UNIVERSITY OF BIRMINGHAM

MULTIFUNCTIONAL NANOPIPETTES: FROM DNA CROWDING TO NANOSCALE SENSING

by

RAND AHED MAHMOUD AL-WAQFI

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

April 2025

PUBLISHED MATERIAL

Both Chapters 3 and 4 of this thesis were adapted from the following publication, where I am the first author:

Rand A. Al-Waqfi, Cengiz J. Khan, Oliver J. Irving, Lauren Matthews, and Tim Albrecht. "**Crowding Effects during DNA Translocation in Nanopipettes**", ***ACS Nano***, **2025**, *19*(17), 16803–16812.

DOI: 10.1021/acsnano.5c01529

ABSTRACT

This work is concerned with single-molecular detection of DNA, RNA and other target systems. First, this thesis showcases the utilisation of nanopipettes' asymmetric geometry to capture and retain double stranded kilobase pair DNA, then subsequently release the DNA by reverse translocation into the bulk solution. By utilising different DNA lengths through multiple experiments, the accumulation of DNA inside the nanochannel has been shown to induce a crowding effect. Further, under certain circumstances no DNA translocation observed, and the pipette was reversibly blocked. This crowding effect is further correlated with the length of the translocated DNA molecules; shorter DNA occupies less space and requires more translocation events to fill the pipette compared to longer DNA molecules. Interestingly, with the introduction of functional moiety to the DNA, this thesis suggests that this newfound effect could be utilized to trap DNA as a capture probe and incubate it with certain analytes inside the nanopipettes, thus, analyte detection by 'functionalised' DNA can be further assessed after reverse translocation.

Beyond local trapping, carrier DNA was also explored in this thesis for long noncoding RNA (lncRNA) detection as a distinct aim. While the results in this specific aspect remained inconclusive, overall, the work highlights the importance of asymmetric geometry of nanopipettes toward opening the door to new capabilities for biophysical and bioanalytical research at the single-molecule level.

ACKNOWLEDGEMENT

Finally, it is the end of this journey after 3.5 years of hard work and dedication. But this work would not be done without people around me whom I am profoundly grateful to each of you. Thank You!

First, I want to express my deepest and most heartfelt gratitude to my supervisor, Tim Albrecht. Your guidance, insight, and patience have enabled me to overcome my fears and trust my knowledge in a complex field called nanopores. I still remember your reaction and support once I sent you the 7 kbp data about DNA trapping.

People in the Albrecht group thank you for everything, thank you for your help always. Nashwa thank you for help and advice always, Jake thank you for help in the lab. Cengiz and Oliver thanks always for fixing the P2000 and making the nanopore lab alive. I would also like to thank Linda from the Friscic group as being a very nice and supportive friend. All of you I wish you the best.

This work is dedicated to my husband, my kids, my family and my husband's family, without your support I would not get into here:

Thank you, Khaled. You carried the weight of this journey with me, held me through moments of doubt, and celebrated even the smallest victories. Your belief in me never wavered, and your love gave me courage. I couldn't have done this without you.

My great mother Abeer, you are my source of strength and power. Thank for taking care of my kids and supporting me always. Gaith and Bisher, my little ones, you are my motivation.

Finally, I am very grateful to the university I graduated from Jordan University of Science and Technology, whose funding made this research possible.

To my family , for their endless love and

support

Thank you

LIST OF PUBLICATIONS

1. **Al-Waqfi, R. A.**, Khan, C., Irving, O. J., Matthews, L., & Albrecht, T. (2025). *Crowding Effects during DNA Translocation in Nanopipettes*. ***ACS Nano***, 19(17), 16803–16812. Also available as arXiv preprint: arXiv:2501.14347 (2025).

2. **Al-Waqfi, R. A.,** Irving, O. J., & Albrecht, T. (2024)**.** Label-free long non-coding RNA assay on a string for bladder cancer detection. ***Biophysical Journal****, 123*(3), 439a.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ΔI_e, $\Delta I_{effective}$	Effective current (relative)
ΔI_{Max}	Current maxima or maximum current (relative)
Δq	Event charge
μ	Electrophoretic mobility
μ_i	Electrophoretic mobility (conventional)
A	Ampere
AC	Alternating current
ADC	Analogue to digital conversion
AFM	Atomic force microscopy
AuNP	gold nanoparticles
bp	Base pair
$C_{DNA, in}$	DNA concentration inside the nanopipette
$C_{DNA, out}$	DNA concentration in the bulk solution
Cf	Cell-free
cf.	See
CSgG	Curli Secretion Protein G
D	Diffusion coefficient
d, d_p	Pore ‘s diameter
DB	Dielectric breakdown
DC	Direct current
DNA	Deoxyribonucleic acid
dsDNA	Double Stranded DNA
EDL	Electric double layer
EDTA	Ethylenediamine tetraacetic acid
ES	Electrochemical-based sensor
f	Mean translocation frequency
G_{pore}	Pore conductance
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IVT	*In-vitro* transcription
K_B	Boltzmann constant

kbp	Kilobase pair
kHz	Kilohertz
l	Length of nanochannel or chip membrane thickness
L	Taper length
L_C	Contour length
lncRNA	Long noncoding RNA
L_p	Persistence length
miRNA	Micro RNA
MspA	Mycobacterium smegmatis porin A
N	Number of monomers
N_{evt}	Total number of detected events
NFW	Nuclease free water
nt	Nucleotide(s)
PCR	Polymerase chain reaction
Phi29	Bacteriophage phi29
qPCR	Quantitative PCR
r	Ion current rectification ratio
r_c	Capture radius
RE	Restriction Enzyme
R_H	Hydrodynamic radius
RMS	root mean squared
RNA	Ribonucleic acid
R_{pore}	Pore resistance
rRNA	Ribosomal RNA
RT	Reverse transcriptase
SA-AuNP	Streptavidin functionalised gold nanoparticles
SPR	Surface plasmon resonance
ssDNA	Single stranded DNA
T	Kelvin temperature
TAE	Tris-Acetate-EDTA buffer
TBE	Tris-Borate-EDTA
TE	Tris-EDTA buffer

tRNA	Transfer RNA
UCA1,UCA-1	Urothelial carcinoma associated 1
UV-Vis	Ultraviolet-visible spectroscopy
$\mathbf{v}$	Velocity of translocation
V	Volt
$\mathbf{V_{bias}}$	Applied voltage
$\mathbf{V_{ch}}$	Volume of nanochannel
$\mathbf{V_{DNA}}$	Volume of DNA in solution
$\boldsymbol{\lambda}$	Debye length
λ DNA	48.5 kbp DNA
$\boldsymbol{\sigma_f}$	Surface charge density
$\boldsymbol{\sigma_s}$	Solution conductivity
τ	Translocation duration or dwell time
I_0, I_b	Open pore current

CHAPTER 1
INTRODUCTION

CONTENTS

Synopsis: This chapter introduces the fundamental principles and theoretical background of nanopore sensing, with a specific focus on DNA translocation through solid-state nanopores, particularly nanopipettes.

1.1. Single Molecule Detection

Traditional biosensors generate a signal based on the interaction of a sufficient number of molecules with the sensor probe. The resulting signal is an ensemble-average quantity. Accordingly, ensemble-based measurements offer several advantages including measurement reproducibility, user friendly experimental operation and high throughput. These advantages have made ensemble-based measurement ideal for large-scale application in diagnostics such as mass screening for infectious diseases, high-throughput drug testing, and population-wide biomarker analysis.[1-3] Figure 1.1A illustrates a sensor device setup where the sensor probe is bound to a certain number of analyte's molecules, while Figure 1.1B shows a typical response of an ensemble averaging-based sensor after analyte binding. Hence, when we have a certain number of molecules in the sample the response will be averaged as a one signal. Further, the intensity of the signal produced will be proportional to the number of molecules bound to the sensor probe. Thereby, averaged measurements are more robust, with less noise interference than individual single molecule measurements. Despite all these advantages, the resulting signal is an ensemble-averaged quantity, which limits the sensor detection threshold.[2] Further, ensemble measurements averages the signal based on macroscopic properties of the analyte ignoring the specific properties of each entity within the system.[2] For instance, if the sensor probes were partially damaged, this is difficult to detect from the overall ensemble-averaged signal. Thus, a lower signal response may be misclassified as a false negative. This could be problematic in the field of diagnostics, as a lower response resulting in missed disease detection. On the other hand, a single molecule sensor response is presented in Figure 1.1C where each analyte molecule interacting with the sensor produces a distinct signal. These signals vary based on analyte's shape, charge and

conformation. Moreover, single molecule methods enable the detection of heterogeneous states of a single molecules.[2, 3] Conversely, single molecule methods has some drawbacks, including complex experimental setup, increased noise and complexity of the signals produced impacting their wide applicability for routine use. Interestingly, single molecule methods provide a plethora of information at the molecular level in terms of microscopic states.[2, 3] In the light with this, single molecule methods were previously applied in studying the structure and the function of important biomolecules like DNA, RNA and proteins.[4, 5]This led to better understanding of the underlying diseases mechanisms associated with a particular mutation or interaction with other biomolecules.[5]

Importantly, single molecule sensors exhibit high sensitivity for low abundance biomarkers. An interesting research work in single entity detection has revealed an attomole-level detection of DNA, proteins and micro-RNA associated with several disease.[6, 7] This is a significant advancement in terms of disease diagnostics, enabling the earlier detection of low abundance biomarkers.

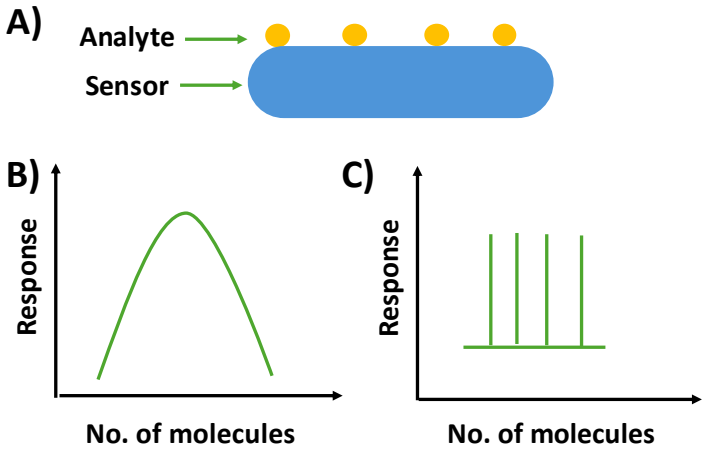

Figure 2.1: provides a (A) depiction of a sensor probe bound to the analyte, (B) representation of typical response of an ensemble averaging based sensor and (C) representation of signals retrieved by single molecule method.[3]

1.2. DNA as a Single Molecule

Deoxyribonucleic acid (DNA) is a special molecule with a unique double-helix structure that encodes genetic information within an organism. Understanding DNA structure is crucial for applications involved drug delivery, diseases diagnostics, studying genetic mutations, and gaining deeper insight into biological functions.[8] Basically, linear double stranded DNA structure contains two antiparallel polynucleotide chains stabilised through complementary base pairing, base stacking interactions, hydrogen bonding, and hydrophobic effects. Each polynucleotide chain is composed of a sugar phosphate backbone and four types of nitrogen bases attached to the deoxyribose sugar, including adenine(A), thymine(T), guanine(G) and cytosine(C). DNA bases forms two distinct base pairs types (A-T and C-G) through hydrogen bonding, which base pair to connect both opposite DNA strands.[9] Nitrogen bases have purine and pyrimidine rings within their structure that are stacking together in a sequence specific manner twisting double stranded DNA. Regardless of base-pairing, base-stacking is important for DNA structure, which contributes significantly to the stability and the shape of DNA as a macromolecule (helical structure). Due to the presence of phosphate groups, DNA is considered as highly charged polyelectrolyte with an elementary charge of $2e^-$ per 3.4 Å (single base pair). Hence, DNA has a high charge density affecting its overall structure in terms of rigidity and bending.[10, 11] Particularly, the length of DNA molecule in its extended B-form form is estimated by multiplying the number of base pairs (*N*) by the helical rise per base pair, which is approximately 0.34 nm, this length is defined as DNA contour length ($\boldsymbol{L_c}$), given by equation (1.1)[12]:

$$L_c = 0.34 * N \tag{2.1}$$

DNA macromolecular rigidity, bending and stability are all governed by DNA length, salt concentration and environmental conditions.[13, 14] Thus, the longer DNA molecule, the more repulsion along its structure will occur, this in turn will allow the DNA molecule to bend, increasing flexibility. A term displays the elasticity/stiffness of DNA as a polymer is referred to as DNA persistence length (L_p). Persistence length values indicate that a polymer at certain length would be a stiff rod-like structure, while, beyond this length, the polymer tend to bend more deviating from stiff rod-like structure. At physiological conditions (0.1 M NaCl), the persistence length of double-stranded DNA is approximately ~ 50 nm (~ 150 bp).[13] It is shorter at higher ionic strengths, due to the screening of DNA negative charge.[14]

Nanopore electrical readout is a novel single molecule technique that has been developed to study DNA structures and other biomolecules with high throughput.[15] In pursuit of this, next sections in this chapter will provide brief introduction to the history, principles and types of nanopores, with focus on solid state nanopores.

1.3. The Genesis of Nanopore Sensing

The concept of nanopore sensing was first implemented by Wallace Coulter, when he filed a patent under the name 'Coulter principle 'in 1953.[16] Nowadays, the Coulter counter is a standard device in medical laboratories for blood count testing. In simple terms, the device works by the principle of resistive pulse sensing.[17, 18] Figure 1.2 presents an illustration diagram of a resistive pulse sensor.

