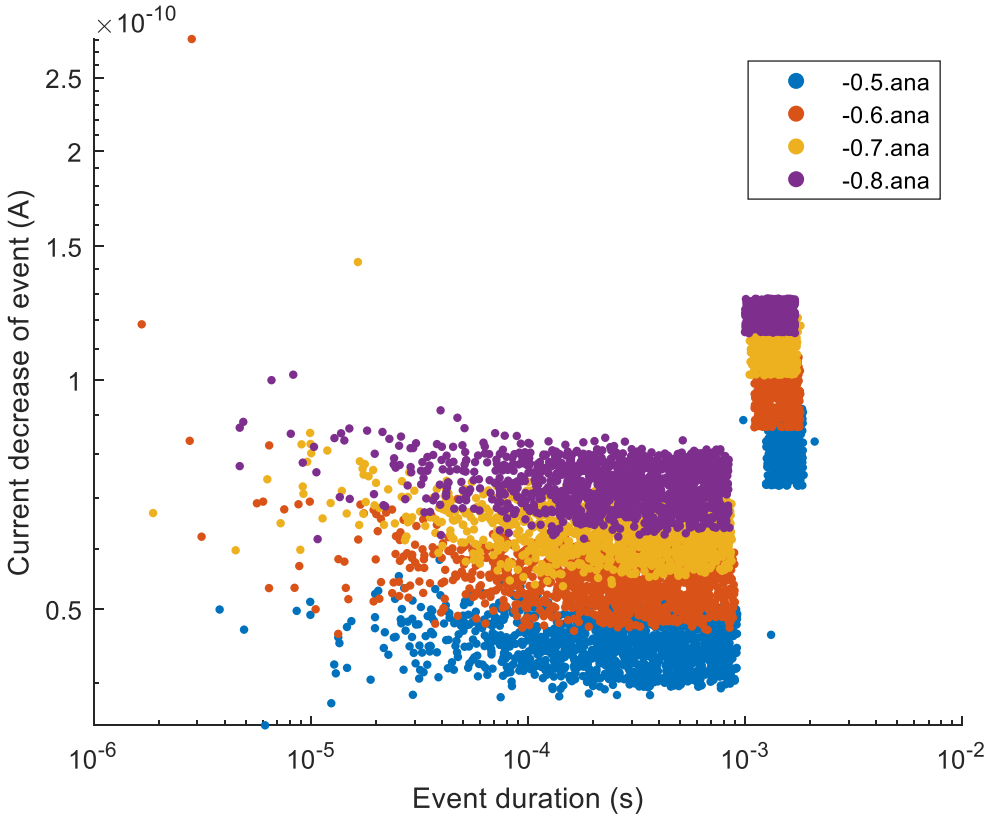

Figure A1.4. In silico event extraction results from 400 scans (100 per bias) illustrating the dependence of event duration and current decrease by bias (filter frequency: 10 kHz; 10 kbp [DNA] = 300pM).

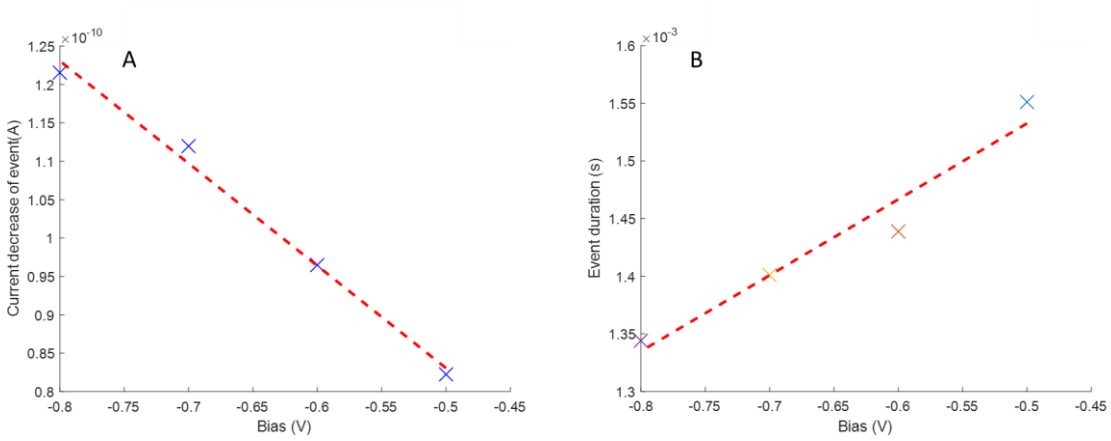

Figure A1.5. In silico results illustrating the increase in intensity of translocation peaks with bias (A) and the decrease in event duration with an increase in bias (B), (filter frequency: 10 kHz; 10 kbp [DNA] = 300pM).

A1.2 DNA filtration

Patient samples are a complex mix of cells, proteins, circulating DNA and RNA, and metabolites. Isolation of DNA from a complex mixture is often time consuming or is isolated with other contaminants. This presents an issue with single molecular sensing. The research conducted aimed to create a filter device which could integrate with nanopore sensing and SCoNE's biomarker capture properties for direct DNA sampling. It was decided that the sample should be placed below the isolation vial and DNA forced to move upwards to prevent biomolecules following the flow of gravity and contaminating the isolate. Due to this approach, the technique was named DNA anti-gravity translocation apparatus (DATA).

A1.2.1 Methodology

A 0.22 µm syringe filter was saturated using 1 mL of 10% skimmed milk solution in 4 M LiCl. The syringe filter was then washed twice using 20 ml of nuclease free water then using 10 ml of 4 M LiCl. 2, 1 mL syringes had the plungers removed, and 1 of the syringes was heat sealed at the needle attachment site, such that no liquid could escape. The filter was then fixed between the 2 syringes in a custom-made holding cell, as illustrated in Figure A1.6. Syringes were filled with 1 mL of 4 M LiCl. A 10 kbp DNA fragment was used for experimental purposes and was added into the lower syringe at 300 pM, comparable to translocation experiments previously conducted. Ag/AgCl electrodes were inserted into both syringes and a voltage of 1 V applied across the filter, such that the upper electrode had a positive charge.

From the upper syringe, samples of 50 µL were collected every 15 minutes for 1 hour, then at hour 2, and finally at 2.5 hours, replacing the removed sample with 4 M LiCl. Samples were then incubated with 2 µL of SYBR gold and heated to 96°C for 2 minutes and allowed to cool to 60°C in a water bath for 10 minutes. Samples were then measured on a fluorimeter (Cary Eclipse Fluorescence Spectrophotometer) with excitation at 498 nm and emission measured at 540 nm. Excitation slit width was set at a value of 10.

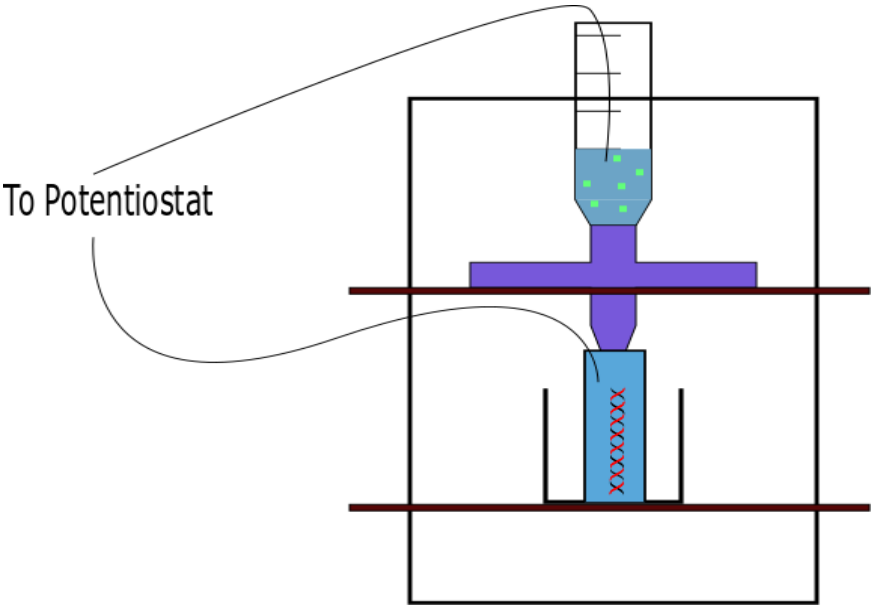

Figure A1.6. Example drawing of DATA constructed

A fluorescence calibration using 10 kbp DNA was also performed. From the 300 pM stock solution, the DNA was diluted from 300 pM to 4.69 pM in a serial dilution manner, using 4 M LiCl as the dilutant.

A1.2.2 Results and discussion

From the results of the calibration curve show that the method for DNA determination was successful, as illustrated in Figure A1.7, and that the use of 4 M LiCl as the running buffer did not affect the fluorescence measurements, even at low concentrations, as

illustrated in Figure A1.8. This initially was a concern as LiCl is not recommended as a buffer for SYBR gold.

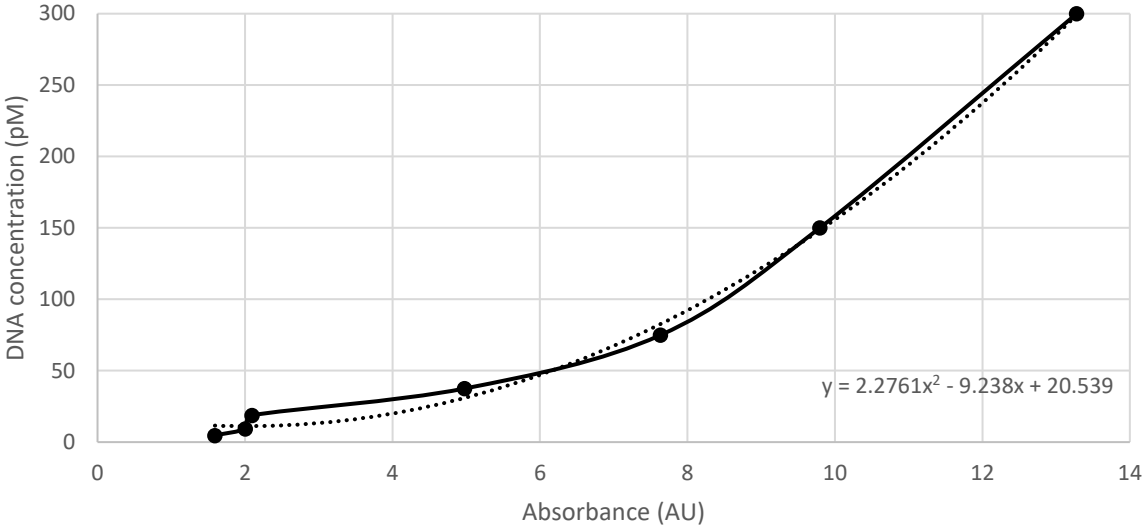

Figure A1.7. Calibration curve generated for absorbance against DNA concentration, with the equation of the line displayed.

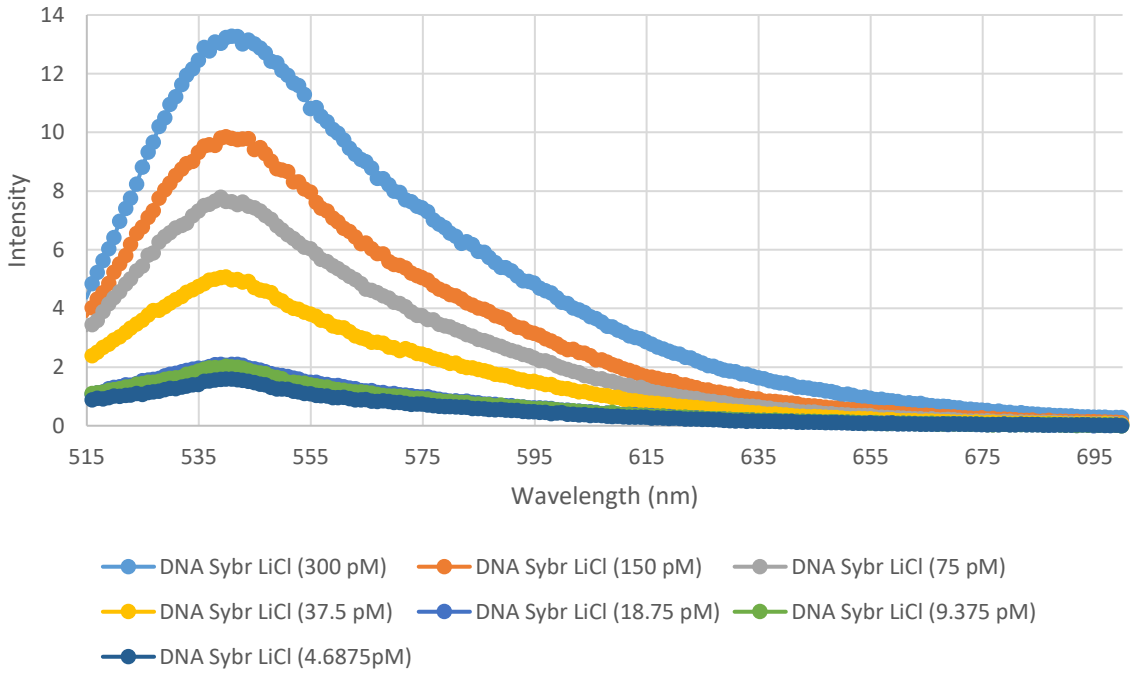

Figure A1.8. Fluorescence emission intensity using different DNA concentrations.

From the results of the DNA purification experiment, it is possible to conclude that DNA could be isolated using the method developed, as shown in Figure A1.9. Using the equation of the line for the calibration curve it was possible to calculate the

concentration isolated from the experimental set up. When analysing the concentration of each individual isolate, it is possible to see that concentration remained approximately the same across all time points. When analysing the cumulative concentration obtained over time, it is possible to conclude that 74.31 pM DNA was isolated from the initial vial over 2.5 hours.

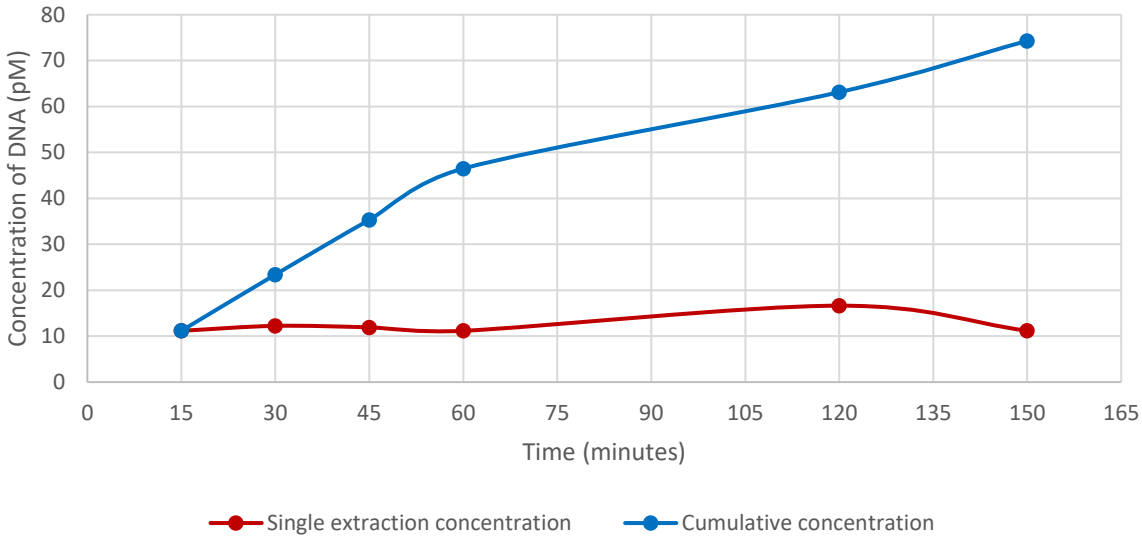

Figure A1.9. Concentration over time calculated using the calibration curve for single extraction timepoints (red) and the accumulative concentration (blue).

From the results of this experiment, it can be suggested that while partially successful, the final yield was 24.77%. Whilst the results of this proof-of-concept experiment show that it is possible to isolate DNA, and could be integrated with nanopore sensing, the long incubation time and low yield indicate that further optimisation is required. It can be suggested that using a higher voltage could improve speed of isolation and yield, however higher voltages would not interface directly with current single molecule sensing equipment. Therefore, potential redesign to include a third electrode in the upper syringe could allow for isolation at higher voltages and translocation at lower

voltages to be conducted independently. Altering the design to handle lower volumes could also prove to be useful in further developments.

A1.3 AFM imaging of SCoNE structures

AFM is a scanning probe technique for molecular imaging with a high, nanometre resolution. AFM imaging employs a cantilever with a super sharp tip which scans over the surface of a sample. A laser is focused on the tip and a photodiode used as the detector, with a piezo crystal controlling the fine movement. Imaging is split into contact and tapping mode. Contact imaging involves dragging the tip over the surface, using the deflection of the laser or feedback from the piezo to determine the surface contours. In tapping mode, the cantilever is oscillated at its resonance frequency using the piezo crystal. The tip is held at a constant distance from the surface, and as the surface contours change, the influence of different forces on the tip causes the amplitude of oscillation to change. The amplitude feedback is used to adjust the height of the tip creating an image of the surface (Binnig, 1988). AFM has many applications for biomolecule research, including DNA imaging and protein study (Kreplak, 2016).

A1.3.1 Methodology

Unbound SCoNE concentration was determined by UV-vis spectroscopy and the absorbance value at 260 nm used in the Beer-Lambert Law to determine DNA concentration. The DNA was suspended in a 10% 1 mM Tris-HCl, 10 mM NaCl solution. AFM magnetic sample plates were prepared by using double sided tape to adhere a 12 mm, mica disk to its surface. A minimum of 3 layers were stripped from the surface of the mica and 40 µl of a 100 mM $NiCl_2$ pipetted onto the surface. This was left in a covered glass petri dish at room temperature for 1 hour. The surface was then washed gently 3 times using nuclease-free water and dried using flowing nitrogen

gas. The surface was then coated with 40 µl of the SCoNE solution and allowed to adhere for 2 hours in a covered petri dish at room temperature. The surface was then gently washed 3 times using nuclease-free water. The plate was then dried using flowing nitrogen gas and left overnight in a vacuum desiccator to ensure as little moisture remained as possible. The surface was then imaged using a Keysight 9500 AFM, using AC tapping mode and 10% I gain.

A1.3.2 Results and discussion

The results obtained from AFM imaging proved to be inconclusive. Whilst it was possible to image trimer SCoNE structures, the DNA fragments coagulated together during deposition, making imaging single structures difficult, as highlighted in Figure A1.10. This occurred over several experiments with A1.10.A representing the initial AFM imaging experiment, and A1.10.B representing an experiment performed after a separate assembly reaction.

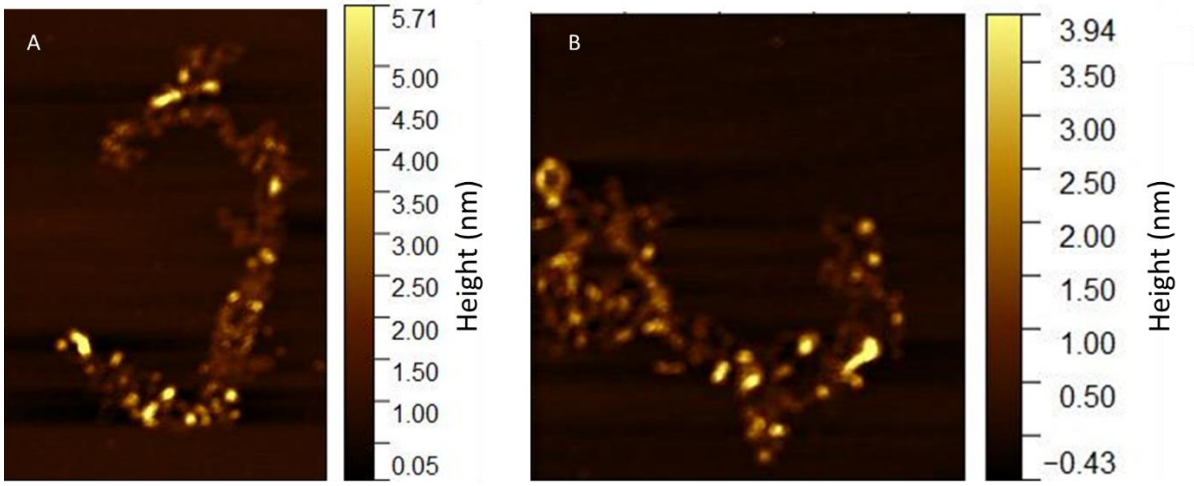

Figure A1.10. AFM images highlighting timer SCoNE structures coagulated together for N=1 (A) and N=2 (B).