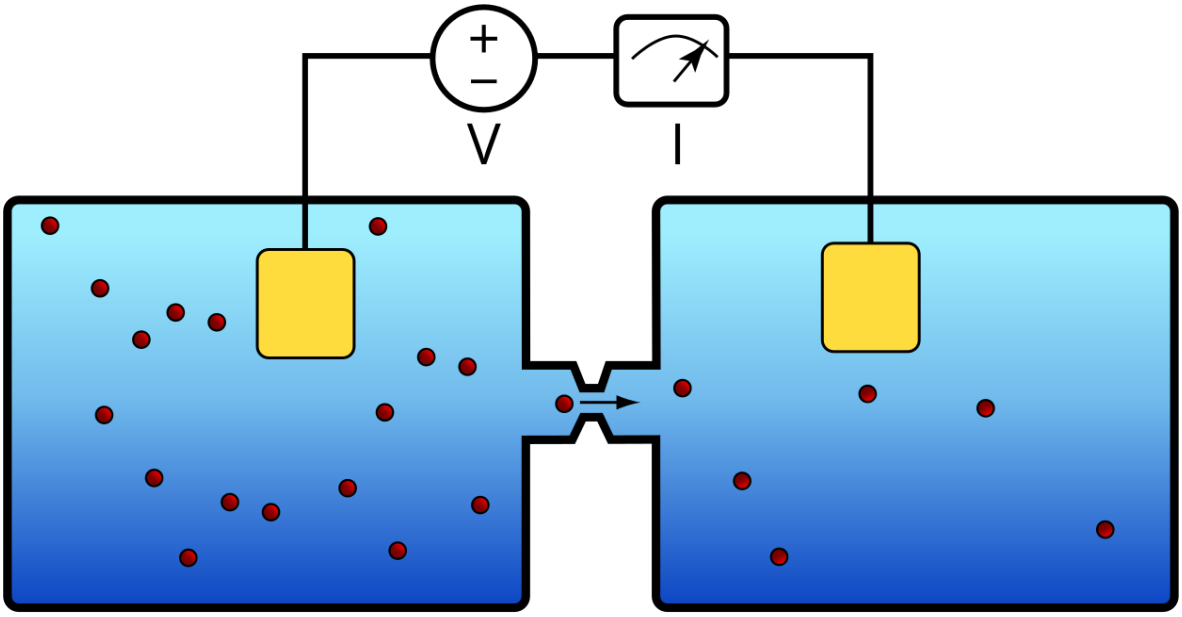

Figure 2.2: Schematic presentation of a resistive pulse sensor, charged analytes (red) are placed in one chamber. Once, the electrodes are biased (yellow), an electric current will be formed moving the analyte through the pore. Accordingly, the size of the analyte is detected as change in resistance. Image by *Mymathematix(2018)*, licensed as: CC-BY-SA 4.0, retrieved from (https://upload.wikimedia.org/wikipedia/commons/f/f1/NC_diagram_version_1.svg)

Specifically, a micro-aperture with a diameter ranges approximately from 0.4 µm to 800 µm is formed between two separated chambers. Each chamber contains a conductive solution and one electrode. Charged particles are placed in only one of the chambers. The two electrodes are connected to the main circuit. Once the electrodes are biased by an applied voltage (V_{bias}), an electrical current is formed inducing the movement of charged analytes between the two chambers. As soon as the analyte passes through the aperture, a detectable change in pore's resistance proportional to the size of the analyte is recorded. Consequently, the size and the amount of the particles in the solution can be quantified.[17, 18]

Having this idea in mind, single particle capture by resistive pulse sensing was further advanced into a miniaturised version to capture single molecules or even ions.[19] This idea came into light after the 3D crystal (Figure 1.3) determination of the pore forming

protein(α–hemolysin). The crystal of α–hemolysin structure has revealed a pore diameter of 4.6 nm narrowing down to 1.4 nm toward its end.[20] Thus, in 1996 Kasianowicz et al. was the first who utilised these transmembrane proteins for sensing ssDNA and RNA, this utilisation ushered in a new era of single molecule detection.[21] These biological pores were small enough to accommodate only polynucleotide chains suggesting their potential use in DNA sequencing. This approach was later commercialised by Oxford Nanopores Technologies in 2015 (MinION©).[22] Aside from biological nanopores, artificial membranes were shortly suggested in 2001 and 2003 as more durable substitute.[23, 24]

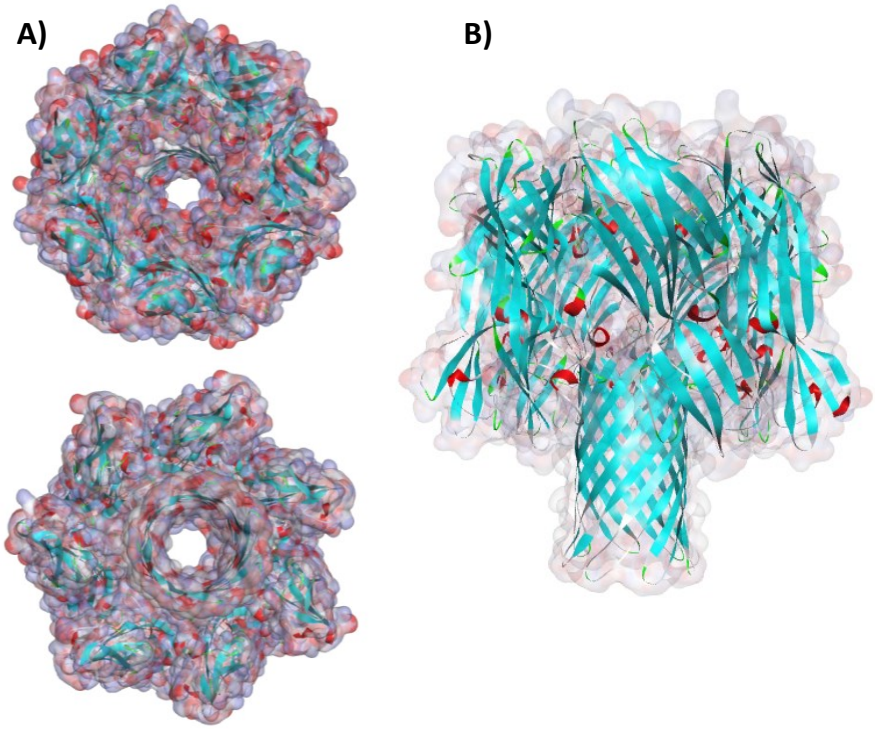

Figure 2.3 The crystal structure of α–hemolysin transmembrane protein (PDB code 7AHL) (A) cross sectional view from the upper constriction(top) and lower constriction (bottom) (B) longitudinal view of the protein channel.[20]

1.4. Different Classes of Nanopores

Types of nanopores can be classified into three classes based on the material they are fabricated from. Therefore, nanopores can be biological, synthetic (solid state) nanopores, or a hybrid of both.[15, 19]

1) **Biological nanopores:** as discussed earlier biological nanopores are pores formed naturally in transmembrane proteins. Prominent example is α-hemolysin, a wild type protein-based nanopore. This transmembrane protein produced by Staphylococcus aureus and Escherichia coli bacteria. α-hemolysin is commonly used in DNA sequencing.[25, 26] Another intriguing protein-based pore is MspA (Mycobacterial porins), which is well recognized for its shorter sensing region with (0.6 nm in length) and a pore diameter of 1.2 nm. This structure provides superior spatial resolution for ssDNA sequencing.[27] The former pores are not suitable for studying dsDNA molecules, which have a diameter of approximately 2 nm. The bacteriophage Phi29 nanopore has a length of 7.2 nm length and a width ranging from 6 nm to 3.6 nm, making it capable of accommodating dsDNA.[28] Aerolysin has a longer channel with very narrow opening of approximately of 0.9 nm. Aerolysin channels had been suggested for protein sequencing and small molecules sensing.[28] Oxford Nanopore Technologies now employs genetically engineered protein channels optimized for lower error rates in sequencing results (CSgG pore).[22, 29] In terms of pore size reproducibility and low noise properties, biological nanopores are robust. However, they suffer from several drawbacks. First, the limited pore size range (< 7 nm), restricting their application only for smaller analytes size. Additionally, the complex fabrication and the requirement of lipid membrane impacts their stability and durability. [19, 30-32].

2) **Solid state nanopores** offer a variable range of pore sizes, as well as higher mechanical stability of the artificial membranes used for fabricating the pores. Hence, solid state nanopores can be applied for studying complex analyte systems beyond sequencing. Solid state nanopores can be categorized into two types: [33, 34]

A) **Chip-based nanopores (figure 1.4A)** which are usually formed in thin membranes (thickness < 20 nm). Membranes for chip-based nanopores are usually made from silicon nitride (Si_3N_4) or silicon dioxide (SiO_2). Further, other materials were also employed like HfO_2, MoS_2 and TiO_2.[35-37] Pores are created by exposing a silicon nitride thin film to a beam of electrons so that a nano-hole can be drilled precisely. Such a strategy gives much more control during the fabrication process. However, In terms of cost effectiveness, the production of nanopores using this method requires an expensive instrumentation setup which makes them less accessible.[19, 35-37] Dielectric breakdown (DB) can be a cheaper strategy to form nanopores within silicon nitride membranes. Briefly, the introduction of a trans-membrane potential generates an electric stress inside the insulating membrane. This stress causes defects in random shape in silicon nitride membrane. Subsequently, the defects connect to form a physical path of the nanometre size through electroporation.[38] Although DB can be cost effective in terms of pore production, however it is hard to upscale (multiple pores production in the same membrane).[38]

B) **Nanopipettes (figure 1.4B)** are either quartz or glass capillary based nanopores. Usually, nanopores are fabricated by laser assisted pulling. Although they sometimes lack reproducibility due to variation of pore sizes within the same batch, Simple fabrication procedure of nanopipettes can outweigh this problem. Additionally, quartz based nanopipettes possess low capacitance, having lower noise in comparison with chip based

nanopores, not to mention wettability and tailored surface properties.[33, 39] Nanopipettes have an interesting conical geometry with a tapered shape. The length of ion channel is longer which is narrowing down to the tip, introducing asymmetry in comparison with chip based nanopores.[40]

3) **Hybrid nanopores:** another important class of nanopores that merges the low noise properties of biological nanopores and the mechanical stability of solid state nanopores are hybrid nanopores. These nanopores are typically designed by integrating a biological component (biological pore like α-hemolysin, MspA or functionalised DNA origami) with solid-state materials (silicon nitride, graphene, or glass). However, this combination usually complicates the fabrication process, reducing their scalability and applicability.[41-46]

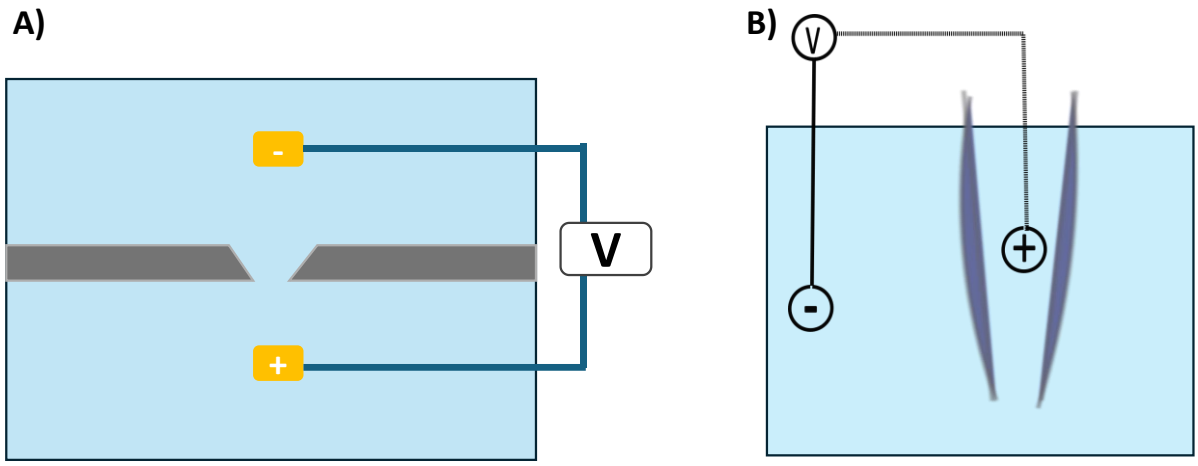

Figure 2.4 Depiction of a (A) chip based nanopore device setup and (B) nanopipette device setup

1.5. The operating principle of nanopore sensor

After the introduction of pore sensing and types of nanopores, it is important to understand the sensing principle of a nanopore. Similar to Coulter counter devices, charged analytes are driven by an applied electric field to pass through the pores, resulting in a change in the pore

resistance. However, when scaling down to the nanometre regime, the nanopore's surface charge effect becomes greater.[47, 48] In a nanopore sensing device setup, first a liquid cell is filled with an electrolyte solution, for example potassium chloride (KCl), which is commonly known as a standard solution used in nanopore experiments. In Figure 1.4A, two Ag/AgCl electrodes are placed in both sides of the liquid cell, having the nanopore membrane separating each side.[48, 49] Hence, it is important to make sure that only the nanopore is the only connection between both sides of the liquid cell setup. Like chip based nanopore, a nanopipette device setup is shown in Figure 1.4B, where one of the electrodes is placed inside the bulk solution (cis) while the other one inside the nanopipettes (trans). Silver electrodes are preferred, as electrodes of this type are non-polarisable providing stable current measurement. Additionally, the electrode reaction is compatible with a chloride-containing electrolyte. In terms of connection to the electronics, one electrode is connected to a power supply (for biasing), while the other one (inside the nanopipettes) connected to the electronics (amplifier) for measurement. Once a voltage difference is applied (V_{bias}) across Ag/AgCl electrodes, a redox electrochemical reaction is formed at the electrodes' surfaces:

Oxidation at the anode: $Ag_{(s)} + Cl^- \rightarrow AgCl_{(s)} + e^-$ (1.2)

Reduction at the cathode: $AgCl_{(s)} + e^- \rightarrow Ag_{(s)} + Cl^-$ (1.3)

The oxidation reaction in equation 1.2 at the anode (positively charged) involves the formation of silver chloride (AgCl) by attracting Cl^- ions toward the anode (reacting with Ag). Two concurrent processes as result of this reaction will happen, first a potential difference is applied between Ag/AgCl electrodes, causing a depletion in the concentration of chloride ions in one side. Further, this potential difference will induce the movement of ions from the other

side through the nanopore as compensation mechanism, resulting in an ionic current (I_0) referred as baseline or background current. Since nanopore is the major resistance of the system, I_0 is considered as the open pore current. Figure 1.5 presents an illustration of the direction of ion movement (panel A) and the formed background current (panel B). [15, 39, 48, 49]

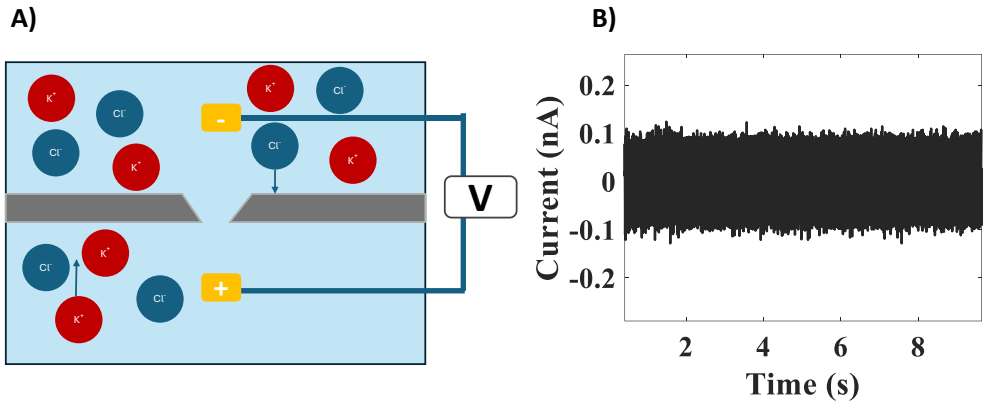

Figure 2.5 (A) depiction of ion movement induced after potential difference causing a (B) baseline current as detected by the electronics (The pore current appears centred at zero because the AC channel is zero-mean, please refer to Chapters 2 and 3 for further details).