When analysing the AFM images, it was possible to isolate potential SCoNE structures by using the 0 nm height as a condition to separate out strands. As illustrated in Figure

A1.11, a few isolates were located, and their length measured. Assuming perfect linearity of the DNA backbone, the length should be equal to 1.19 µm. The lengths of the probe (approximately 5 nm) and the linker group (3.28 nm) should equal 8.28 - 8.62 nm overall. However due to folding of the probe and linker groups, the size is assumed to be approximately 4X smaller than the unfolded structure (Ivanov *et al.*, 2014; Xu *et al.*, 2018). This is confirmed by the size difference between the sequence length and the folded aptamer (approximately a decrease from 20 nm to 5 nm respectively). Therefore, the length of a probe would be expected to be 2.07 - 2.16 nm in length. Due to the coiling of the DNA, length measurements were taken following the curvature of the backbone observed. The results of the isolates show that the observable backbone have lengths of 1.35, 1.19, and 1.14 µm respective to Figure A1.11.A, B and C. Whilst this cannot be absolutely confirmed, this is a good indicator that the fragments observed have similar lengths to the expected size of the DNA backbone. This provides further evidence that the assembly method performs as previous experiments indicated.

From the imaging experiments conducted, it was possible to measure one of the probe groups attached to the DNA backbone. A probe group from Figure A1.11.C was measured with a length of 2.82 nm. The length measured was similar to that expected from a probe group, however due to the deposition conditions, it is possible that some unfolding of the structures occur. Whilst the resolution does not allow for complete conformation, this measurement is a good indicator that a probe group is attached. The increase in hight at the site of the probe group is also similar to that observed in other AFM aptamer imaging experiments of similar sized aptamers (Ivanov *et al.*,

2014). This provides further evidence that the SCoNE method protects the probes during assembly steps.

Whilst it was not possible to completely confirm individual shapes of SCoNE structures through AFM imaging, the images indicate that the size, height, and shape of the structures are similar to those expected.

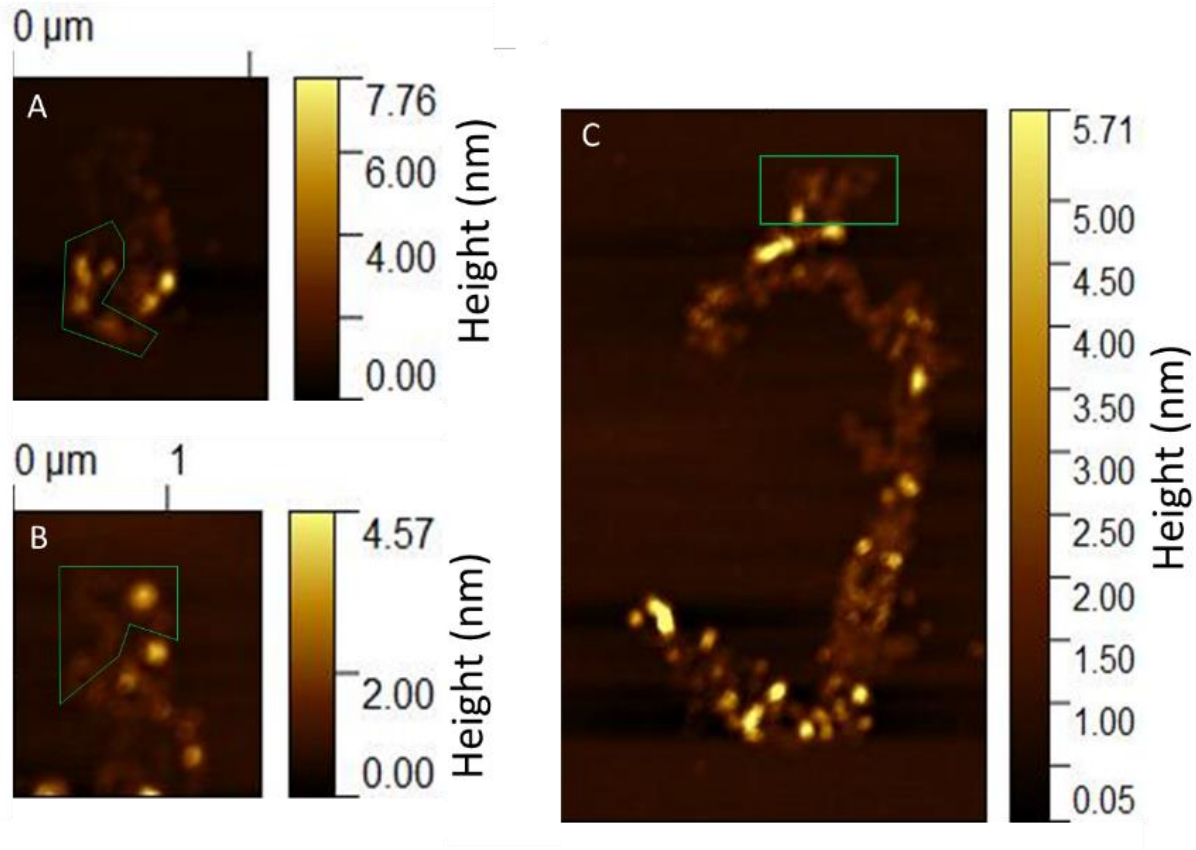

Figure A1.11. AFM images of SCoNE structures with possible single structures highlighted in green boxes (N=1, 2, and 3 respective to A, B and C).

Appendix 2 – SCoNE sequences and additional information

A2.1 Sequence generating code

```
data = {};
v = {'A','C','G','T'};
while length(data) < 10
   ind = randsample(4,500,1);
   V = strjoin(v(ind),'');
   if ~contains(V,'ACGT')
      if ~contains(V,'TGCA')
         data{length(data)+1} = [V,flip(V)];
      end
   end
end
```

A2.2 pDNA fragment sequence

pDNA/ iDNA fragment	Sequence (5'-3')
1	CTGCAACGCAGCGCTTGTAGGATTCACATCGAACGGATTCAGTGCTTATCGCATAGACAG
2	GAACGGCCCGGTGTGCTTGTGATTCTAGTCGACGCGATTCAACGTTTGGAAATCCCATAC
3	GTATTGAGGGGTGTGCCACTGATTCTGATCGAAACGATTCTAGCGCACGACCTGGCAGAG
4	GATCAATGCTCCAACGAACTGATTCATCTCGAGGCGATTCAAAGCGGGAATGATACGGCG
5	AACCACTATAACTAGAGCAGGATTCACATCGAAGCGATTCACGTGTTGGCACCGGATATG
6	CTGTAGCAGCCTGAGTAGGTGATTCAATTCGATATGATTCGCGCTCTCTAATTTGCGTCG
7	TGCCAGTCTTGGCCTCTTCAGATTCGTGTCGACCAGATTCATTCGTGCATTACGTTATCG
8	TCGGCTGACTAGAGAATAGGGATTCTTGTCGACCCGATTCCCCCGCTTCTGCTTGAACAC
9	CGGACAAAGAGGCAGCGAATGATTCATGTCGACCCGATTCGGTTAGGACCGTCAGTTATG
10	ATTGACACTATTAGTCCAGAGATTCAATTCGATAAGATTCCGCAGGACCGTGCTCGTAGA

A2.3 sDNA fragment sequence

sDNA fragment	Sequence (5'-3')
1	AGTGCTTATCGCATAGACAGCGAGGACACTTGGCACGACCAAGCATACTTGTGA GCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGAT AGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGCA TCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAAG ATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCGC CGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTACC TGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGTG TGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAACA ACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTACA AATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCTG AAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGAT GATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAATT ATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTCA ATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTGT GGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCTT TACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATACC CCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGGA AGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGT GAACGGCCCGGTGTGCTTGT
2	AACGTTTGGAAATCCCATACCGAGGACACTTGGCACGACCAAGCATACTTGTGA GCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGAT AGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGCA TCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAAG ATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCGC CGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTACC TGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGTG TGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAACA ACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTACA AATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCTG AAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGAT GATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAATT ATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTTA ATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTGT GGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCTT TACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATACC CCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGGA AGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGT GTATTGAGGGGTGTGCCACT
3	TAGCGCACGACCTGGCAGAGCGAGGACACTTGGCACGACCAAGCATACTTGTG AGCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGA TAGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGC ATCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAA GATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCG CCGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTAC CTGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGT GTGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAAC AACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTAC AAATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCT GAAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGA TGATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAAT TATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTT AATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTG TGGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCT

	TTACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATAC CCCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGG AAGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGT GATCAATGCTCCAACGAACT
4	AAAGCGGGAATGATACGGCGCGAGGACACTTGGCACGACCAAGCATACTTGTG AGCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGA TAGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGC ATCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAA GATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCG CCGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTAC CTGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGT GTGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAAC AACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTAC AAATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCT GAAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGA TGATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAAT TATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTT AATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTG TGGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCT TTACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATAC CCCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGG AAGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGT AACCACTATAACTAGAGCAG
5	ACGTGTTGGCACCGGATATGCGAGGACACTTGGCACGACCAAGCATACTTGTGA GCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGAT AGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGCA TCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAAG ATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCGC CGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTACC TGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGTG TGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAACA ACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTACA AATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCTG AAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGAT GATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAATT ATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTTA ATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTGT GGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCTT TACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATACC CCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGGA AGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGTC TGTAGCAGCCTGAGTAGGT