When a voltage is applied (V_{bias}), an electrical current is driven through the system. Hence, circuit components will resist current flow by means of total resistance ($\boldsymbol{R_{Total}}$) of the system following Ohms law:[50]

$$\boldsymbol{R_{Total}} = \frac{V}{I} \tag{1.4}$$

Thus, it is important to understand factors affecting ion current flow within the system. Accordingly, we have three main sources of resistance: charge transfer resistance at the electrode/electrolyte, resistance of the bulk solution and resistance of the pore. As indicated earlier, Ag/AgCl electrodes are nonpolarizable electrodes that exhibit minimal resistance. Hence, the resistance can be denoted as in equation (1.5): [39, 48]

$$\boldsymbol{R_{Total}} = R_{pore} \tag{1.5}$$

In nanopipette, pore's geometry displays high surface-to-volume ratio, the resistance emerges from bulk solution is often neglected (at high salt concentration).Accordingly, pore's resistance is dominant, since it is the main site of the voltage drop ($V_{bias} \approx \Phi_{pore}$).[33, 48, 51]

It is more intuitive to describe the system by pore conductance ($\mathbf{G_{pore}}$) rather than resistance. Hence, considering pore geometry as a cylinder, $\mathbf{G_{pore}}$ can be calculated using equation (1.6): [48-50]

$$\boldsymbol{G_{pore}} = (\boldsymbol{R_{Total}})^{-1} = \sigma_s \frac{\pi\, d^2}{4l} \tag{1.6}$$

Where σ_s is solution conductivity, d is pore diameter and l is the thickness of membrane (nanochannel length in nanopipettes)

Accounting for accurate assumption of pore resistance needs careful consideration of all pore surroundings. This would aid in understanding the mechanism of ion transport through nanopores. To offer more insights, nanopores' surfaces are often negatively charged such as quartz nanopipettes and Si_3N_4. When a V_{bias} is applied, counter ions are attracted to the charged surface, generating an electric double layer (EDL) at the interface between the charged solid surface and the electrolyte solution. EDL Thickness (λ) or "Debye length" is dependent on ionic strength of the solution. Based on surface charge effect, an extra term would be added to the pore conductance calculation including surface charge ($\boldsymbol{G_{surf}}$), thus total pore conductance can be estimated using equation (1-7)[39, 48-50]:

$$\boldsymbol{G_{pore}} = \sigma_s \frac{\pi\, d^2}{4l} + \frac{\pi d}{l} \sigma_{\mathrm{f}} \cdot \mu_i \tag{1.7}$$

(σ_{f} and μ_i are surface charge density and electrophoretic mobility, respectively.)

It is important to mention, that at sufficiently high salt concentration (> 0.1 M) and relatively larger pore size($d>>\lambda$), surface charge density is constant with bulk ions controlling ion transport through the pore ($V_{bias} \alpha I$).[33, 34, 48, 52, 53]

Therefore, in this thesis all DNA translocation experiments were conducted using high salt concentration.

1.5.1. DNA Translocation

Now imagine having the nanopipettes as an example with a charged biomolecule such as DNA injected to the bulk solution side (cis). Thus, the application of an external negative V_{bias}, will set the electrode in the bulk solution to a negative potential (The other way around when a positive bias is applied). Since the negatively charged DNA is suspended in the bulk solution (not inside the pipettes), the DNA will move in response to the electrophoretic force from the bulk near the tip to enter the nanopipettes. This process is known as ***DNA translocation***. At high salt concentration with equal salt concentration in both compartments.[34, 52, 53] When a DNA molecule translocates through the pore a relatively partial blockage of ionic current flow will happen.[40, 51-53] This blockage appears to be a transient drop in current signals as presented in Figure 1.6B, along with a description of the setup loaded with DNA molecules (figure 1.6A).

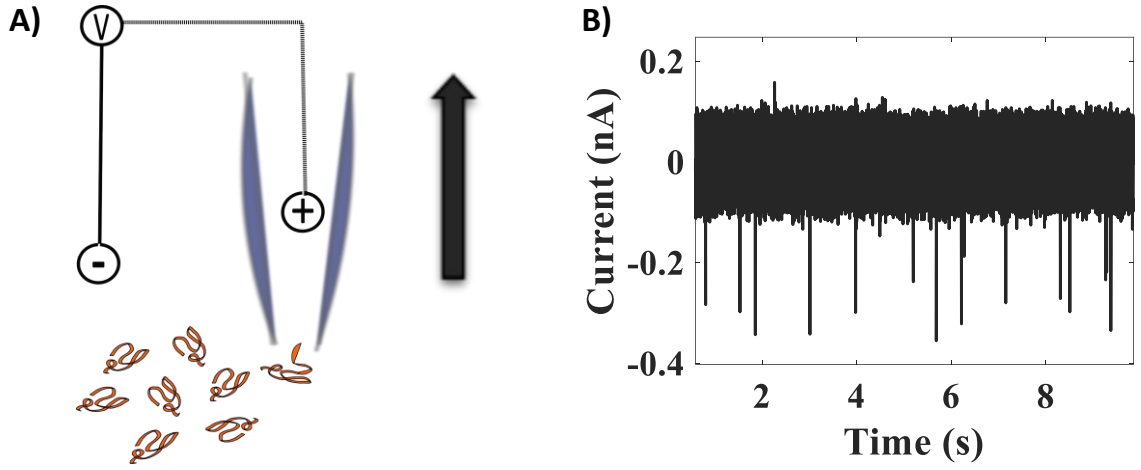

Figure 2.6 (A) Depiction displaying nanopipette translocation setup with DNA molecules in the bulk solution (B) Representative current time trace with transient spikes, indicating blockages of ion current flow.

The former description of DNA translocation was overly simplified. Nevertheless, translocation is composed of three consecutive stages: diffusion, drift and threading. In solution, movement of DNA molecules will be random and dominated by Brownian motion.[39, 54, 55] After the application of an external V_{bias}, the electric field will be the strongest at the pore opening. However, after a defined distance, again, Brownian motion becomes dominant making DNA molecules' movement less controlled by electrophoresis. This distance is known as the capture radius ($\boldsymbol{r_c}$). The capture radius is obtained by equation (1.8):[39, 54]

$$r_C = \frac{d^2 \mu}{8lD} \cdot V_{bias} \tag{1.8}$$

(Where D is the diffusion coefficient and μ is the electrophoretic mobility of DNA)

It should be noted that the capture radius of nanopipettes is hemispherical from the outside. however, it is confined and smaller from the inside. This would in turn, affect the translocation characteristics when DNA molecules are driven from the cis compartment to the trans compartment compared to the opposite direction, adding direction as a factor affecting

translocation characteristics.[56-59] As the DNA molecule enters the capture (hemispherical) volume, the transport mechanism shifts from diffusion to a drift mechanism, and DNA molecule is mainly transported by electrophoresis. [39, 54, 55] From equation (1.8), one would assume that r_C scales positively with both V_{bias} and pore size. While r_C is inversely proportional to diffusion coefficient D. Hence, longer DNA molecules (L_c>>L_p), tend to diffuse slower than smaller molecules.[60] This means that longer DNA molecules have an increased chance to enter the capture volume compared to the shorter DNA molecules. This can be further emphasized by understanding the relationship between DNA effective dimension and D. From D ,we can obtain DNA molecule effective dimensions within the solution as well as occupying volume by estimating DNA hydrodynamic radius (R_H), using Stokes-Einstein equation (equation 1.9):[39, 60]

$$R_H = \frac{k_B T}{6\pi\eta D} \tag{1.9}$$

Where k_B is Boltzmann constant, T is kelvin temperature and η is solution viscosity.

Further the hydrodynamic volume (V_H) of DNA molecule is calculated by equation (1.10):[39]

$$V_H = \frac{4}{3}\pi R_H^3 \tag{1.10}$$

Importantly, the frequency at which the DNA molecules enter the capture volume is often described by mean translocation frequency (f) as determined by equation (1.11):[54]

$$f = \frac{\pi\, d^2 \mu}{4l} \cdot V_{bias} \cdot C_{DNA,out} \tag{1.11}$$

Where $C_{DNA,out}$ is the concentration of the DNA in the bulk solution.

It is evident from equation (1.11) that the mean translocation frequency (f) depends on various parameters such as applied voltage, nanopore dimensions and DNA concentration in the bulk. At higher voltages, the translocation frequency is expected to increase linearly ($f \propto V_{bias}$) which indicates a drift dominated transport of DNA molecules rather than diffusion (assuming other factors controlled).[51, 61, 62] However, any deviation from this aspect (drift regime) indicates subtle factors that need to be investigated.

When the DNA molecule enters the capture radius; it will be subject to electrophoretic force. As mentioned, the electric field is the strongest at the pore opening, and the DNA molecule begins attempting to enter the pore. [51, 55] Particularly, the entry process requires overcoming energy barriers.[51, 55] Since DNA adopts multiple conformations in the solution, DNA molecule must align properly and overcome entropic barrier to initiate the pore entry process (threading). At higher voltage regimes, the effect of energy barrier is insignificant were DNA is mainly driven by electrophoresis.[51] The way the DNA enter the pore has been previously studied, as DNA molecule prefers to enter from one of its ends rather than between those ends, disregarding partially unravelled states.[63-65]

From the electrophoretic force we can estimate the duration at which DNA fully transversing through the pore. This duration is often referred as dwell time (τ) or translocation duration. Thus, at high voltage regime, the velocity (v)of DNA translocation can be obtained from equation (1.11), where $v = \frac{\mu V_{bias}}{l}$. Accordingly, in case of kilobase-pair DNA where $L_c \gg l$, $\tau = \frac{\boldsymbol{l.Lc}}{\mu V}$.[66, 67]

In a drift dominated regime, it is imperative to understand that once DNA molecule enter the pore, DNA translocation inside the confinement is driven by electrophoresis.[66, 67] This can be

interpretated as an inverse relationship where $\tau \propto 1/V_{bias}$.[63, 67-69] However, there is a possibility of increasing τ with the opposing electroosmotic force (EOF), since the direction of EOF usually follows the direction of the counter ions in the EDL for the negatively charged surfaces. Additionally, DNA molecule upon threading may experience a hydrodynamic drag due to the viscosity of the solution and the surrounding solvent molecules.[39]

1.5.2. DNA Electrical Signatures Analysis

As indicated earlier in figure 1.6B at high salt concentrations, when a DNA molecule translocates through the pore, a partial blockage of ionic current flow is observed as a transient drop relative to the baseline current. Each current drop is directly proportional to molecule's characteristics such as its size, charge or configuration. Hence, accurate assignment of DNA translocation signal from the background noise is crucial to obtain reliable data. At least, two characteristic features can be extracted from each spike. The first feature is event amplitude (ΔI_{Max}).Hence, ΔI_{Max} is usually calculated as the maximum decrease or increase in ionic current (I_{Max}) relative to the baseline current (I_b) as defined by equation (1.12): [58, 70-72]

$$\Delta I_{Max} = I_b - I_{Max} \tag{1.12}$$

The second element is spike's width or duration (τ). [58, 70-72] Figure 1.7 shows an actual current time trace of a DNA translocation event with arrows indicating both ΔI_{Max} and τ.

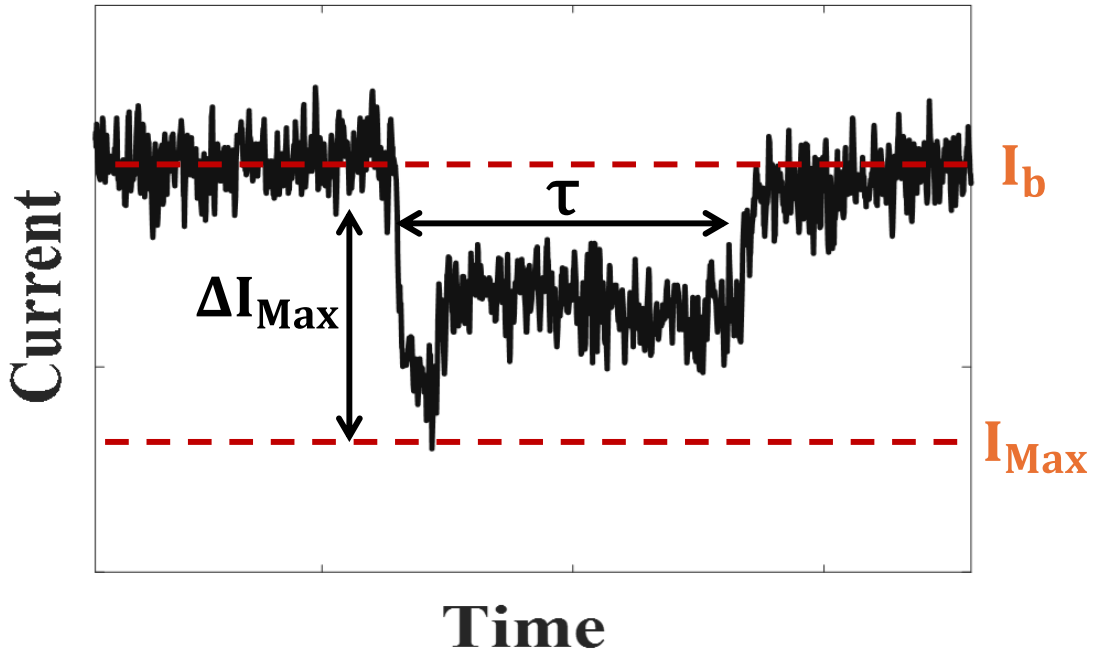

Figure 2.7 Current time trace of an actual DNA translocation event along with its amplitude (ΔI_{Max}) and duration (τ).