6	GCGCTCTCTAATTTGCGTCGCGAGGACACTTGGCACGACCAAGCATACTTGTGA GCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGAT AGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGCA TCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAAG ATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCGC CGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTACC TGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGTG TGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAACA ACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTACA AATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCTG AAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGAT GATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAATT ATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTTA ATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTGT GGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCTT TACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATACC CCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGGA AGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGTT GCCAGTCTTGGCCTCTTCA
7	ATTCGTGCATTACGTTATCGCGAGGACACTTGGCACGACCAAGCATACTTGTGA GCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGAT AGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGCA TCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAAG ATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCGC CGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTACC TGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGTG TGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAACA ACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTACA AATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCTG AAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGAT GATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAATT ATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTTA ATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTGT GGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCTT TACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATACC CCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGGA AGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGTT CGGCTGACTAGAGAATAGG
8	CCCCGCTTCTGCTTGAACACCGAGGACACTTGGCACGACCAAGCATACTTGTGA GCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGAT AGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGCA TCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAAG ATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCGC CGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTACC TGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGTG TGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAACA ACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTACA AATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCTG AAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGAT GATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAATT ATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTTA ATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTGT GGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCTT TACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATACC CCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGGA AGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGTC GGACAAAGAGGCAGCGAAT

9	GGTTAGGACCGTCAGTTATGCGAGGACACTTGGCACGACCAAGCATACTTGTGA GCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGAT AGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGCA TCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAAG ATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCGC CGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTACC TGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGTG TGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAACA ACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTACA AATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCTG AAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGAT GATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAATT ATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTTA ATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTGT GGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCTT TACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATACC CCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGGA AGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGTA TTGACACTATTAGTCCAGA
10	CGCAGGACCGTGCTCGTAGACGAGGACACTTGGCACGACCAAGCATACTTGTG AGCGGAAGTCTCGCAATCCTTCGTGCCAGGACCTGGCGGCCCGTGCTCGCTGA TAGCCGTTAAAATGGTTTTGGGGTTGTTCATCGGACACTATCCCACTACTATTGC ATCACGTACACTACACCCCTGGGGAGTAATCGGGACTGAAACCCACTTCGTGAA GATATGCGAGATCTCATAAACGTACGGCATGGTACACGACTATATTCATGCCCCG CCGTCATATGAAAGGGGAACATTATGAAAACAATGTGTAGTATCCCGCCACCTAC CTGTCTTTTCAGGCTAAGGCGCGCGACACAGGATTGGGAAGGATGCCCATATGT GTGCTCATCTCAAACAACCTTAACCCATTAGCCAGGTTACAGCAGAGCGCATAAC AACGACGGACCAGCATGATCTTCCAACGAACACTCTCTTGCAACGACACGTTAC AAATAACGTGCGTCACGCTGTTACATGTTTTCTAGCATCAGCGGGCTTGAGACCT GAAGCAACCTTGTATCTAGGATCTCGCCGGAAATCTACGAACCTCACCCGACGA TGATACTTTGTTGCCCAGAGCGAGGACACTCAATAGACCTTTCAGACTGTAGAAT TATCACACCAAAGAGTCAGCACGGTTTCGCGAGAAATGCACACGGCGGTCGCTT AATCCATGGTATAATTATTCACTGCTCAGCGTAGGGTCATGAACTGTGGCGCTTG TGGCTCCCCGTGAAGTACAATGGGAACCGGCAACGGTCTAGCATGGGAATTTCT TTACTCTGCAATCGCGAGTATGAGAAAGTCGGCAACGCAGCCAGTGTAAAATAC CCCTTTCACGGGAATTCCGGAACACAGCTCGCGGTGGCTAAATCGCCCTAAGGG AAGGGTGTTCCAACTTCAAATTTGTAAATTTTCTTAATTGTATATAAAGTTTCCGGT AATTCGTCGGACTGCTGGAT

Appendix 3 – Conductance data analysis

```
clear all;close all;clc;
dn = 'C:\Users\OXI762\Documents\PhD Electro\Conductance Data\Oliver\20181211'
;%file location
cd(dn) %access data
basename = input('file number?','s'); %starting point for analysis
Y = []; %empty matrix for average current input
for i = 1:5
    filename = sprintf('%s_%d',basename,i); %assign the filename
    tmp = importdata(filename); %import the data for one file
    data = tmp.data; %save the data
    y = data(:,3); %pull out the current values
    Y = [Y y]; %assign the values and increase the size of the matrix with each loop
end
X = data(:,2); %extract the bias values
y = mean(Y,2); %average the current per bias value
xlabel 'Bias (V)'
ylabel 'Current (A)'
plot(X,y); %plot the average of the 5 scans
[m,err] = fit(X,y,'poly1') %fit a linear average line to the plot
Q = coeffvalues(m); %extract the gradient of the line
q = Q(1);
hold on;
plot(m);
G = q; %conductance
LENG = input('length in um?'); %input the length obtained from microscope readings
I = (LENG/1000000); %taperlength conversion to m
Di = 0.0005; %innerdiameter 0.5 mm= 0.0005
gc = 17.3; %conductivity of 4M LiCl mS cm-1
Num = ((4*(G*I))+((pi/2)*(G*Di))); % Create the numerator value to the equation
Den = (Di*pi*gc)-((pi/2)*G); % create the denominator for the equation
dpore = Num/Den; %perform the calculation
dporenm = dpore*1e9; %multiply by 1x10^9 for values in nanometers
title(['Size is ' ,num2str(dporenm),'nm']); %create title for these values
ylabel({'Current (A)'}); % y label for current
% Create xlabel
xlabel({'Bias (V)'}); % x axis, bias values
hold off;
```

Appendix 4 – RMS and FT analysis code

A4.1 RMS

```
Figure
   for G =1:size(X,2) %calculate based on number of bias values for import
      FileName3= sprintf('%s_%d%s',basename,fileidentout{1,G},'.mat'); % create the
file name from data import
      load (FileName3); %load in the data
      RMS (G)= rms(ch2(3000000:4000000)); %calculate the RMS values between 3
and 4 seconds
   end
   x = []; %blank x variable
   y = []; %blank y variable
   for G =1:size(X,2)
      x (G) = biasorig(1,G); %import bias name based on repeat
      y (G) = RMS(1,G); % import rms values based on repeat value
      scatter (x,y, 200, 'x') %scatter plot the results
      hold on %hold on to the plot
   end
g1 = fit(x',y','poly1') % fit a linear fit line
plot (g1, x,y); %plot the line
ylabel({'Noise (pA)'});
xlabel('Bias (V)');
legend('off');
hold off
```

A4.2 FT

```
Fs = 1000000;           % Sampling frequency (10000000 for 10 s for)
T = 1/Fs;             % Sampling period
L = 1e7;             % Length of signal
t = (0:L-1)*T;
Y = fft(ch2);
P2 = abs(Y/L);
P1 = P2(1:L/2+1);
P1(2:end-1) = 2*P1(2:end-1);
f = Fs*(0:(L/2))/L;
loglog(f,P1)
xlabel('Hz')
```

```
%% Cluster Time O.Irving
close all;clc;clear all
disp ('select files with .ana at the end then click open')
[FileNameorig,PathNameorig]=uigetfile('*.ana','Event Results','Multiselect','on');% find
the ana files for input
disp ('Click the folder containing all the.all files then click ok/enter')
PathNameraw = uigetdir(); %pathname for all raw unprocessed files
extrafolder = input('are the raw files stored in a different folder to the all files?','s');
if extrafolder == 'y'
    disp ('Click the folder containing all the raw .mat, .evt and .log files then click
ok/enter')
   Pathnamerawmatfiles = uigetdir();
end
cd(PathNameraw);
SES = input('do you want to look at single events?','s');
%% Magic numbers time
decider = input('do you want to use FWHM?','s');
if decider == 'y'
    %use for raw DNA
    durationcolumn = 14; %charge column we use later
    currentcolumn = 16; %fwhm of current column
else
    % use for strutted DNA
    durationcolumn = 13; %duration 1s1s
    currentcolumn = 9; %1s1s duration
end
timecol = 1; %event start time column
F = 1;  %Figure starting point
gaussEqn = 'a*exp(-((x-b)/c)^2)+d' ;% gaussian equation
startvalue = 957; %the value of the -0.4bias all/ana file
basename = 'dataout'; %scan base name (you wrote this, it's the beginning of its
name before it starts numbering up)
polarity = -1;
cutoff =0.25e-10;
%% Useful cells/ matrix output location
biasorig = []; % the bias
ALLDATA = {}; %aLL ANA data to play with
XandYdata = {}; %extracted plotting values
XandYdatacorr = {}; %extracted plotting values
lgcombi = {}; %log values for values-- histo work
counts = {}; %histo counts for same scale
averagedur = {}; %average duration per bias
averageduri = {}; %intensity per bias
evtout = {}; %the index of the processed data, ie its location in the matrix
resfile = {}; %output res file with the event you picked
residx = {}; % the index of the row it came from
```

```
Y = [];%event intensity (defined in the loop by either column 13 or 16
X = []; %event duration
%% File locations
filelocations{1} = startvalue:startvalue+101; %-0.4
filelocations{2} = (startvalue-840):(startvalue-739); %-0.5
filelocations{3} = (startvalue-420):(startvalue-319);% -0.6
filelocations{4} = (startvalue-210):(startvalue-109); %-0.7
filelocations{5} = (startvalue-630):(startvalue-529); %-0.8
%% Basic data set up
% import the whole data
for G = 1:length(FileNameorig)
    ALLDATA {G} = importdata([PathNameorig,FileNameorig{G}]); %all ana files
import
end
% import the filenames
biasorig = [];
for G = 1:length(FileNameorig)
    fname = FileNameorig{G}; %extract filename
    biasorig = [biasorig sscanf(fname,'%g')]; %creates bias variable by extracting only
the numbers from the filename
end
for G = 1:length(FileNameorig)
    data = cell2mat(ALLDATA (:,G)); %imports all data for particular bias into a more
useable format
    xlp = data(:,durationcolumn); %imports data for duration
    if decider =='y' %FWHM
        N = abs((data(:,currentcolumn))*0.000001); %convert to absolute value in A
        O = abs(N./xlp); % convert charge to time
        X(1:length(xlp),G)= O; %creates X values for FWHM calculations
        Y(1:length(N),G) = N; %creates Y values for FWHM calculations
    else %1s1s
        X(1:length(xlp),G)= xlp; %creates X values for 1s1s calculations
        Y(1:length(xlp),G) = abs((data(:,currentcolumn))); %creates Y values for 1s1s
calculations
    end
end
Y(Y==0)=NaN; %replace 0's with NaN's to be ignored by the code later on
X(X==0)=NaN; %same as previous
%remove the NaN's
for G = 1:size(X,2) % repeats for the entire dataset
    COMBI = []; %set up combined information store
    indi = ~isnan(Y(:,G)); %creates an arbitrary index for eliminating any NAN's left so
they are removed later
    COMBI(:,1) = (X(:,G)); %access the x data from the initial X aquired
    COMBI(:,2) = (Y(:,G)); %access the x data from the initial X aquired
    COMBI = COMBI(indi,:); %remove the values from both datasets where one or
both of the values were NAN
    XandYdata{G} = COMBI;
```

```
end
%remove all start below 1000000 datapoints
corrdatas = {};
snuse = {}; %evt datapoints
sause = {}; %all file index above
cd(PathNameraw);
for G = 1:size(X,2)
    corrdata = [];
    mindata = 1000000;
    indi = find(ALLDATA{:,G}(:,2)<mindata);
    corrdata (:,1) = XandYdata{:,G}(:,1);
    corrdata (:,2) = XandYdata{:,G}(:,2);
    corrdata(indi,:) = [];
    corrdatas {G} = corrdata;
    FileName1=
sprintf('%s_%d%s',basename,filelocations{1,G}(1,1),'.mat.res_events.all'); %get the
all file name
    sn = importdata(FileName1);
    SA2 = ALLDATA{G};
    SN2 = sn(:,[2:end]);
    SN2(:,indi) = [];
    SA2 (indi,:) = [];
    snuse {G}= SN2;
    sause {G} = SA2;
end
%% Thresholding
%add x thresh!!
limdatas = {};
eventrawdata = {};
anadataout = {};
for G = 1:size(X,2)
    limdata = [];
    limdata2 = [];
    Figure (F)
    scatter (corrdatas{1,G}(:,1),corrdatas{1,G}(:,2),4,'filled')%show the scan so you can
decide
    set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log
scale
    dec2 = input ('limit value in pA?','s'); %the threshold limit
    dec3 = input ('upper limit in pA?','s');
    limitvalue = (str2double(dec2))*1E-12; %turn input value into s
    limitvalue2 = (str2double(dec3))*1E-12; %turn input value into s
    indi = find(corrdatas{1,G}(:,2)<limitvalue);
    limdata(:,1)=corrdatas{1,G}(:,1);
    limdata(:,2)=corrdatas{1,G}(:,2);
    limdata(indi,:) = [];
    SN = snuse {G};
    SN(:,indi) = [];
```

```
    SA = sause {G};
    SA (indi,:) = [];
    indi2 = find(limdata(:,2)>limitvalue2);
    limdata2 (:,1) = limdata(:,1);
    limdata2 (:,2) = limdata(:,2);
    limdata2(indi2,:) = [];
    SN(:,indi2) = [];
    SA (indi2,:) = [];
    limdatas {G} = limdata2;
    eventrawdata {G} = SN;
    anadataout {G} = SA;
end
%%
%plot result
Figure (F)
for G = 1:size(X,2)
    scatter
(limdatas{1,G}(:,1),limdatas{1,G}(:,2),4,'filled','DisplayName',(num2str(biasorig(G))))
%scatterplot all initial data
    set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log
scale
    hold on %hold on to plot all of the bias' together
end
legend show %show the legend for bias identification
xlabel('Event duration (s)')
ylabel('Event intensity (A)')
hold off
%% lets look at the scatter
Figure (F)
for G = 1:size(X,2)
    scatter
(corrdatas{1,G}(:,1),corrdatas{1,G}(:,2),4,'filled','DisplayName',(num2str(biasorig(G))))
%scatterplot all initial data
    set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log
scale
    hold on %hold on to plot all of the bias' together
end
legend show %show the legend for bias identification
xlabel('Event duration (s)')
ylabel('Event intensity (A)')
hold off
%% Individual bias' look
Figure (F+1)
for G = 1:size(X,2)
    subplot((size(X,2)),1,G)
    scatter
(limdatas{1,G}(:,1),limdatas{1,G}(:,2),4,'filled','DisplayName',(num2str(biasorig(G))))
%scatterplot all initial data
```

```
    set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log
scale
end
%% focus on one bias
decsing = input('do you want to analyse one specific bias?','s');
if decsing == 'y'
    Figure (F+2)
    %singlechoise = input('do you want to look atr a single bias?','s');
    %if singlechoise == 'y'
    clusterinfostorage = {};
    maxpeakheight = {};
    G = 5;
    numberofclusters = 1;
    %scatter
(corrdatas{1,G}(:,1),corrdatas{1,G}(:,2),4,'filled','DisplayName',(num2str(biasorig(G))))
%scatterplot all initial data
    C1 =
clusterdata(limdatas{1,G}(:,2),'Linkage','ward','SaveMemory','on','Maxclust',numberof
clusters);
    for i= 1:numberofclusters
        clusteredvalues = limdatas{1,G}(:,:);
        indexpullout = find(C1 == i);
        CIS = clusteredvalues(indexpullout,:);
        clusterinfostorage {1,i} = CIS;
        scatter( clusterinfostorage {:,i}(:,1),clusterinfostorage
{:,i}(:,2),4,'filled','DisplayName',(num2str(i))) %scatterplot all initial data
        hold on
    end
    legend ('Location','southeast')
    set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log
scale
    hold off
    Figure (F+3)
    for i= 1:numberofclusters
        subplot((size(X,2)),1,i)
        if size(clusterinfostorage {1,i} <2)
            clusterinfostorage {1,i}(2:10,2) = zeros;
        end
        tmpdta = log10(clusterinfostorage {1,i}(:,1));
        bsize =  (-11:0.05:0); %bin size
        options = fitoptions(gaussEqn); %sets gaussian equation for fit
        options.StartPoint = [100 -4 0.5 0.5]; %works in sets of 4, this is where the
equation starts a fit
        options.Lower = [-5 -11 0 0]; % to not go below this value
        countsub = histc(tmpdta,bsize) ; %get histo data
        g1 = fit(bsize',countsub,gaussEqn, options); %fit using the info given above
        maxpeakheight {1,i} = g1.b;
```

```
        averageclusdurcellx {1,i} = 10^(g1.b); %gets real number values from cluster
and stores it
        averageclusdur = 10^(g1.b); %outputs only this peak value for title labling
        bar(bsize,countsub); %bar plot
        hold on
        plot(g1, bsize,countsub); %plots the fit line
        title (['average duration upper cluster ',num2str(i)]);
        ylabel({'Frequency'});
        xlabel(['Log event duration']);
        hold off
        legend off
    end
end
%% All bias analysis
clusterinfoexport = {}; %datapoints for clusters
snclust = {}; %raw datapoints for clusters
saclust = {}; %ana output for clusters
for G = 1:1:size(X,2)
        Figure (F+(4+G))
    numberofclusters = 1;
    C1 =
clusterdata(limdatas{1,G}(:,1),'Linkage','ward','SaveMemory','on','Maxclust',numberof
clusters); %separate by x = 1 by y = 2 in limdatas{1,G}(:,*)
    clusteredvalues = limdatas{1,G}(:,:);
    clusteredvaluesevtraw = eventrawdata {G};
    clusteredvaluesanaout = anadataout {G};
    for i= 1:numberofclusters
        indexpullout = find(C1 == i); %for each cluster create an index value for the row
corresponding
        CIS = clusteredvalues(indexpullout,:); %take all the rows out using indexing
        saforclust = clusteredvaluesanaout(indexpullout,:);%all the rows for ana values
        saclus {1,i} = saforclust;% export the values
        snforclust = eventrawdata{1,G}(:,indexpullout); %get the info via the indexing
        snclus {1,i} = snforclust; %separate evt data into clusters
        clusterinfostorage {1,i} = CIS; %pulls the values for each
        scatter( clusterinfostorage {:,i}(:,1),clusterinfostorage
{:,i}(:,2),4,'filled','DisplayName',(num2str(i))) %scatterplot all initial data
        hold on
    end
    title (['Bias',(num2str(biasorig(G)))]);
    legend ('Location','southeast')
    set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log
scale
    hold off
    clusterinfoexport {1,G} = clusterinfostorage; %cluster intensity and duration export
    snclust {1,G} = snclus; %export raw evt data for use