Current maxima (ΔI_{Max}) may not fully present the overall picture of translocated DNA molecules, as ΔI_{Max} captures only the peak blockade value at a single point in time. Accordingly, ΔI_{Max} is prone to interference due to transient fluctuations, accompanied by the stochastic nature of translocation process and the electric noise of the system. since τ is defined by event stop and start, noise will also affect its accurate determination.[70, 73] In this respect, the event charge (Δq) can be calculated by integrating the current blockade over the event duration, given by equation (1.13): [70, 73]

$$\Delta q = \int_{t0}^{t1} \Delta I\ (t)dt \qquad (1.13)$$

Where $t0$ and $t1$ are the starting and the end data time point of an event, respectively.

Δq presents the amount of charge displaced during a DNA translocation event. Interestingly, Δq has been found to be constant for a specific DNA length regardless of DNA conformation.[73]

By dividing event charge over duration ($\Delta q/\tau$) we can obtain an effective current (ΔI_e). Opposed to ΔI_{Max}, ΔI_e represent overall current magnitude. Accordingly, in case of determination of DNA conformation, ΔI_e is more representative as it does reflect the overall current reduction observed during a DNA translocation event. However, for identifying a modification site or a binding event (protein bound to the DNA) within single DNA molecules, ΔI_{Max} is more representative in this case. [39, 70, 73]

After the introduction of the characteristic features of a single DNA event, understanding DNA translocation process requires recording a considerable number of translocation events (10^2-10^3 events).[39] Previously, dsDNA translocation process and events shape has been studied and several hypotheses of how the DNA enters the pore were formulated.[65, 74, 75] When analyzing complex molecules such as DNA using relatively larger pores (>8nm), there is an increased challenge in measuring and interpreting the results due to the ability of the molecule to fold and knot upon itself. [74, 76] The chance of folding or knotting increases as the length of the DNA strand grows.[74, 76] Therefore, capturing a statistically meaningful number of events is crucial for accurate post-analysis of the signal.[77, 78] As a result, overall event features can be further studied by using clustering and statistical analysis to determine the underlying trends, variations, and representative configurations of DNA events.

For typical analysis of dsDNA events conformation, events may represent two subpopulations' ones with shorter duration (τ) and higher current blockade ΔI_e, these events are usually referred to "folded" and the ones with less intensity and longer durations are called "linear" events (figure 1.8A). Folded events tend to produce current response twice the intensity of a linear event (figure 1.8B). Specifically, for understanding DNA structure, accurate analysis and comparison of DNA molecules must be conducted properly. This is particularly important

when studying DNA lengths or defining a location of a modification within the same molecule. Hence, (unfolded) linear events are more representative. As, it is often complex to identify the point of DNA entry during a folded event.[33, 39, 65, 75, 79, 80]

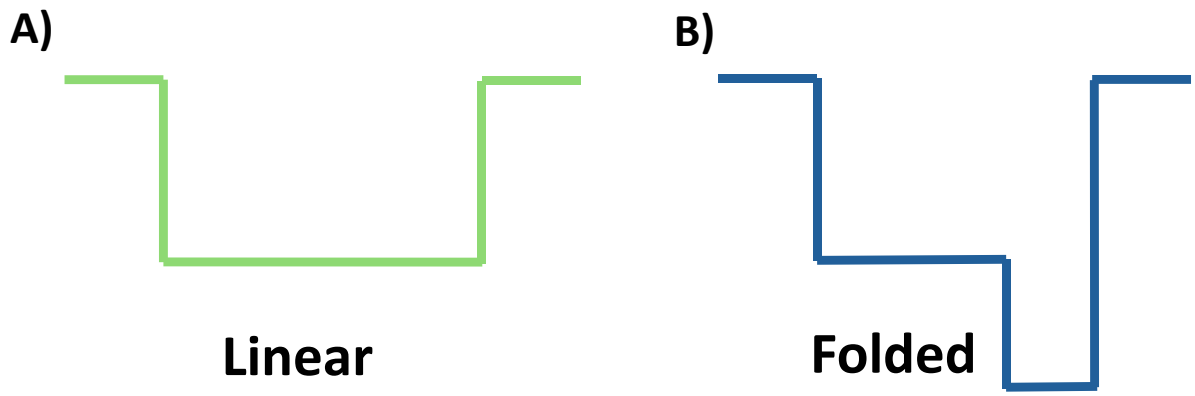

Figure 2.8 presentation of the shape of (A) linear translocation event and (B) folded translocation event.

1.6. Analyte Capture Strategies in Solid State Nanopores

Since the initial introduction of solid state nanopores, the primary purpose was to sequence biopolymers such as DNA. However, DNA translocation through solid state nanopores was fast, impacting proper temporal resolution for single base identification. To achieve this aim, initial efforts focused on slowing down DNA translocation and enhancing DNA capture rate. Concurrently, this is still being performed to achieve analyte characterization at high resolution. To slow down DNA translocation, earlier strategies focused on tailoring pore geometry to match the analyte's dimension, optimizing electronics, or modifying solution conditions.[15, 36, 61, 81] Optimized solution conditions included increasing viscosity, reducing temperature and modifying salt concentration.[81] Apart from sequencing, solid state nanopores shifted toward detecting charge analytes like proteins, particles and nucleic acids, however, a reduction in analyte capture rate was encountered. Since analyte capture is

controlled by diffusion, simple approaches to enhance capture rate include increasing DNA concentration and applied voltage (equation 1.11). Increasing DNA concentration raises the risk of pore clogging and crowding, while higher voltage induces DNA deformation along with clogging.[82-84]

Several novel strategies were applied to enhance capture rate. One of the strategies employed salt gradient between both compartments surrounding the pore. For instance, using KCl concentration in the cis side 20x times higher than the trans side has been demonstrated to increase capture rate by approximately a factor of 40. The increased translocation frequency using salt gradient was attributed to electric field effect expansion, accompanied by increased capture radius.[54] Another approach included the application of functionalized lipid-based surface coatings. Further, the lipid coatings around nanopore membrane were functionalized with a specific binding probe to capture the analyte of interest.[39, 83, 85] Other strategies included confining the analyte into the nanopipette through controlled delivery[59], analyte preconcentration via dielectrophoresis[39, 86], and analyte capture by thermophoresis induced through focused plasmonic excitation.[39, 87]

1.7. Carrier-Based Analyte Capture

Capturing analytes present at low concentrations and of small size can be achieved using larger pores (> 10 nm). Benefiting from the fact that dsDNA (>2kb) is linear molecules with permanent charge, supported by the well-studied system in solid state nanopores, the design of carrier DNA modified with capturing probes at a known position had been explored for detecting proteins, shorter DNA and RNA fragments.[88, 89]

As discussed in section 1.6, when a linear DNA structure translocates through the pore, a constant drop in current level correlates with the length of that DNA molecule.[40] However, when a DNA carrier modified with capturing probes along its double structure translocate through the pore, additional electrical component will contribute to the electrical readout of the used nanopore. Figure 1.9 presents the translocation characteristics of both linear unmodified DNA and DNA modified with capturing probes bound to the analyte (forming protein complex as shown in Figure 1.9A (bottom)), additional spikes ("subevents") appear corresponding to the position of that modification within the carrier (figure 1.9B, (bottom)).[39, 79, 89, 90]

This idea had already been demonstrated as novel potential to promote the biosensing application of nanopores in clinical settings.[39, 91-94] The DNA probes were designed using DNA assembly, hybridization or chemically modifying DNA molecule with protein recognition elements. [57, 79, 80, 95-99] The Keyser group has pioneered this concept by designing a library of DNA carriers. Each carrier was modified with analyte specific capture probe in addition to a sequence barcode to precisely distinguish that carrier. They designed their carrier through DNA assembly inspired from DNA origami. To this end, M13 circular single-stranded DNA is linearised by restriction digestion enzymes (RE). To form a dsDNA carrier, complementary oligonucleotides with protrusions or secondary structures are hybridised to the linearised ssDNA carrier through controlled process of heating and gradual cooling. Modified oligonucleotides in their design are either modified with capture probes to the analyte of interest or modified with a special secondary structure "dumbbell".[79] The dumbbell structure can be distinguished by nanopore readout. To generate a set of barcoded carrier, the dumbbell structures' location and number is shuffled to ensure precise detection.[89] Recently,

they have explored the potential of their method to detect micro-RNA and long-noncoding RNA.[97, 100-102] They have further proposed an enhanced repeated capture strategy for analytes present in low concentration.[103] Importantly, our group previously has demonstrated the potential of carrier-based method to capture short DNA fragments for tuberculosis diagnostics.[72]

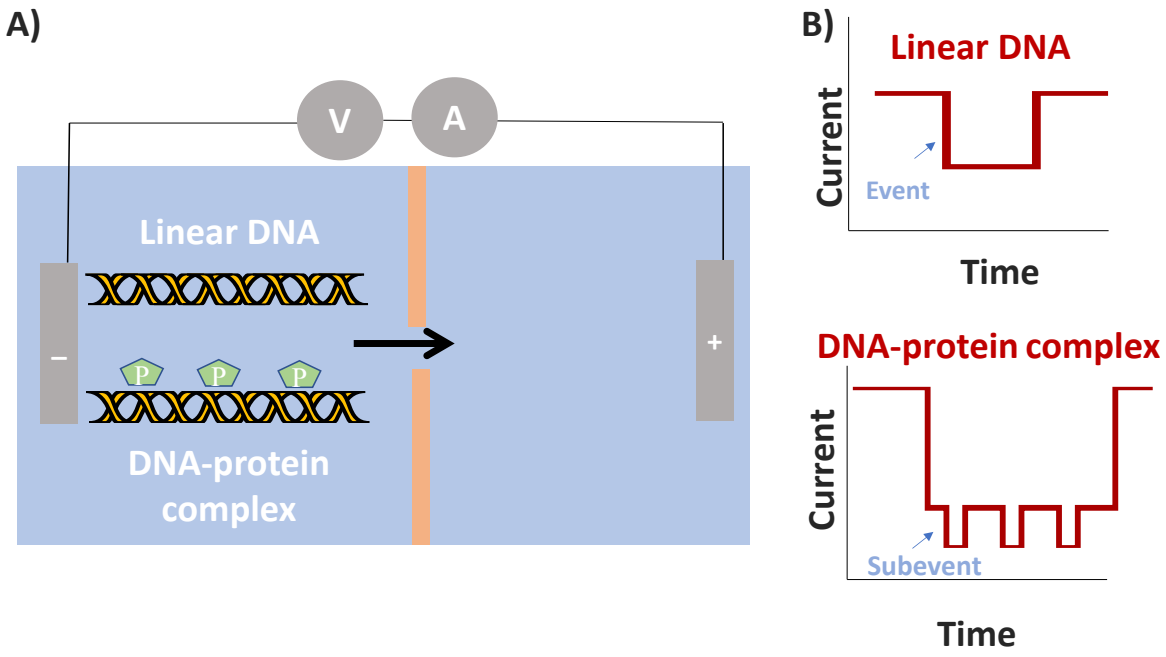

Figure 2.9: illustration of carrier-based DNA strategy. (**A**) shows nanopore setup has unmodified DNA (**top**) and modified DNA-bound to protein (**bottom**) both pass through a nanopore (**B**) shows the blockage of ionic current signal when linear DNA passes through the pore (**top**) modified DNA structure contains modification sites and bound to the analyte electrical signatures as subevents (small, squared protrusions).

Evolving from this point, the above mentioned strategies have constituted the basics to further promote the application of solid state nanopores in capturing smaller analytes like proteins, short DNA fragments, RNA and nanoparticles.[35, 36, 39, 94, 104]

1.8. Nanopipette Geometry

The geometry of the pore has an important impact on the translocation characteristics. As we discussed in section 1.5, nanopipette has a conical shape with a taper. This tapered shape creates a strong electrophoretic force gradient which aids in directing DNA molecules into the capture radius to enter the pore. The sensing region or nanochannel is the narrowest part located at the end of pipette tip. Pore inner diameters typically range from 5 nm to 20 nm, and gradually widen to the (inner) diameter of the capillary.[105] This convergence created at opening angles usually smaller than 10°, with taper lengths extending to several millimetres.[72] Within the first 50 nm from the tip, the sensing region remains confined to the narrowest part of the channel. As a result, when DNA passes through pore (nanochannel), the speed of translocation slows down, achieving higher temporal resolution. The asymmetry between the cis (wider opening) and trans (narrower tip) sides significantly influences DNA transport dynamics. This geometric difference creates a spatial restriction on DNA molecules once it exits the pore. Thus, asymmetry has been found to affect the duration of DNA translocation and confirmation when DNA molecules move from the inside to the outside of the pore (the trans compartment).[56, 57, 103, 106] However, less work has been implemented to understand the effect of geometry on DNA capture rate and translocation dynamics on both direction.

For instance, does the asymmetric geometry of the pipettes make it possible for the pipette to trap DNA molecules inside the sensing region. Figure 1.10 illustrates how asymmetry might affect DNA translocation. Furthermore, in Figure 1.10A we can treat the pipette as a storge device that can accommodate a certain number of DNA molecules. Accordingly, this device might take a relatively longer time to fill the narrower tip with DNA molecules (loading).

After the translocation of sufficient number of DNA molecules as shown in Figure 1.10B, assuming that these molecules were trapped inside, do they reach a crowded state where they affect the translocation characteristics of the succeeding translocating molecules from the wider region.

Accordingly, asymmetric trapping of DNA molecules inside would also provide valuable insights into the interactions between molecules within a confined space, potentially advancing our understanding of molecular structures under confinement and opening the door toward real time reaction monitoring inside a nanoconfined space. If the asymmetric pipette were able to trap DNA, can it be used as an incubation tool for small analytes such as proteins to facilitate their capture in unconventional ways. Interestingly, the asymmetry of the pipette can present an interesting opportunity for novel bioassays.

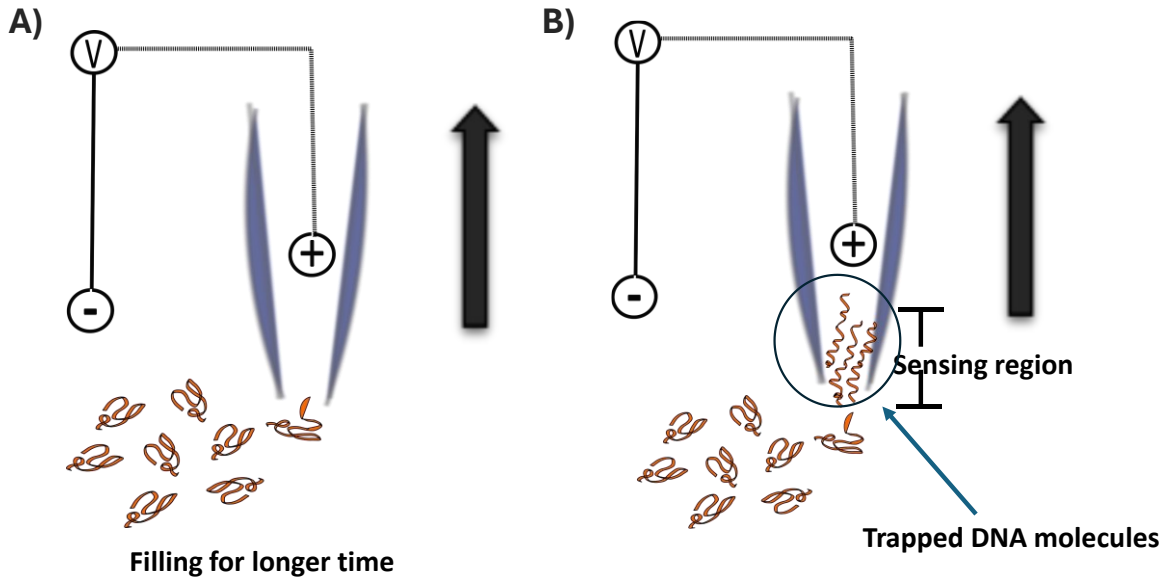

Figure 2.10 illustration of the idea of treating nanopipette as storage device where first (A) DNA to test whether nanopipette (B) can trap DNA after longer time of the translocation experiment. The black arrow signifies the direction of DNA movements (cis to trans).

1.9. Thesis Objectives and Outlook

The main goal of this thesis was to explore the effect of the asymmetric geometry of nanopipettes on DNA capture rate and translocation dynamics. We detailed in the introduction the theoretical framework behind DNA translocation in solid state nanopores and the factors involved in controlling the process. Since nanopipettes exhibit an asymmetric geometry, there is limited research on how nanopipette asymmetry affects DNA translocation characteristics. Further, previous work explored the effect only on translocation duration and current intensity, ignoring that asymmetry might constrain diffusion and induce accumulation. From this point, the motivation of this thesis was to initiate a systematic analysis in this regard. Accordingly:

Chapter 3 details a systematic analysis of DNA translocation of a range of DNA lengths (4 kbp to 48.5 kbp) within the nanopipettes in both directions (cis and trans). We have observed an increase in the local concentration of DNA inside the nanopipette near the tip due local trapping. Thus, the objectives aimed to analyse the effect of DNA accumulation on translocation frequency and translocation duration in both compartments. Furthermore, the study considered different aspects including the DNA length, baseline current fluctuation, pore conductance and nanopipette's effective volume. [107, 108]

Chapter 4 After demonstrating that DNA can remain trapped inside the nanochannel affecting DNA translocation dynamics, Chapter 4 aimed to implement a proof-of-concept study. Accordingly, nanopipettes were employed as a local storage device for 40 nm streptavidin functionalised Au-nanoparticles. For this purpose, biotinylated 10 kbp DNA was designed through PCR amplification. Capture of these nanoparticles was achieved by first loading the pipette with modified DNA (in the cis compartment) and allowed it to incubate with particles

inside the pipette (trans). Then, successful capture was demonstrated by reverse translocation and subevent analysis. [107, 108]

Both **Chapter 3** and **Chapter 4** resulted in successful publication of the results:

***Al-Waqfi, R. A**.; Khan, C.; Irving, O. J.; Matthews, L.; Albrecht, T**. Crowding Effects during DNA Translocation in Nanopipettes**. ACS Nano* **2025**, *19* (17), 16803–16812. The work was also made available as a preprint: arXiv:2501.14347 (2025).

Chapter 5: beyond the primary investigation, a secondary objective was to assess the feasibility of nanopipette-based analysis of long, noncoding RNA (lncRNA) without fragmentation. While distinct from the main focus, this study aimed to explore whether nanopipettes could serve as a potential platform for RNA sensing. Thus, the approach involved the combination of both carrier DNA and nanopore readout to assess if lncRNA could be detected. The translocation characteristics of the carrier, free RNA and carrier-RNA hybrid were analysed by nanopipettes. Finally, electrical signals of each were compared; wherein, the challenges and the potential of this strategy were further discussed. However, the results in this respect were not conclusive.

Chapter 6: this chapter presented the overall conclusions from the findings of this thesis, in terms of understanding DNA translocation in asymmetric pores, and suggests future work.

1.10. References

1. S. K. Vashist, *Biosensors*, 2017, **7**, 62.
2. E. Barkai, Y. Jung and R. Silbey, *Annu. Rev. Phys. Chem.*, 2004, **55**, 457-507.
3. J. J. Gooding and K. Gaus, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 11354-11366.

4. Y. L. Liu, T. Y. Bian, Y. Liu, Z. M. Li, Y. F. Pei and J. Song, *Engineering-Prc*, 2023, **24**, 276-291.
5. Q. P. Su and L. A. Ju, *Biophys Rev*, 2018, **10**, 1349-1357.
6. N. Akkilic, S. Geschwindner and F. Hook, *Biosens. Bioelectron.*, 2020, **151**, 111944.
7. N. Chauhan, K. Saxena and U. Jain, *Int. J. Biol. Macromol.*, 2022, **209**, 1389-1401.
8. C. R. Calladine, H. Drew, B. Luisi and A. Travers, *Understanding DNA: The Molecule and How it Works*, Elsevier, 2004.
9. A. Travers and G. Muskhelishvili, *FEBS J*, 2015, **282**, 2279-2295.
10. M. Frank-Kamenetskiĭ, V. Anshelevich and A. V. Lukashin, *Soviet Physics Uspekhi*, 1987, **30**, 317.
11. G. S. Manning, *Q Rev Biophys*, 1978, **11**, 179-246.
12. M. Doi, S. F. Edwards and S. F. Edwards, *The theory of polymer dynamics*, oxford university press, 1988.
13. N. C. Stellwagen, Y. J. Lu and B. Weers, *Biophys. J.*, 2002, **82**, 127a-127a.
14. C. G. Baumann, S. B. Smith, V. A. Bloomfield and C. Bustamante, *Proc Natl Acad Sci U S A*, 1997, **94**, 6185-6190.
15. Y. Zhou, X. Long, Y. Zhang, D. Zheng, Y. Jiang and Y. Hu, *Langmuir*, 2025, **41**, 5736-5761.
16. J. P. Robinson, *Cytometry A*, 2013, **83**, 424-438.
17. H. R. Walter and C. H. Wallace, *US Patent 3,557,352*, 1971.
18. S. M. Bezrukov, *J. Membr. Biol.*, 2000, **174**, 1-13.
19. W. Shi, A. K. Friedman and L. A. Baker, *Anal. Chem.*, 2017, **89**, 157-188.
20. L. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley and J. E. Gouaux, *Science*, 1996, **274**, 1859-1866.
21. J. J. Kasianowicz, E. Brandin, D. Branton and D. W. Deamer, *Proc Natl Acad Sci U S A*, 1996, **93**, 13770-13773.
22. H. Lu, F. Giordano and Z. Ning, *Genomics Proteomics Bioinformatics*, 2016, **14**, 265-279.
23. J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz and J. A. Golovchenko, *Nature*, 2001, **412**, 166-169.
24. A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen and C. Dekker, *Nat Mater*, 2003, **2**, 537-540.
25. M. Wanunu, *Phys Life Rev*, 2012, **9**, 125-158.
26. M. M. Mohammadi and O. Bavi, *Biophys Rev*, 2022, **14**, 99-110.
27. A. H. Laszlo, I. M. Derrington and J. H. Gundlach, *Methods*, 2016, **105**, 75-89.

28. P. Chen, Z. Sun, J. Wang, X. Liu, Y. Bai, J. Chen, A. Liu, F. Qiao, Y. Chen, C. Yuan, J. Sha, J. Zhang, L. Q. Xu and J. Li, *Front Microbiol*, 2023, **14**, 1043967.

29. A. Dorey and S. Howorka, *Nat Chem*, 2024, **16**, 314-334.

30. L. Luo, S. R. German, W. J. Lan, D. A. Holden, T. L. Mega and H. S. White, *Annu Rev Anal Chem (Palo Alto Calif)*, 2014, **7**, 513-535.

31. D. Deamer, M. Akeson and D. Branton, *Nat. Biotechnol.*, 2016, **34**, 518-524.

32. M. N. Islam, M. K. Masud, M. H. Haque, M. S. Al Hossain, Y. Yamauchi, N. T. Nguyen and M. J. A. Shiddiky, *Small Methods*, 2017, **1**, 1700131.

33. L. J. Steinbock, O. Otto, C. Chimerel, J. Gornall and U. F. Keyser, *Nano Lett.*, 2010, **10**, 2493-2497.

34. L. J. Steinbock, A. Lucas, O. Otto and U. F. Keyser, *Electrophoresis*, 2012, **33**, 3480-3487.

35. C. Wen and S.-L. Zhang, *J. Phys. D: Appl. Phys.*, 2020, **54**, 023001.

36. Y. Luo, L. Wu, J. Tu and Z. Lu, *Int J Mol Sci*, 2020, **21**, 2808.

37. Q. Chen and Z. Liu, *Sensors (Basel)*, 2019, **19**, 1886.

38. H. Kwok, K. Briggs and V. Tabard-Cossa, *PLoS One*, 2014, **9**, e92880.

39. T. Albrecht, *Annu Rev Anal Chem (Palo Alto Calif)*, 2019, **12**, 371-387.

40. J. B. Edel and T. Albrecht, *Engineered nanopores for bioanalytical applications*, William Andrew, 2013.

41. P. Sen, H. Hoi and M. Gupta, *ACS Appl Bio Mater*, 2021, **4**, 5416-5424.

42. A. R. Hall, A. Scott, D. Rotem, K. K. Mehta, H. Bayley and C. Dekker, *Nat Nanotechnol*, 2010, **5**, 874-877.

43. F. P. Qin, M. Q. Sheng, R. J. Li, D. X. Liu, M. L. Ren, R. Tian, L. Douadji, D. Q. Wang, Q. Gan and L. Y. Liang, *Chem. Eng. J.*, 2024, **489**, 151006.

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

45. S. Pal, A. Naik, A. J. A. Rao, B. Chakraborty and M. M. Varma, *ACS Appl. Nano Mater.*, 2022, **5**, 8804-8810.

46. R. Wei, T. G. Martin, U. Rant and H. Dietz, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 4864-4867.

47. M. Firnkes, D. Pedone, J. Knezevic, M. Doblinger and U. Rant, *Nano Lett.*, 2010, **10**, 2162-2167.

48. T. Albrecht, T. Gibb and P. Nuttall, in *Engineered Nanopores for Bioanalytical Applications*, eds. J. B. Edel and T. Albrecht, William Andrew Publishing, Oxford, 2013, pp. 1-30.

49. V. Tabard-Cossa, in *Engineered Nanopores for Bioanalytical Applications*, eds. J. B. Edel and T. Albrecht, William Andrew Publishing, Oxford, 2013, pp. 59-93.

50. W. Tang, J. P. Fried, R. D. Tilley and J. J. Gooding, *Chem. Soc. Rev.*, 2022, **51**, 5757-5776.

51. O. Otto and U. F. Keyser, in *Engineered Nanopores for Bioanalytical Applications*, eds. J. B. Edel and T. Albrecht, William Andrew Publishing, Oxford, 2013, pp. 31-58.

52. R. M. Smeets, U. F. Keyser, D. Krapf, M. Y. Wu, N. H. Dekker and C. Dekker, *Nano Lett.*, 2006, **6**, 89-95.

53. Y. Zhang, G. S. Wu, W. Si, J. Ma, Z. S. Yuan, X. Xie, L. Liu, J. J. Sha, D. Y. Li and Y. F. Chen, *Nanoscale*, 2017, **9**, 930-939.

54. M. Wanunu, W. Morrison, Y. Rabin, A. Y. Grosberg and A. Meller, *Nat Nanotechnol*, 2010, **5**, 160-165.

55. M. Muthukumar, *J. Chem. Phys.*, 2010, **132**, 195101.

56. K. Chen, N. A. W. Bell, J. Kong, Y. Tian and U. F. Keyser, *Biophys. J.*, 2017, **112**, 674-682.

57. N. A. W. Bell, K. Chen, S. Ghosal, M. Ricci and U. F. Keyser, *Nat Commun*, 2017, **8**, 380.

58. R. L. Fraccari, P. Ciccarella, A. Bahrami, M. Carminati, G. Ferrari and T. Albrecht, *Nanoscale*, 2016, **8**, 7604-7611.

59. A. P. Ivanov, P. Actis, P. Jonsson, D. Klenerman, Y. Korchev and J. B. Edel, *ACS Nano*, 2015, **9**, 3587-3595.

60. A. E. Nkodo, J. M. Garnier, B. Tinland, H. Ren, C. Desruisseaux, L. C. McCormick, G. Drouin and G. W. Slater, *Electrophoresis*, 2001, **22**, 2424-2432.

61. M. Wanunu, J. Sutin, B. McNally, A. Chow and A. Meller, *Biophys. J.*, 2008, **95**, 4716-4725.

62. N. A. Bell, M. Muthukumar and U. F. Keyser, *Phys Rev E*, 2016, **93**, 022401.

63. P. Chen, J. Gu, E. Brandin, Y. R. Kim, Q. Wang and D. Branton, *Nano Lett.*, 2004, **4**, 2293-2298.

64. A. J. Storm, J. H. Chen, H. W. Zandbergen and C. Dekker, *Phys Rev E*, 2005, **71**, 051903.

65. M. Mihovilovic, N. Hagerty and D. Stein, *Phys. Rev. Lett.*, 2013, **110**, 028102.

66. D. Y. Ling and X. S. Ling, *J Phys Condens Matter*, 2013, **25**, 375102.

67. J. Li and D. S. Talaga, *J Phys Condens Matter*, 2010, **22**, 454129.

68. L. P. Liu, X. Xie, J. L. Kong, H. W. Wu and Q. J. Liu, *Sci Adv Mater*, 2013, **5**, 2032-2038.

69. J. Li, M. Gershow, D. Stein, E. Brandin and J. A. Golovchenko, *Nat Mater*, 2003, **2**, 611-615.

70. C. Plesa and C. Dekker, *Nanotechnology*, 2015, **26**, 084003.

71. R. L. Fraccari, M. Carminati, G. Piantanida, T. Leontidou, G. Ferrari and T. Albrecht, *Faraday Discuss.*, 2016, **193**, 459-470.