    saclust {1,G} = saclus;
    Figure (F+(10+G))
```

```
    for i= 1:numberofclusters
        subplot((size(X,2)),1,i) %create subplot
        SZC = size(clusterinfostorage {i},1); %find size of data
        if size(SZC == 1) %if its too small to cluster, we add a second point at 0
            clusterinfostorage {1,i}(2,:) = (zeros); %adding zeros
        end
        tmpdta = log10(clusterinfostorage {1,i}(:,1)); %log the data for histos
        bsize =  (-11:0.05:0); %bin size
        options = fitoptions(gaussEqn); %sets gaussian equation for fit
        options.StartPoint = [1000 -5 0.5 0.5]; %works in sets of 4, this is where the
equation starts a fit
        options.Lower = [100 -6 0 0]; % to not go below this value
        countsub = histc(tmpdta,bsize) ; %get histo data
        g1 = fit(bsize',countsub,gaussEqn, options); %fit using the info given above
        maxpeakheight {G,i} = g1.b;
        averageclusdurcellx {1,i} = 10^(g1.b); %gets real number values from cluster
and stores it
        averageclusdur = 10^(g1.b); %outputs only this peak value for title labling
        bar(bsize,countsub); %bar plot
        hold on
        plot(g1, bsize,countsub); %plots the fit line
        title (['average duration upper cluster ',num2str(i)]);
        ylabel({'Frequency'});
        xlabel(['Log event duration of',(num2str(((biasorig(G)))))]);
        hold off
        legend off
    end
end
%% Compare the values of the clusters
Figure (F+16)
x = biasorig;
lgmaxpeakheight = {};
for G = 1:1:size(X,2)
    for i = 1:numberofclusters
        lg = 10^(maxpeakheight{G,i}(:,:));
        lgmaxpeakheight{G,i} = lg;
    end
    for i = 1:numberofclusters
        labels = {num2str(i)};
        scatter (x(G),lgmaxpeakheight{G,i},'x','DisplayName',(num2str(i)))

text(x(G),lgmaxpeakheight{G,i},labels,'VerticalAlignment','bottom','HorizontalAlignme
nt','right')
        hold on
    end
end
xlabel('Bias (V)')
ylabel('Event duration (s)')
```

```
legend off
hold off
%% cluster select output view
Clustval = input('which clusters do you want to look at?(write values separated by
,)','s');
Clustersofinterest = str2num(Clustval);
Figure (F+17)
for G = 1:1:size(X,2)
    Q = clusterinfoexport{1,G}{1,Clustersofinterest(G)};
    scatter (Q(:,1),Q(:,2),4,'filled','DisplayName',num2str(biasorig(G)))
    hold on
end
legend ('Location','southeast')
set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log scale
hold off
%% Specific cluster comparrison 1/bias
Figure (F+18)
ignorebias = 1;
for G = ignorebias:1:size(X,2)
    scatter
((1/x(G)),lgmaxpeakheight{G,Clustersofinterest(G)},'x','DisplayName',(num2str(x(G)))
)
    hold on
end
legend ('Location','northeast')
G1 = [];
for G = ignorebias:1:size(X,2)
    GQ = cell2mat(lgmaxpeakheight(G,Clustersofinterest(G)))
    G1(G) = GQ;
end
g2 = fit(((1./x(ignorebias:size(X,2))))',G1(ignorebias:size(X,2))','poly1')
plot (g2);%,(x(ignorebias:size(X,2)))',G1(ignorebias:size(X,2)));
ylabel({'Event duration (s)'});
xlabel(['1/bias']);
hold off
%% bias alone fit
Figure (F+19)
for G = ignorebias:1:size(X,2)
    scatter
((x(G)),lgmaxpeakheight{G,Clustersofinterest(G)},200,'x','DisplayName',(num2str(x(G
))))
    hold on
end
legend ('Location','northeast')
ylabel({'Event duration (s)'});
xlabel(['bias (V)',num2str(biasorig(G))]);
G1 = [];
for G = ignorebias:1:size(X,2)
```

```
    GQ = cell2mat(lgmaxpeakheight(G,Clustersofinterest(G)))
    G1(G) = GQ
end
g2 = fit((x(ignorebias:size(X,2)))',G1(ignorebias:size(X,2))','poly3')
plot (g2);%,(x(ignorebias:size(X,2)))',G1(ignorebias:size(X,2)));
hold off
hold off
ylabel({'Event duration (s)'});
xlabel(['Bias (V)']);
%% Intensity values
Figure (F+20)
intensityofclust = {};
for G = ignorebias:1:size(X,2)
    for i= 1:numberofclusters
        avintperclust = mean(clusterinfoexport {1,G}{1,i}(:,2));
        intensityofclust {G,i} = avintperclust;
    end
end
G1 = [];
for G = ignorebias:1:size(X,2)
    GQ = cell2mat(intensityofclust(G,Clustersofinterest(G)))
    G1(G) = GQ
end
for G = ignorebias:1:size(X,2)
    scatter
(x(G),intensityofclust{G,Clustersofinterest(G)},200,'x','DisplayName',(num2str(x(G))))
    hold on
end
g2 = fit(((x(ignorebias:size(X,2))))',G1(ignorebias:size(X,2))','poly3')
plot (g2);
ylabel({'Event intensity (A)'});
xlabel(['Bias (V)']);
legend
hold off
%% Data point finding
if SES == 'y'
    Figure (F+21)
    disp('Click the event you want to look at, then hit "Enter"')
    for G = 1:size(X,2)
        Q = clusterinfoexport{1,G}{1,Clustersofinterest(G)};
        scatter (Q(:,1),Q(:,2),4,'filled','DisplayName',num2str(biasorig(G)))
        hold on
        set(gca,'XMinorTick','on','XScale','log','YMinorTick','on','YScale','log'); %sets log
scale
        datacursormode on    %enable cursor mode
        dcm_obj = datacursormode(Figure (F+21)); %display the scatter with data
cursor
        pause
```

```
        info_struct = getCursorInfo(dcm_obj); %create a structure witht the info in
        points {G}= info_struct.Position; %gets the location
        pt = points{1,G}; %points extraction
        D = find(XandYdata{1,G}== pt); %finds where the data is equal to wher you
asked it to find
        D(2,1) = D(2,1)-length(XandYdata{1,G}(:,1)); %counts the entire column then
continues into the next column, this makes it equal to check it hasnt messed up
        ix = D(1,1);  %column index where the points are found
        evtout {G} = ix; %the event location index
        hold off
     end
    close (Figure(F+21)); %closes the datapoint finding window
    brush off
    datacursormode off

    % so youve found your event.... now lets see what you actually found!
    timesearchvalue = {};
    %first find the data in the ana file, relate it to the event start and stop
    % we need to go back into the ana file and extract column 1 and 2 from the
    % index line
    for G =1:size(X,2)
        data = importdata([PathNameorig,FileNameorig{G}]); %imports the data into a
useable form
        strtdtapoint = data(evtout{G},timecol); %the initial time starting value to find
        enddtapoint =  data(evtout{G},timecol+1);% the end time value to find
        timesearchvalue {G} = {strtdtapoint enddtapoint}; %here are the values we need
to locate
    end
    %%
    % now lets look in the res file and find the file identity
    pathway = convertCharsToStrings(PathNameraw);%set filepath for search
    cd (pathway); %change directory
    for G =1:size(X,2)
          bias = biasorig(G); %know which bias we are searching through
          searchdata = cell2mat(timesearchvalue {1,G});
             if     bias ==-0.4 %sets file names for bias'
                    filelocation =  filelocations{1,1};
             elseif bias ==-0.5
                    filelocation = filelocations{1,2};
             elseif bias ==-0.6
                    filelocation = filelocations{1,3};
             elseif bias ==-0.7
                    filelocation = filelocations{1,4};
             elseif bias ==-0.8
                    filelocation = filelocations{1,5};
             elseif bias ==-0.52
                    filelocation = filelocations{1,6};
             elseif bias ==-0.54
```

```
                filelocation = filelocations{1,7};
            elseif bias ==-0.56
                filelocation = filelocations{1,8};
            elseif bias ==-0.58
                filelocation = filelocations{1,9};
            end
        tmpdtta = {}; %temporary data storage for search
    for j = filelocation(:,1):filelocation (:, end)
        if extrafolder == 'y'
            cd(Pathnamerawmatfiles);
        end
            tmpfname=sprintf('%s_%d%s',basename,j,'.mat.res'); %set up the file
name for data import
            tmpdtta{j} = importdata(tmpfname); %import the data for search
            fndloc = find(tmpdtta {1,j}(:,[1 2]) == searchdata(:,:)); %find in the files
where the values are equal
            if fndloc>=1 %if its done it....
                resfile {G} = sprintf('%s_%d%s',basename,j,'.mat.res');%gives the res file
name
                fileidentout {G} = j; %file identity number
                residx {G} = fndloc; %index in the file for the event
            end
    end
  end
  % Now we've found the file location, lets look at the scan it came from
  Figure (F+22)
if extrafolder == 'y'
  cd(Pathnamerawmatfiles);
elseif extrafolder == 'n'
  cd (PathNameraw) %make sure the directory is correct
end
  for G =1:size(X,2)
        subplot((size(X,2)),1,G)
        %file name import
        FileName3= sprintf('%s_%d%s',basename,fileidentout{1,G},'.mat.cor.mat');
        FileName2= sprintf('%s_%d%s',basename,fileidentout{1,G},'.mat.res');
        FileName1= sprintf('%s_%d%s',basename,fileidentout{1,G},'.mat.evt');
        %file data import
        Qf = importdata(FileName3);
        QFF = importdata(FileName2);
        QFFF = importdata(FileName1);
        H = 1/1000000; %time step
        FQU = (length(Qf.ch2)/1000000); %converts datapoints to time values
        Xvale = (0.000001:H:FQU)'; %sets x axis
        plot (Xvale,Qf.ch2, 'k'); %plots the scan
        xlabel('Time (s)');
        ylabel('Current (A)');
        title(num2str(biasorig(G)))
```

```
            hold on %we want the events overlay on this
            for k = 1:size(QFF,1) %loop needs to be for whole scan
                X1st = QFF(k,1):QFF(k,2); %creates x values for event only plotting
                X1q = X1st/1000000; %converts datapoints to time
                Y1 = QFFF(:,(k+1)); %sets up the intensity values
                Y1 = Y1(Y1~=0); %gets rid of any 0's imposed by rhe event finder
                if length(X1st)==length(Y1) %makes sure the lengths are the same
                    plot (X1q,Y1,'LineWidth',2); %plots the event on top
                end
            end
             hold off
    end
    % Time to plot only events we looked for
    %same as above but only for the event we picked
    Figure (F+23)
        for G =1:size(X,2)
            subplot((size(X,2)),1,G);
            FileName3= sprintf('%s_%d%s',basename,fileidentout{1,G},'.mat.cor.mat');
            FileName2= sprintf('%s_%d%s',basename,fileidentout{1,G},'.mat.res');
            FileName1= sprintf('%s_%d%s',basename,fileidentout{1,G},'.mat.evt');
            Qf = importdata(FileName3); %scan data
            QFF = importdata(FileName2); %res data
            QFFF = importdata(FileName1); %evt data-- want to plot the respective event
            H = 1/1000000;% timestep
            FQU = (length(Qf.ch2)/1000000);% sets max time
            Xvale = (0.000001:H:FQU)'; % sets x axis values
            plot (Xvale,Qf.ch2, 'k'); %plots the scan against real time values
            xlabel('Time (s)'); %
            ylabel('Current (A)'); %
            title(num2str(biasorig(G))) %title - bias
            hold on %keep the scan open
            if residx{1,G}(1,1)> size(QFF,1)
                residx{1,G}(1,1) = size(QFF,1);
            end
            X1st = QFF(residx{1,G}(1,1),1):QFF(residx{1,G}(1,1),2); %Timestart and stop
            X1q = X1st/1000000; % changing this for real seconds time