72. A. Y. Y. Loh, C. H. Burgess, D. A. Tanase, G. Ferrari, M. A. McLachlan, A. E. G. Cass and T. Albrecht, *Anal. Chem.*, 2018, **90**, 14063-14071.

73. D. Fologea, E. Brandin, J. Uplinger, D. Branton and J. Li, *Electrophoresis*, 2007, **28**, 3186-3192.

74. C. Plesa, D. Verschueren, S. Pud, J. van der Torre, J. W. Ruitenberg, M. J. Witteveen, M. P. Jonsson, A. Y. Grosberg, Y. Rabin and C. Dekker, *Nat Nanotechnol*, 2016, **11**, 1093-1097.

75. C. Plesa, L. Cornelissen, M. W. Tuijtel and C. Dekker, *Nanotechnology*, 2013, **24**, 475101.

76. R. Kumar Sharma, I. Agrawal, L. Dai, P. S. Doyle and S. Garaj, *Nat Commun*, 2019, **10**, 4473.

77. Y. H. Sen, T. Jain, C. A. Aguilar and R. Karnik, *Lab Chip*, 2012, **12**, 1094-1101.

78. M. A. Edwards, S. R. German, J. E. Dick, A. J. Bard and H. S. White, *ACS Nano*, 2015, **9**, 12274-12282.

79. N. A. Bell and U. F. Keyser, *Nat Nanotechnol*, 2016, **11**, 645-651.

80. C. Plesa, N. Van Loo, P. Ketterer, H. Dietz and C. Dekker, *Nano Lett.*, 2015, **15**, 732-737.

81. D. Fologea, J. Uplinger, B. Thomas, D. S. McNabb and J. Li, *Nano Lett.*, 2005, **5**, 1734-1737.

82. G. Ando, C. Hyun, J. Li and T. Mitsui, *ACS Nano*, 2012, **6**, 10090-10097.

83. O. M. Eggenberger, C. Ying and M. Mayer, *Nanoscale*, 2019, **11**, 19636-19657.

84. S. Li, S. Zeng, C. Wen, L. Barbe, M. Tenje, Z. Zhang, K. Hjort and S. L. Zhang, *J. Phys. Chem. B*, 2020, **124**, 11573-11583.

85. E. C. Yusko, J. M. Johnson, S. Majd, P. Prangkio, R. C. Rollings, J. Li, J. Yang and M. Mayer, *Nat Nanotechnol*, 2011, **6**, 253-260.

86. K. J. Freedman, L. M. Otto, A. P. Ivanov, A. Barik, S. H. Oh and J. B. Edel, *Nat Commun*, 2016, **7**, 10217.

87. M. Zhang, C. Ngampeerapong, D. Redin, A. Ahmadian, I. Sychugov and J. Linnros, *ACS Nano*, 2018, **12**, 4574-4582.

88. C. Yuan, F. Zhou, Z. Xu, D. Wu, P. Hou, D. Yang, L. Pan and P. Wang, *Chembiochem*, 2024, **25**, e202400227.

89. Y. Xing, A. Rottensteiner, J. Ciccone and S. Howorka, *Angew. Chem. Int. Ed. Engl.*, 2023, **62**, e202303103.

90. T. Albrecht, *Curr Opin Electroche*, 2017, **4**, 159-165.

91. A. Singer, M. Wanunu, W. Morrison, H. Kuhn, M. Frank-Kamenetskii and A. Meller, *Nano Lett.*, 2010, **10**, 738-742.

92. A. Singer, S. Rapireddy, D. H. Ly and A. Meller, *Nano Lett.*, 2012, **12**, 1722-1728.

93. L. Xue, H. Yamazaki, R. Ren, M. Wanunu, A. P. Ivanov and J. B. Edel, *Nat Rev Mater*, 2020, **5**, 931-951.

94. R. M. Kuanaeva, A. N. Vaneev, P. V. Gorelkin and A. S. Erofeev, *Biosensors (Basel)*, 2024, **14**.

95. N. A. Bell and U. F. Keyser, *J. Am. Chem. Soc.*, 2015, **137**, 2035-2041.

96. J. Kong, N. A. Bell and U. F. Keyser, *Nano Lett.*, 2016, **16**, 3557-3562.

97. J. L. Kong, J. B. Zhu and U. F. Keyser, *Chem. Commun.*, 2017, **53**, 436-439.

98. J. Y. Y. Sze, A. P. Ivanov, A. E. G. Cass and J. B. Edel, *Nat Commun*, 2017, **8**, 1552.

99. E. Beamish, V. Tabard-Cossa and M. Godin, *ACS Sens*, 2017, **2**, 1814-1820.

100. F. Boskovic, C. Maffeo, G. Patino-Guillen, R. Tivony, A. Aksimentiev and U. F. Keyser, *ACS Nano*, 2024, **18**, 15013-15024.

101. F. Boskovic and U. F. Keyser, *Nat Chem*, 2022, **14**, 1258-1264.

102. J. Zhu, R. Tivony, F. Boskovic, J. Pereira-Dias, S. E. Sandler, S. Baker and U. F. Keyser, *J. Am. Chem. Soc.*, 2023, **145**, 12115-12123.

103. F. Zheng, M. Alawami, J. Zhu, C. M. Platnich, J. Sha, U. F. Keyser and K. Chen, *Nano Lett.*, 2023, **23**, 11145-11151.

104. Y. L. Ying, Z. L. Hu, S. Zhang, Y. Qing, A. Fragasso, G. Maglia, A. Meller, H. Bayley, C. Dekker and Y. T. Long, *Nat Nanotechnol*, 2022, **17**, 1136-1146.

105. J. Stanley and N. Pourmand, *APL Mater*, 2020, **8**.

106. F. Zheng and Q. Han, *Analyst*, 2024, **149**, 5131-5138.

107. R. A. Al-Waqfi, C. Khan, O. J. Irving, L. Matthews and T. Albrecht, *arXiv preprint arXiv:2501.14347*, 2025.

108. R. A. Al-Waqfi, C. J. Khan, O. J. Irving, L. Matthews and T. Albrecht, *ACS Nano*, 2025, **19**, 16803-16812.

CHAPTER 2
MATERIALS AND METHODS

CONTENTS

Synopsis: This chapter introduces the methods and materials for all practical experiments conducted in this thesis.

The first part of this chapter includes the molecular biology methods for the design and assembly of DNA carriers, in-vitro RNA synthesis, gel electrophoresis and UV-Vis spectroscopy. The second part includes the analytical chemistry part which describes the fabrication and characterization of nanopipettes, DNA translocation set up and instrumentation. Finally, this chapter describes the method used for characterizing DNA and RNA by atomic force microscopy (AFM).

2.1. Molecular biology methods

2.1.1. Gel Electrophoresis

Gel electrophoresis experiments were performed using the following protocol for DNA samples: 1% TAE agarose gel was prepared by adding 1g of Agarose (Biotium) to 98 mL of Mili-Q water and 2 mL 50X Tris-Acetate-EDTA buffer (TAE, Fisher bioreagents) to the final solution, all were placed in a glass vial which was heated in the microwave to dissolve the agarose, the solution was kept till it cooled to 50 ^{0} C. To form the gel, 20 mL of the abovementioned solution was poured in a tray with a well comb and then were left to solidify at room temperature after 30-45 minutes. For RNA samples, 1% agarose TBE gel was prepared by adding 1g of agarose to 95 mL of NFW water with 5 mL of 10X TBE buffer (Tris-Borate-EDTA, Merck) and then dissolved and poured as the TEA gel steps.

DNA samples were diluted to lower than 50ng/µL and then 2-5µL of DNA sample was mixed with 1µL of 6X DNA loading dye (ThermoFisher) and NFW water was added up to 6 µL of total volume. Two sources of standard DNA ladders were used depending on availability GeneRuler 1 kb DNA Ladder, ready-to-use (ThermoFisher) or 1 kb DNA Ladder (NEB). In both cases, 5 µL of each ladder was mixed with 1 µL of 6XDNA loading dye. For ssRNA ladder (NEB), 2 µL were mixed with 4µL NFW. After the gels solidified, all samples were loaded into the wells, starting with the DNA ladder in the first well. The gels were run in a BIORAD electrophoresis chamber filled with the same buffer used to prepare the gels. A 1% agarose 1xTAE gel was run at 80 V for 45-50 minutes, and a 1% agarose 0.5xTBE agarose gel was run at 100 V for 60 minutes. TAE buffered DNA Gels were then placed in tray containing staining solution which was prepared freshly of 10 µL of 10000X SYBR^{TM} Gold (Invitrogen, part of Thermo fisher Scientific) in 100 mL

of 1X TAE buffer. For gels buffered with TBE, 8 µL of peqGREEN (VWR International) was mixed in 100 mL of 0.5X TBE for staining. The gel tray was placed on a shaker (Merck) for 40-50 minutes. Finally, images were captured using a UV illumination box (Bio-Rad Instruments). Saved images were analysed using Fiji-ImageJ.[1]

2.1.2. UV-Vis Spectroscopy and Fluorescence Analysis of DNA/RNA for The Characterisation of DNA and RNA Samples

For measurement and evaluation of DNA purity and PCR success, the concentration was measured using 1-4 µL of the sample at 260/280 nm absorbance ratio (BioSpec-nano, Shimadzu). The concentration of in vitro-transcribed RNA samples was quantified using a Qubit™ 3.0 Fluorometer (Invitrogen, Fisher Scientific) and the Qubit™ RNA Broad Range Assay Kit (Invitrogen, Fisher Scientific), according to the manufacturer's protocol. Briefly, the assay procedure involved the preparation of a working reagent by diluting Qubit© RNA BR reagent 1:200 in Qubit© RNA BR buffer. Using the standard reagents provided by the assay, two standards were prepared by diluting 10 µL of each standard in 190 µL. Standards were prepared for the purpose of calibrating the instrument. Hence, this will ensure consistent measurement of the sample. Finally, 1 µL of RNA sample was added to 199 µL of the working reagent in the assay tubes. After a 2-minute incubation at room temperature, fluorescence was measured for both sample and standards using the Qubit™ 3.0 Fluorometer and RNA concentration reading was obtained as ng/ µL.

2.1.3. DNA Standards and Constructs

2.1.3.1. DNA Standards

The commercially available dsDNA standards used in this thesis were 4 kbp, 7 kbp and 10 kbp (Thermo Fisher, no limits DNA fragments). 48.5 kbp (λ) DNA was purchased from NEB, standards' purity and length were assessed using gel electrophoresis and UV-Vis spectroscopy as described in section 2.1.1 and 2.1.2, respectively. All DNA samples were provided at a stock concentration of 0.5 µg/µL and kept in the freezer at -20^{0}C.

2.1.3.2. UCA-1 gene plasmid

UCA-1 gene (2314 bp, NR_015379)[2, 3] was synthetically assembled as a plasmid by Invitrogen (part of Thermo Fisher Scientific) through insertion into pMA-RQ (AmpR) circular vector with Sp6 promotor sequence (5' ATTTAGGTGACACTATAG) at the beginning of the top strand, and two SfiI specific Enzyme Restriction sites (RE) at both ends of specific sequence (refer to Appendix 1 to view plasmid map and sequences), the plasmid construct was reconstituted as 200 ng/µL and kept in the freezer at -20^{0}C.

2.1.3.3. PCR of 10 kbp

10 kbp PCR prepared DNA used in chapter 4 was optimised and synthesized by Cengiz Khan a PhD student in the Albrecht group.

To amplify biotinylated 10kbp DNA, the PCR reaction was prepared using 2 primers provided by Integrated DNA Technologies (IDT); the forward (5'-TCATCAGGGCGAGATGCTCAATG) and the reverse (5'-[Biotin]-AAGGCGTTTCCGTTC TTCTTCGT). The primers were reconstituted as a stock solution (100 µM) in nuclease-free water (VWR international). Briefly, the reaction

mixture contained 25 µL of Q5 High-Fidelity 2X Master Mix (NEB), 2.5 µL of each primer (10 µM) , 1µL of λ DNA as template DNA (1 ng/ µL) and NFW (VWR international) was added to bring the final reaction volume up to 50 µL in LoBind© 0.2 mL tubes (Eppendorf). The reaction components were gently mixed by pipetting and briefly centrifuged to ensure homogeneity. The PCR reaction tube was placed inside a thermocycler (PrimeG©, Cole Palmer) and the following parameters were used to amplify the 10 kbp DNA as shown in Table 2.1:

Table 2.1 PCR thermocycling parameters for the synthesis of ~10 kbp DNA

Step	**Temperature(°C)**	**Time (s)**
Initial denaturation	98	30
30 cycles	98 70 72	10 30 300
Final extension	72	120
Hold	5	

Finally, PCR product was purified using Monarch PCR and DNA Clean-Up kit (NEB) and eluted in nuclease-free water. The length and purity were assessed as provided in section 2.1.1 and 2.1.2, respectively.

2.1.3.4. M13 constructs:

Preparation and Characterization of the DNA hybrid structure (capturing probe): M13 constructs were assembled as previously reported by Loh[4] with adaptation of staple sequences from Bell.[5] M13mp18 circular ssDNA (250 ng/µL), and M13mp18 RF (100 ng/µL) circular dsDNA were acquired from New England Biolabs (NEB). 190 staple oligos (<36 bases) and ssDNA-overhang sequences (>100 bases) were provided by Integrated DNA Technologies (IDT) with overhang sequences listed in Appendix 1.