            Y1 = QFFF(:,(residx{1,G}(1,1)+1)); % getting the evt data out
            Y1 = Y1(Y1~=0);
            plot (X1q,Y1,'LineWidth',2);
        end
end
%% how do the average events look
% extend and average these
%output and plot
%Clustersofinterest = [1,1,1,2,2];

for  G = 1:size(X,2)
    COI = Clustersofinterest(:,G);
```

```
    tmpdta = snclust {1,G}{1,COI};
    [r,c]=size(tmpdta); % determine overall size of input matrix
    Nrs=zeros(r,c); % preallocate results matrix
    for i=1:c % running from the first column to the last of the input matrix
        rc=find(tmpdta(:,i),1,'last'); % finds the last 'non zero' element in each column
        Nrs_tmp=resample(tmpdta(:,i),ceil(r/rc),1); % resamples acc. to scaling factor
r/rc, decimation = 1
        Nrs(:,i)=Nrs_tmp(1:r,:);
    end
    NrsAV = median(Nrs);%(:,(1:end)));  %change median to mean for diferent
analysis
    Nrsaddzero = [0,NrsAV,0];% returns ends to zero
    Figure (F+(24+G))
    plot(Nrsaddzero)
    gca
    title (FileNameorig{G});
    xlabel ('Normalised datapoints')
    ylabel ('Current decrease (A)')
end
%%  sub evt search
%cutoff =0.75e-10;
subeventinfo = {};
warning off
for G =1:size(X,2)
    COI = Clustersofinterest(:,G);
    FileName1 = snclust {1,G}{1,COI};
    [subeventinf, SIGarr] = subeventsearch(FileName1,polarity, cutoff); %find all the
sub event info
    subeventinfo {G} = subeventinf; %save it outside of loop for handling later
end
%% Sub event info
 subeventoverallinfo = {};
for G =1:size(X,2)
   NEIF = size(unique(subeventinfo{1,G}(:,1))); %number of events in file
   [C,IA] = unique(subeventinfo{1,G}(:,1),'first'); %first point in the event
   [C,IB] = unique(subeventinfo{1,G}(:,1),'last'); %last point of data inthe event
   CURSUB = (subeventinfo{1,G}(:,6)); %max current of subevent
   DURSUB = (subeventinfo{1,G}(:,4)); %duration of subevent
   DPMEDIAN = subeventinfo{G}(:,5); %midpoint of each subevent
   startpointsofsubevt =  subeventinfo{1,G}(:,2);
   tstep = 1e-06; %timestep between values
   DNAleng = 4080; %dna length for 4mer aprox
   NOSE = []; %blank variable for number of subevents occuring
   DISEV = {}; %vector for distance(number of datapoints) between events occruing
   AVDISEV = []; %average distance between subevent
   DURAV = []; %average subevent duration
   CURSUBED=[];
   DNAlengtodtapoint = [];
```

```
   basestart = [];
   basesbetween = [];
    for GG = 1:size(C,1)
        NOSE (GG) = (IB(GG) - IA(GG))+1; %number of subevents found per event per
scan
    end
    for GG = 1:size(C,1)
        IAA = IA(GG); %sets start index
        IBB = IB(GG);%sets end index
        CURSUBED (GG) = mean(CURSUB(IAA:IBB)); %AVERAGE MAX CURRENT
of each event
    end
    for GG = 1:size(C,1)
        IAA = IA(GG); %sets start index
        IBB = IB(GG);%sets end index
        DURAV(GG) = mean(DURSUB(IAA:IBB)); %average subevent duration in
seconds
    end
    for GG = 1:size(C)
       IAA = IA(GG); %sets start index
       IBB = IB(GG);%sets end index
       HH = DPMEDIAN(IAA:IBB); %get the values correspoindng to each subevent
       DISEV{GG}  = diff(HH); %distance(number of datapoints) between events
occruing, assumed median point between
    end
    for GG = 1:size(C)
        AVDISEV (GG)= mean(round(DISEV{1,GG}(:))); %average distance between
subevents per event
    end
    %convert for DNA length
    COI = Clustersofinterest(:,G);
    for GG = 1:size(C,1)
%         tmpdta = abs(snclust {1,G}{1,COI}(:,GG)); %convert tempor to positive for
min val
%         [~,ind] = min(tmpdta);%find first 0
        DNAlengtodtapoint (GG) = DNAleng/saclust{1,G}{1,COI}(GG,7); %DNA length
converted from number of datapoints
        IAA = IA(GG);
        startofsubevt = startpointsofsubevt(IAA);
        basestart = startofsubevt*DNAlengtodtapoint(GG);
        basesbetween (GG)= DNAlengtodtapoint(GG)*AVDISEV(GG);
    end
    % cluster distribution positioning
    Subprobdistnorm = [];
    Subprobsingleevt = {};
    for GG = 1:size(C)
       IAA = IA(GG); %sets start index
       IBB = IB(GG);%sets end index
```

```
        tmpdta = abs(snclust {1,G}{1,COI}(:,GG)); %convert tempor to positive for min
val
        [~,ind] = min(tmpdta);%find first 0
        EvtDur = (ind-1)*tstep; %find the event duration
        LocV = (DPMEDIAN(IAA:IBB)).*tstep;% get location value
        SubProbdist = LocV./EvtDur;%multiple by duration of whole event
        Subprobdistnorm = [Subprobdistnorm SubProbdist'];
        Subprobsingleevt {G} = SubProbdist;
    end
    subeventoverallinfo {1,G} = NOSE; %number of subevents per event
    subeventoverallinfo {2,G} = round(nanmean(NOSE)); %average number of
subevents over all events
    subeventoverallinfo {3,G} = CURSUBED; %AVERAGE MAX CURRENT of each
event
    subeventoverallinfo {4,G} = nanmean(CURSUBED);% average overall max current
experienced
    subeventoverallinfo {5,G} = nanmedian(CURSUBED); %median overall max
current experienced
    subeventoverallinfo {6,G} = DURAV; %average subevent duration in seconds
    subeventoverallinfo {7,G} = nanmean(DURAV); %average overall duration
    subeventoverallinfo {8,G} = nanmedian(DURAV); %median event duration
    subeventoverallinfo {9,G} = DISEV{:,:}; %distance between subevent datapoints
    subeventoverallinfo {10,G} = AVDISEV; %average distance between subevent
datapoints
    subeventoverallinfo {11,G} = nanmean(round(AVDISEV));%overall average
distance between datapoints
    subeventoverallinfo {12,G} = basestart; %the starting point for each subevent in
basepairs
    subeventoverallinfo {13,G} = DNAlengtodtapoint;% DNA length calculated from
number of datapoints
    subeventoverallinfo {14,G} = basestart; %the base number that the events start at
    subeventoverallinfo {15,G} = basesbetween; %average number of bases between
each subevent
    subeventoverallinfo {16,G} =  Subprobsingleevt {:,:};% the start points for each
event
    subeventoverallinfo {17,G} = Subprobdistnorm; % the normilised start points for
each event
    subeventoverallinfo {18,G} = log10(Subprobdistnorm); % the log normilised start
points for each event
end
%% Extract the useful columns, because its annoying
subinfobybias = {};
for G =1:size(X,2)
    subinfobybias {1,G}(:,1) = subeventoverallinfo {1,G};
    subinfobybias {1,G}(:,2) = subeventoverallinfo {3,G};
    subinfobybias {1,G}(:,3) = subeventoverallinfo {6,G};
    subinfobybias {1,G}(:,4) = subeventoverallinfo {15,G};
    subinfobybias {2,G}(:,1) = subeventoverallinfo {17,G};
```

```
end
%% Time to look for DNA directionality
FBthresh = 0.2;
Direct = {};
ForwardDNA = {};
BackwardsDNA = {};
for G =1:size(X,2)
   NE = subeventoverallinfo {17,G}; %all normalised events
   [C,IA] = unique(subeventinfo{1,G}(:,1),'first'); %first point in the event
   [C,IB] = unique(subeventinfo{1,G}(:,1),'last'); %last point of data inthe event
   Direc = [];
   forw = [];
   backw = [];
   for GG = 1:size(C,1)
       IAA = IA(GG); %sets start index
       IBB = IB(GG);%sets end index
       if NE(IAA)>FBthresh
          Direction = 1; %sets indec for direction
          backwa = NE(IAA:IBB); %forward results are then encorporated into matrix
          backw = [backw backwa];
       else% NE(IAA)<FBthresh %if its less
          Direction = 0;
          forwa = NE(IAA:IBB);
          forw = [forw forwa];
       end
       Direc  = [Direc Direction];
   end
    Direct {G} = Direc;
    ForwardDNA {G} = forw;
    BackwardsDNA {G}= backw;
end
%take to origin for data analysis here
%% Save the files
save('datafoircomparrison25pA.mat','subinfobybias');
%% stats analysis, do manually for different groups based on distribution
G = 4;
addon = 1;
    x1 = clusterinfoexport{1,G}{1,Clustersofinterest(G)}(:,2);
    y1 = clusterinfoexport{1,G+addon}{1,Clustersofinterest(G+addon)}(:,2);
    [h,p,ci,stats] = ttest2(x1,y1)
%% sub event finder -- this is modified for clustered data input
function [subeventinf, SIGarr] = subeventsearch(FileName1,polarity, cutoff)
    FilesEvt=strcat(FileName1);
    FilesEvt=cellstr(FilesEvt);
    j=size(FilesEvt);
    m3=0; m4=1;
    % Determine duration of each event in the evt file and calculate median
    % [20/80] (based on the central 60% of each event)
```

```
    for m=1:j(2)
        RESdata=FileName1;
        REScur=RESdata(:,1:end)*polarity; % remove first column (time) and project as
positive currents (to make seach easier)
        [r,c]=size(REScur);
        RESlast=zeros(1,c);
        RESmed=zeros(1,c);
        for m2=1:c
            RESlast(:,m2)=find(REScur(:,m2)~=0,1,'last'); % find last non-zero data point
            if RESlast(:,m2)<5 % makes sure that too short events cause issues
                RESmed(:,m2)=median(REScur(1:RESlast(:,m2),m2));
            else % here only the event from 20 to 80% in time is considered

RESmed(:,m2)=median(REScur(ceil(0.2*RESlast(:,m2)):floor(0.8*RESlast(:,m2)),m2)
);
            end
        end
    end
    r1s=1;
    tstep=1e-06;
    for evtc=1:c
     N=REScur(:,evtc);
        r1=RESlast(:,evtc); % number of data points in file (Q: Should this include zeros
at the end or not?)
        r1=r1-r1s;
        mu=RESmed(1,evtc);
        SIGstart=1;
        m=1;
        m2=1;
        while m <= r1
            m=m+1; %new
            if N(m,1)<=abs(mu)
                SIGstart=m;
                SIGend=1;
            elseif N(m,1)>=cutoff
                    SIGend=find((N(SIGstart+1:end,1))<=abs(mu),1,'first');
                    TF=isempty(SIGend);
                    if TF==1
                        SIGend=1;
                    end
                    SIG(m2,1,evtc)=evtc; % This adds the event column to each event, to
facilitate further processing
                    SIG(m2,2,evtc)=SIGstart; % data pt. number, event start
                    SIG(m2,3,evtc)=SIGend+SIGstart; % data pt. number, event end
                    SIG(m2,4,evtc)=(SIG(m2,3,evtc)-SIG(m2,2,evtc))*tstep; % duration in
seconds
                    [~,mxi]=max(N(SIGstart:(SIGend+SIGstart),1)); % determines index of
maximum within subevent, relative to SIGstart
```

```
                SIG(m2,5,evtc)=SIGstart+mxi(1); % write index of maximum relative to
event index (mxi(1) catches error when multiple values found)
                SIG(m2,6,evtc)=max(N(SIGstart:(SIGend+SIGstart),1))-
(N(SIGstart,1)+N(SIGend+SIGstart,1))/2; % max current within sub-event, rel to
approximate lin. baseline
                SIG(m2,7,evtc)=trapz(N(SIGstart:(SIGend+SIGstart),1)-
(N(SIGstart,1)+N(SIGend+SIGstart,1))/2); % charge [A*unit]; multiply with tstep to
charge [C}
                m=SIG(m2,3,evtc)+1;
                SIGstart=m+1;
                m2=m2+1;
            end
        end
    end
    % Process result matrix
    evtc=1;
    SIGarr=zeros(1,7);
    [~,~,f3]=size(SIG);
    for evtc=1:f3
        tSIG=SIG(:,:,evtc);
        evtc;
        d=find(tSIG(:,1)>0,1,'last');
        tSIG=tSIG(1:d,1:7);
        %tSIG=reshape(tSIG,[],6);
        SIGarr=[SIGarr;tSIG];
    end
    SIGarr=SIGarr(2:end,:);
    subeventinf = SIGarr; %export the info for later use
end
```

Appendix 6 – DNA simulation

```
%%synthetic data
close all;clear all;clc;
x = 1:1:10000000; %sets comparable time size to actual data
%% create ui for synthesis
scrsz = get(0,'ScreenSize');
S.fh=Figure(1);
set(S.fh,'toolbar','Figure');
set(S.fh,'Name','Initialisation parameters','Position',[10 (scrsz(4)/2-100) (scrsz(3)/3-25)
(scrsz(4)/3)+50]);
S.tx(1)=uicontrol('Style','Text','String','Save data? ','Position', [20 50 100 25]);
S.cb(1)=uicontrol('Style','checkbox','Position', [180 50 25 25],...
   'KeyPressFcn','A');
S.tx(2)=uicontrol('Style','Text','String','Filename: ','Position', [20 80 100 25]);
S.cb(2)=uicontrol('Style','edit','String','SynthDNA','Position', [180 80 300 25]);
S.tx(3)=uicontrol('Style','Text','String','DNA length [bp]: ','Position', [20 240 100 25]);
S.cb(3)=uicontrol('Style','edit','String','10000','Position', [180 240 80 25]);
S.tx(4)=uicontrol('Style','Text','String','Bias voltage [V]: ','Position', [20 300 100 25]);
S.cb(4)=uicontrol('Style','edit','String','0.8','Position', [180 300 80 25]);
S.tx(5)=uicontrol('Style','Text','String','Filter frequency? [kHz]: ','Position', [20 185 100
25]);
S.cb(5)=uicontrol('Style','edit','String','10','Position', [180 185 80 25]);
S.tx(6)=uicontrol('Style','Text','String','Number of structures wanted?: ','Position', [20
120 100 25]);
S.cb(6)=uicontrol('Style','edit','String','10','Position', [180 120 80 25]);
S.tx(7)=uicontrol('Style','Text','String','Type of DNA added? [R,B,SUB,SB]: ','Position',
[5 300 150 25]);
S.cb(7)=uicontrol('Style','edit','String','B','Position', [180 300 80 25]);
QQQQ = input('values ready to continue?','s');
%% set initial values from input
bias=str2double(get(S.cb(4),'string'));
typeofstrut = (get(S.cb(7),'string'));
ff=str2double(get(S.cb(5),'string'))*1e3;
numberstruts=str2double(get(S.cb(6),'string'))*1e6;
DNAlegngth =str2double(get(S.cb(4),'string'))*1e6;
A=get(S.cb(1),'Value'); % Save data?
QQQQ = input('values ready to continue?','s');
%% set up values for change
sevtdec= (-160E-12); %normal event
doublebent = (-160E-12)*1.5; %bent event
strut = (-200E-12); % with unbound strut
boundstrut = (-250E-12); %with bound strut
%% noise set up
nois = 45E-16; % create artificial noise baseline
nois1 = (nois*ff);
noise = (nois1/(1/bias));
Ya =zeros((length(x):1));
```

```
for i = 1:2:length(x)-1
    Ya(i) = (-1*rand)*noise; %positive values
    Ya(i+1) =(1*rand)*noise; %negative values
end
corrnoise = (Ya); %multiplies values based on filter frequency used
noisy = awgn(corrnoise,10, 'measured'); % adds gaussian noise
maxtime = 20; %normal max time for event, assumed to be 20ms but feel free to
change
%% Deciding factors

if typeofstrut == 'R';
    noi = sevtdec;
elseif typeofstrut == 'B';
    noi = doublebent;
elseif typeofstrut == 'SUB';
    noi = strut;
elseif typeofstrut == 'SB';
    noi = boundstrut;
end
%% assume translocation frequency based on other data
TF = (1); %translocation frequency
TT= (1); %translocation time
ED = (1); %event decrease average
%% Running data show
DC=Figure(2);
lHandle = line(nan, nan); %# Generate a blank line and return the line handle
set(DC,'Name','Synthscan','Position',[scrsz(3)/3+10   (scrsz(4)/2-100)   (scrsz(3)/3-25)
scrsz(4)/3+50]);
%%
 noise = 10E-12;
Ya =zeros((length(x):1));
for i = 1:2:length(x)-1
    Ya(i) = (-1*rand)*noise; %positive values
    Ya(i+1) =(1*rand)*noise; %negative values
end
corrnoise = (Ya*(ff/10000)); %multiplies values based on filter frequency used
noisy = awgn(corrnoise,10, 'measured'); % adds gaussian noise
Figure (1) %create Figure
plot (x,corrnoise);
ylim([-2e-10 2e-10]);
Figure (2)
plot (x,noisy);
ylim([-2e-10 2e-10]);
hold off
%%
evtxns=randi([15,maxtime],1,1); %create random event duration
xleng = (1:evtxns); %create length of event
ev1 = [0,repelem(evtdec,(evtxns-2)),0]; %create y values for event duration
```

```
ydet = awgn(ev1,20, 'measured'); %add artificial noise
% Figure (3)
evtxnsbnt=randi([10,(maxtime*0.75)],1,1);  % create random event duration but
reduced time due to bending
xleng2 = (1:evtxnsbnt);
ev2 = [0,repelem(doublebent,(evtxnsbnt-2)),0];
yevt2det = awgn(ev2,10, 'measured');
evtxs= randi([15,maxtime],1,1);
xleng3 = (1:evtxs);
mtrx = zeros(length(xleng3):1);
for i = 2:((length(xleng3)/numberstruts)):(length(xleng3)-1)
    mtrx (round(i)) = evtdec;
    mtrx (round(i)+1) = strut;
end
mtrx ((length(xleng3))) = 0;
yevt3det = awgn(mtrx,20, 'measured');
evtxbs= randi([15,maxtime],1,1);
xleng4 = (1:evtxbs);
mtrx1 = zeros(length(xleng4):1);
for i = 2:((length(xleng4)/numberstruts)):(length(xleng4)-1)
    mtrx1 (round(i)) = evtdec;
    mtrx1 (round(i)+1) = boundstrut;
end
mtrx1 ((length(xleng4))) = 0;
%ev4 = [0,repelem(boundstrut,(evtxbs-2)),0];
yevt4det = awgn(mtrx1,20, 'measured');
Figure (3)
plot (xleng4,yevt4det);
Figure (4)
plot (xleng4,mtrx1);
%bound and unbound struts
%% intergrate the event with the original data
tmpta = (length(x)/5);
tmptb = length(x);
eventdata = noisy;
tmptpoint = [];
Figure (2)
for i =1:5
    tmeptint = randi([tmpta,tmptb],1,1);
    tmptpoint = [tmptpoint tmeptint];
end

evdtared = eventdata;
for i = 1:length(tmptpoint);
    evdtared(tmptpoint(i):1:(tmptpoint(i)+((length(ev1))-1)))= ydet;
    evdtared = evdtared;
end
plot(x,evdtared);
```