Building construct included two main steps as the following:

a) Linearizing M13mp18 ssDNA: 1 µL of restriction site oligo (100 µM, added to introduce recognition sites for EcoRI and BamHI digestion) was mixed with 5 µL of circular M13mp18 ssDNA, 5 µL rCutSmart buffer(NEB) and 37 µL of nuclease free water(VWR international) in LoBind© 0.2 mL tubes (Eppendorf). Then, the mixture was heated up to 65 °C for 5 min, rapidly cooled for 5 minutes at 23 °C, and finally placed on hold at 10 °C for 10 min in the thermocycler. For RE digestion, 1 µL of EcoRI-HF (20000 U/mL, NEB) and 1 µL of Bam-HF enzymes (20000 U/mL, NEB) were added to the LoBind©PCR reaction tube and the reaction was incubated at 37°C for two hours. Heat liable enzymes were deactivated by heating at 65°C for 30 minutes. The circular M13mp18 dsDNA RF was linearized in another LoBind©PCR reaction tube in similar procedure as control for confirming the success of linearization of M13mp18 ssDNA. Both reactions were purified after digestion using PCR Monarch clean-up kit as the protocol provided by NEB (for linear ssDNA purification the DNA binding buffer volume was added in a ratio of 7:1 to the final volume of the pre-purified mixture), the final sample volume was 50 µL.

b) DNA assembly of the unbound carrier: the unbound construct was assembled by the addition of 21 µL of linearized ssM13mp18 (11.2 nM), 1 µL of staple oligos, 2 µL of each overhang oligos, 20 µL of $MgCl_2$ (25 mM) and 14 µL of nuclease free water. The assembly reaction was placed in the thermocycler and heated up to 73^{o}C, then gradually cooled to 23^{o}C for 49 cycles (at 1^{o}C/4 minutes rate for each cycle). Amicon Ultra 100 kDa cutoff centrifugal filters (Millipore, Sigma-Aldrich) were used to filter out the constructs from excess oligonucleotides. Briefly, the mixture was placed in a filter after mixing with 450 µL 1x TE buffer then centrifuged at room temperature for 8 minutes at 6000 g force. This step was repeated three times. Sample recovery was accomplished by inverting the filter in a fresh collecting column and then centrifuging at maximum force (*9000 g*) for 45 seconds. Success at each step was assessed by gel electrophoresis (1% agarose TAE gel, see section 2.1.1).

c) Binding the carrier with RNA: to bind the carrier with the in-vitro transcribed UCA-1 lncRNA, 25 µL of the unbound carrier were mixed with 2ul of RNA (236ng/ul), 0.6 µL of EDTA (0.5 M) and 3 µL of $MgCl_2$ (100 mM) in LoBind©PCR reaction tube. Then, the reaction was incubated at 40^{O}C for 16 hrs.

2.1.4. In-Vitro Transcription of UCA-1 lncRNA (free target)

a) Deactivation of RNAases in nuclease-free water:

Nuclease free water was pretreated with 1X RNAsecure™ Reagent (Invitrogen, Thermo Fisher Scientific) and heated up to 60 ^{0}C, then left to cool down to room temperature before use in the subsequent steps.

b) Plasmid linearization and Restriction digestion

Restriction digestion of UCA-1 plasmid (section 2.1.3.2) to get the linearized UCA-1 gene was performed as follows: 5 µL of plasmid DNA (200 ng/µL), 1 µL of SfiI enzyme(NEB, 20 unites), 5 µL of 10X rCutSmart Buffer and 38 µL of nuclease-free water (see 1.4/b), then the mixture was incubated at 50^{o}C for one hour in the thermocycler. To remove sticky overhangs, 1 µL Klenow DNA Polymerase I, Large Fragment (neb) and 1.65 µL of dNTP mix (1mM) were added and the digested product was incubated at 25^{0}C for 30 minutes. Klenow DNA Polymerase I enzyme was heat deactivated at 75^{0}C for 20 minutes. Finally, the linearized dsDNA target gene was gel extracted and purified using the Monarch PCR clean up kit into a final volume of 50 µL at a concentration of 20 ng/µL. Complete digestion and purity was confirmed using 1% agarose TAE gel and the concentration of the linearized target gene was determined using UV-Vis spectroscopy as provided in section 2.1.1 and 2.1.2, respectively.

c) In-Vitro Transcription (IVT)

HiScribe®SP6 RNA Synthesis Kit (NEB) was used for in-vitro transcription of the linearised UCA-1 gene into long non-coding RNA. Briefly, the reaction solution was prepared in LoBind©PCR reaction tubes by mixing 2.5 µL of each of the following: ATP (Tris, 50 mM), GTP (Tris, 50 mM), UTP (Tris, 50 mM), ATP (Tris, 50 mM), SP6 Reaction Buffer (10X) and SP6 RNA Polymerase Mix. 10 µL of linearized template was then added to the reaction master mix. The mixture was then centrifuged at 3000 rpm for 10-15 seconds and incubated at 37 ^{0}C for 2 hours. To remove the linearised DNA template, the reaction mixture was diluted with 25 µL NFW and then 2 µL of DNase I (included in the kit) added, after pipetting gently the reaction was incubated at 37°C for 20 s. In-vitro transcribed RNA was purified using PureLink^{TM} FFPE total RNA isolation kit (Invitrogen, Thermo Fisher Scientific) according to manufacturer

protocol. RNA reaction yield and length was verified using Qubit™ 3.0 Fluorometer (236 ng/µL, see section 2.1.2) and 1% Agarose 0.5x TBE Gel (see section 2.1.2 and 2.1.1), respectively.

2.2. Nanopore Device Setup

2.2.1 Nanopipettes Preparation

Nanopipettes are the main type of solid state nanopores used in our lab, as they are cheaper and easier to fabricate in comparison with other devices, demonstrating better surface wettability, and improved electric properties (low noise).[6-8] Nanopipettes were prepared from filamented quartz capillaries (outer diameter: 1 mm, inner diameter: 0.5 mm; length: 7.5 cm, Sutter Instruments©, Novato, USA) by laser-assisted mechanical pulling using a pipette puller (P2000, Sutter Instruments©)[9]. All quartz capillaries were plasma cleaned at 0.8 Torr and medium ionizing power for 2-5 minutes prior to the translocation experiment for cleaning and to enhance wettability.[10]

The pulling parameters included two programs: the first one was previously optimized by the Albrecht group and produced a pore size below 20 nanometres (also shorter taper length).[11-13] The second program was an optimised program. This program was used when inconsistent pore sizes obtained using the first program (both parameters are mentioned in Table 2.2). Furthermore, the main pulling parameters included: Heat; which control the amount of energy required to heat the quartz, Filament; is the actual length of the section of quartz capillary that will be heated; Velocity; is the velocity at which the capillary would be pulled from each side before the final pull, delay is the lag period before the final pull which determines the force needed for taper and fine tip formation.[9] Both programs had two lines; the first line was for taper extension and the second line was for pulling to separate the two connected tip of

the nanopipettes. It was worth noting that controlling lab humidity (setting up a dehumidifier) and temperature (around 20 ^{0}C) were relevant factors for maintaining pore size consistency and reproducibility, as it is well known that pulling parameters are not universal, and they are specific to the used instrument. To prevent surface aging, nanopipettes were pulled on-demand and used immediately after full characterisation.

Table 2.2: P2000 Pulling parameters programs used for pulling pore sizes below 20 nm

Program no.	Heat	Filament	Velocity	Delay	Pull
59	700	5	35	150	75
	700	0	15	128	200
56	700	4	35	150	80
	700	1	15	120	200

2.2.2 Nanopipette Taper Length Measurements

After pulling the pipettes, the taper length of the nanopipettes was determined using optical microscopy (Nikon, T-105, 4X objective lens) as shown in Figure 2.1. Taper is the longitudinal length that describes the gradual narrowing from the wider bulk region to the end of the fine tip.[12]

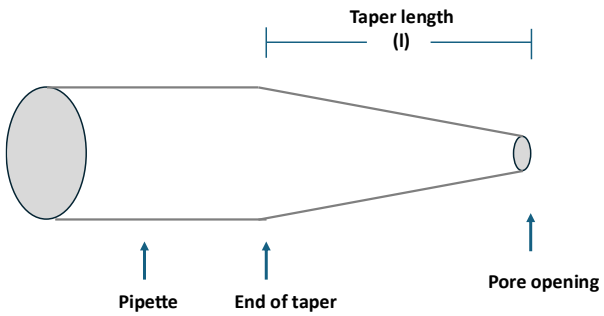

Figure 2.1 Drawing illustrates the region determined from optical microscopy for the measurement of taper length.

2.2.3 Silver/Silver Chloride Electrodes (Ag/AgCl)

4 M LiCl solution (prepared from powder form, Sigma) buffered at pH 8.0 with 1xTE buffer (10 mM Tris-HCl, 1 mM EDTA, Fisher Bioreagents) was the main electrolyte solution for anodization, conductivity measurement and nanopore translocation experiments, unless otherwise specified. Initially, this solution was filtered using 0.2 µm syringe filter (Merck) to remove debris and form a clean surface, the process included oxidation of silver to silver oxide by immersing 8-15 cm cut of silver wire (0.25 mm, 99.99% purity, Goodfellow) in 38% v/v nitric acid (Sigma) for 10 seconds and then washed using Ultra-pure (MiliQ-water©, 18 MΩ, Merck Millipore). Afterwards, sliver wires were soldered to gold contact pins and submerged in 4 M LiCl 1xTE solution. Anodization of electrodes proceeded after connecting sliver wire to an electrochemical cell at 1 mA using gold as counter electrode (99.99% purity, Goodfellow). Finally, the current supply was disconnected when silver electrode colour turned to black/purple. Electrodes were then washed with MilliQ-water and kept away from light until use (but no longer than 3 days*).*[13]

2.2.4 Pore Size Estimation and Conductance Measurement

2.2.4.1 Liquid Cell design

DNA translocation and conductance measurements for pore size determination were initially set up using the previously custom designed electrochemical cell (3 mL volume capacity vertical glass vials).[6] Steps for cell preparation involved: cleaning glass vials by sonication with 70% ethanol for 20 minutes followed by another sonication round with MilliQ-water for 10 minutes, respectively. Following drying, the vials were filled with the filtered 4 M LiCl 1xTE solution (see section 2.2.1). Two holes were drilled in the lid of the vials to act as holder for

both bulk immersed electrode and nanopipettes (see Figure 2.2A). Again, 4M LiCl 1xTE solution was backfilled into the nanopipette slowly using a Microfil™ needle (World Precision Instruments, WPI) for blocking bubble build up near the sensing region, afterwards, another electrode was inserted inside the nanopipette. To reduce the required amount of DNA for translocation experiment, an optimised electrochemical cell design (see Figure 2.2B) was built to accommodate volumes ranging from 100 to 300 µL, based on microscope glass slides (length: 76.2 mm and thickness: 1.2 mm, Fisher Scientific). Slides were cleaned as previously described for the glass vials. To prevent leaks, the clean and dry slide was wrapped in parafilm (VWR International), and a small square was cut in the middle of the slide with a new scalpel (~0.5cmx~0.5cm). Subsequently, the nanopipette and the electrode were fixed to the slide opposing each other as presented in Figure 2.2B 100-300 µL of 4M LiCl 1xTE solution was pipetted slowly into the air exposed region. This will hold the liquid by means of surface tension (Figure 2.2B). Finally, the small cell was kept in a closed petri dish.

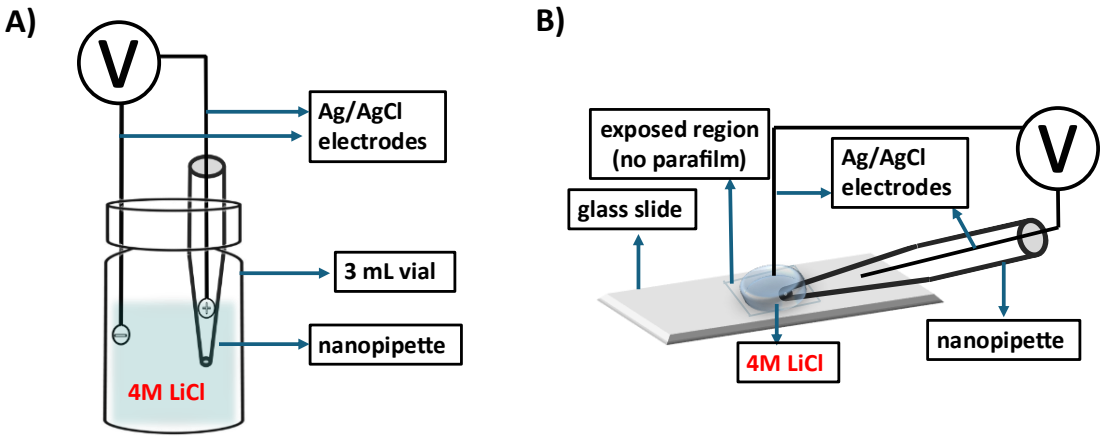

Figure 2.2: Schematic illustration of liquid cells set up used in nanopore experiments (A) represents the 3 mL liquid cell set up and (B) shows the optimised new design that can handle lower sample volume by means of surface tension. The symbol V signifies connection to Potentiostate or translocation set up (refer to section 2.2.5)

2.2.4.2 Conductance measurement

The conductance of the pore was measured by cyclic voltammetry (CV), both electrodes were attached to Potentiostat (Ivium) inside a Faraday cage.[4, 13] The electrode inside the nanopipettes represented the working electrode, whereas the bulk immersed electrode represented the counter electrode. To obtain current-voltage *(I/V)* relationship, a voltage bias was applied using Ivium software with the scanning parameters mentioned in Table 2.3. Liquid cell design incorporated two different setups for the translocation experiment, this could arguably affect the reliability of conductance measurement as ions migrate different distances. Hence, to ensure consistent conductance measurement in both cell setups, the distance between the electrodes in the electrolyte solution was kept approximately identical in both setups.

Table 2.3 : Ivuim software parameters used for CV measurements

Parameter	**Configuration**
Scan rate (mV/s)	50
Step Potential (mV)	10
Number of scans	4
Bias range (V)	-0.5 to +0.5

Nanopipettes with nonlinear or jaggy response were discarded. Further, nanopipettes selection criteria was based on the calculation of the ion current rectification ratio (r) using equation (2.1)[14, 15] provided in Table 2.4. Selected nanopipettes have r values ranged from 0.9 to 1.2. Next, pore size was estimated using equation(2.2)[12] in Table 2.4 :

Table 2.4: Equations for nanopipettes selection and pore size estimation (d_p)

	Formula	Description
Equation 2.1	$r = \left\lvert \frac{I_-}{I+} \right\rvert$	Where I_- is the ion current at V_{bias} =-0.5V and $I+$ is the ion current at V_{bias}=+0.5V.[15] r is the ion current rectification ratio
Equation 2.2	$d_p = \frac{4GL + \frac{\pi}{2} G D_i}{\pi D_i g - \frac{\pi}{2} G}$	Where *G* is the conductance of the pore, L the taper length of the nanopipette (by optical microscopy), Di is the inner diameter of the capillary (0.5 mm as provided by the manufacturer) and *g* is the conductivity of the used electrolyte solution (*173 mS cm^{-1}*)[11].

2.2.5 Translocation Experiment

2.2.5.1 DNA and RNA Concentration

DNA samples (see section 2.1.3.1) were injected into the liquid cells filled with 4M LiCl 1xTE solution (refer to section 2.2.4.1). After injection, final solution was mixed to insure homogeneous distribution of DNA molecules. Depending on DNA length and yield, DNA final concentrations ranged from 40 pM to 600 pM, unless stated otherwise. DNA stock concentrations were determined using UV-Vis spectroscopy as stated earlier in section 2.1.2. It is worth mentioning that a short blank run with only 4M LiCl 1xTE solution before adding DNA was performed, as a step to check up noise level and help stabilise the system. For RNA translocation experiment, in-vitro transcribed UCA-1 lncRNA was diluted up to 100 pM in the

same electrolyte condition (RNA stock concentration was determined using Qubit3 Fluorometer as described earlier).

2.2.5.2 Set up connection and data recording

Liquid cell was connected to the translocation set up which encompasses custom built low noise, wide-bandwidth amplifier.[6, 7] The electrode inside the nanopipette was connected to the amplifier, whereas the electrode outside the nanopipette in the bulk is attached for biasing. For instance, when the bulk electrode is biased negatively, the applied V_{bias} signifies a negative value. The amplifier has two output channels, nominally the direct current "DC" and the alternating current "AC" channels.[6, 7] The DC channel captures the slowly varying components of the electrical current (low-pass filtered at 7 Hz), providing valuable information about the steady-state open pore current (pore conductance). In contrast, the AC channel was designed to have a zero-mean, and is suitable to detect fast, transient modulations, such as those associated with molecular translocation events. The amplifier along with the liquid cell connection were enclosed in a double faraday cage to shield the recorded signals from electromagnetic noise interference. Further, analogue signal of the AC channel was filtered by an eight-pole, 100 kHz low-pass Bessel filter (Krohn-Hite Corporation). Analogue to digital conversion (ADC) was performed through a Picoscope 4262 (Pico Technology), and data were recorded at 1 MHz sampling rate

2.2.5.3 Data collection and analysis

Data collection was conducted with both polarity in which negative and positive biases applied in a range of ±0.4 to ±0.8 V. Each bias was applied for 1000s and recorded for 100 files (10s for each file), unless noted otherwise. Every data trace was saved along with separate log file

containing information about mean (μ) and standard deviation (σ) of both channels recordings, in accordance with their root-mean-square (RMS) noise. Prior to data analysis for longer DNA lengths, traces were digitally low-pass Bessel filtered at 50 kHz for 7 kbp and 10 kbp, and at 10 kHz for 48.5 kbp using a custom written MATLAB code (provided by Prof. Tim Albrecht). Filtering at these frequences was performed to ensure signal clarity and reduce the potential for noise while maintaining the necessary resolution for detecting slower events.[16]

Event detection involved using another custom written MATLAB code (courtesy to Prof. Tim Albrecht). Specifically, the three numbered steps involved in event detection are presented in Figure 2.3 as follows:

1) the first step was to adjust the level of AC channel output to a zero-baseline to eliminate small current offset (usually <10 pA). [4]

2) Afterwards, to define a baseline cutoff through baseline crossing, the standard deviation (σ) for the AC channel output for each trace was calculated. Accordingly, a 5 σ cutoff was chosen to capture events. The 5σ cutoff was adjusted to lower or higher values as needed, this was contingent on the level of AC channel noise and the requirement for each individual experiment. In the line of this, events related segments of the current-time trace were extracted from and up to the adjacent baseline transition point and related event features determined. Extracted features included current maxima (ΔI_{Max}) or effective current (ΔI_e). Effective current (ΔI_e) values of the events were calculated by dividing the event charge deficit (q) by event duration (τ). Where, Event charge deficit (q) is the integral of the current signal over the duration of the detected event.[4, 6, 7, 17]

3) The event duration (τ) was calculated after defining both event start points and stop points, then multiplied by time step. For accurate determination of events shapes and durations, ,1σ cutoff was set to define event start and stop from the baseline.[4] Finally, the extracted features of the events were plotted as 2D scatter plots either as effective current ($\Delta I_{effective}$) or current magnitude (ΔI_{Max}) vs duration (τ). 2D plots were transformed to a log-log plot as this will provide easier separation.[4, 6, 7] All data plotting and analysis were pursued using OriginPro2021[18] and MATLAB software.

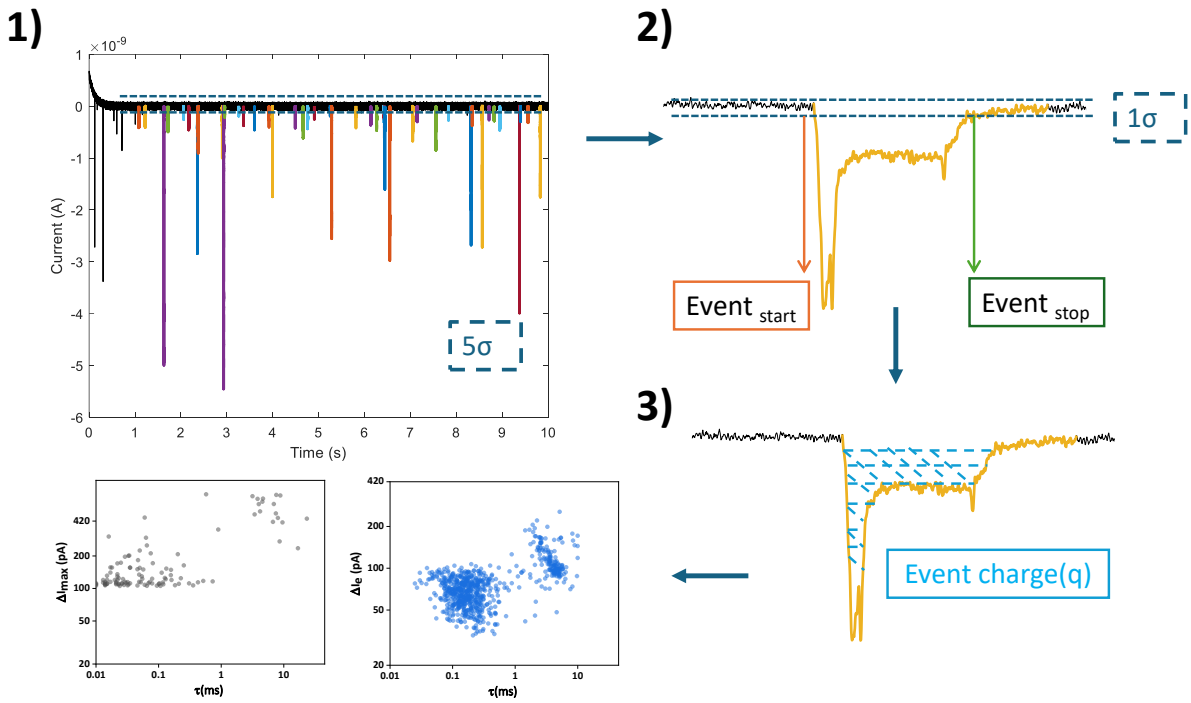

Figure 2.3: schematic diagram of steps performed for nanopore DNA event detection and analysis as explained in the main text (see 2.2.5.4, steps1-3). The data plotted in step3 are illustrative and do not show the actual data.

2.2.5.4 Linear and folded events separation

After plotting events features as described in section 2.2.5.3, DBSCAN algorithm, a density-based clustering method, was used as a built-in function in MATLAB to cluster out DNA

translocation events cluster (manual thresholding was applied in case DNA cluster was well resolved, refer to chapter 3).[4] To distinguish linear from folded for these events, five key event parameters were extracted, including charge (Δq), current maxima (ΔI_{Max}), effective current (ΔI_e), AC channel noise (σ), and event duration (τ). These features were standardised using z-score before being subjected to dimensionality reduction using principal component analysis (PCA). A k-means clustering algorithm was subsequently applied to the first two principal components, segmenting the events into two distinct populations. Later, the mean duration (τ_m) for each cluster was obtained from Gaussian fit. The cluster with longer τ_m was classified as linear DNA translocation events, which was colour-coded in the effective current plots as described in chapter 3.

2.2.5.5 Subevent detection

For subevent detection described in chapter 4 and chapter 5. Separately, a defined subevent cutoff search was employed on each event starting from the median baseline of the event. Defining the subevent cutoff was based on the median of the ΔI_{max} values for the events of interest. After performing the search, relevant subevents features were then extracted including the number, position, charge, magnitude and duration for the subevents detected.

2.2.5.6 Nanoparticles translocation experiment setup

Two types of spherical gold nanoparticles were used in this thesis in separate experiments, with average diameters of 40 nm and 5 nm, respectively (both were purchased from NanoPartz, USA). Further, both types of nanoparticles were polymer-coated spherical gold nanoparticles and are mono-functionalized with streptavidin. The nanoparticles were reconstituted in PBS buffer and provided at a stock concentration of 11.6 nM for the 40 nm

particles and 3.7 μM for the 5nm particles. Both suspensions were kept in the fridge at 4°C and covered with foil to protect them from light exposure.

For translocation experiment, large particles were diluted to a concentration of 500 pM in 4M LiCl 1xTE solution, while smaller particles were diluted to 2.74 nM in the same solution. Before loading particles into the nanopipettes, cells were connected to the translocation set-up as mentioned in section 2.2.4. Blank run with 4 M LiCl 1xTE solution was ran to insure stability and proper formation electric double layer (EDL). Subsequently, after the cell set up was disconnected, diluted nanoparticles were loaded slowly inside the nanopipette using a Microfil™ needle. Prior to introducing biotynlated 10 kbp DNA into the bulk solution, a separate translocation experiment was performed of both types of nanoparticles loaded in the trans compartment. Afterwards, biotinylated 10 kbp DNA was introduced into the bulk solution at a concentration of 300 pM while the translocation experiment procedure as described in section 2.2.5.2 and 2.2.5.3, with modification to the applied bias sequence. Initially, a bias of -0.8 V was applied for 1000 seconds to drive the DNA into the nanopipette. After this, the bias was switched off for ~20 minutes to allow for incubation, during which biotin binding to streptavidin was meant to occur. Subsequently, the bias was switched on at +0.8V to unload the DNA from the nanopipette. Following the initial voltage application, additional biases of ±0.5 V, ±0.6 V, and ±0.7 V were applied to further explore the behavior of the DNA in the nanopipette.

2.3 Atomic Force Microscopy (AFM)

For AFM imaging, two buffer recipes were used [19, 20]: buffer 1 which was used to image M13 unbound construct sample at concentration of 0.2-0.6 nM (10 mM $MgCl_2$ (Sigma), 1xTE buffer

and 10 mM HEPES (Sigma)) and buffer 2 (40 mM HEPES, 1xTE buffer and 10 mM $NiCl_2$) used for imaging RNA samples at a concentration of 0.5 nM. Prior to adding DNA/RNA, both buffer solutions were filtered using a 0.2 µm syringe filter. 10-15 µL of diluted samples were mounted on a freshly cleaved mica (12 mm, Agar Scientific) attached to AFM magnetic disk and left for 10-15 minutes. Then, mica surface was washed by adding 1 mL of Milli-Q water for 2-3 times (repeated more, if extra wash was needed). After washing, the surface was dried by mild flow of N_2 air. Imaging was acquired using non-tapping (AC mode) on Keysight 9500 AF/SPM in air at room temperature with previously adapted parameters. Probes used for imaging were purchased from NanoWorld NCHR-10 (resonance Frequency: 320 kHz, Force: 42 N/m, Length: 125 µM, Width: 30 µm and Thickness: 4 µm). Processing and viewing images were pursued using Gwyddion.[21]

2.4 References

1. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak and A. Cardona, *Nat Methods*, 2012, **9**, 676-682.
2. X. S. Wang, Z. Zhang, H. C. Wang, J. L. Cai, Q. W. Xu, M. Q. Li, Y. C. Chen, X. P. Qian, T. J. Lu, L. Z. Yu, Y. Zhang, D. Q. Xin, Y. Q. Na and W. F. Chen, *Clin Cancer Res*, 2006, **12**, 4851-4858.
3. B. Neve, N. Jonckheere, A. Vincent and I. Van Seuningen, *Cancers (Basel)*, 2018, **10**, 440.
4. A. Y. Y. Loh, C. H. Burgess, D. A. Tanase, G. Ferrari, M. A. McLachlan, A. E. G. Cass and T. Albrecht, *Anal. Chem.*, 2018, **90**, 14063-14071.
5. N. A. Bell and U. F. Keyser, *Nat Nanotechnol*, 2016, **11**, 645-651.
6. R. L. Fraccari, P. Ciccarella, A. Bahrami, M. Carminati, G. Ferrari and T. Albrecht, *Nanoscale*, 2016, **8**, 7604-7611.
7. R. L. Fraccari, M. Carminati, G. Piantanida, T. Leontidou, G. Ferrari and T. Albrecht, *Faraday Discuss.*, 2016, **193**, 459-470.
8. J. Stanley and N. Pourmand, *Apl Materials*, 2020, **8**.
9. Sutter, *Sutter P-2000/F Laser-based Pipette Puller*, last accessed 2024.
10. S. Zhang, M. Li, B. Su and Y. Shao, *Annu Rev Anal Chem (Palo Alto Calif)*, 2018, **11**, 265-286.
11. T. Albrecht, *Annu Rev Anal Chem (Palo Alto Calif)*, 2019, **12**, 371-387.
12. L. J. Steinbock, A. Lucas, O. Otto and U. F. Keyser, *Electrophoresis*, 2012, **33**, 3480-3487.
13. O. J. Irving, University of Birmingham, 2022.
14. P. Actis, A. C. Mak and N. Pourmand, *Bioanal Rev*, 2010, **1**, 177-185.
15. Z. Siwy, E. Heins, C. C. Harrell, P. Kohli and C. R. Martin, *J. Am. Chem. Soc.*, 2004, **126**, 10850-10851.
16. J. B. Edel and T. Albrecht, *Engineered nanopores for bioanalytical applications*, William Andrew, 2013.
17. D. Fologea, E. Brandin, J. Uplinger, D. Branton and J. Li, *Electrophoresis*, 2007, **28**, 3186-3192.
18. OriginPro 2021*: OriginLab Corporation, Northampton, MA, USA*, 2021, **5**, 1-15.

19. J. Zhu, R. Tivony, F. Boskovic, J. Pereira-Dias, S. E. Sandler, S. Baker and U. F. Keyser, *J. Am. Chem. Soc.*, 2023, **145**, 12115-12123.

20. E. Peeters, M. Boon, C. Rollie, R. G. Willaert, M. Voet, M. F. White, D. Prangishvili, R. Lavigne and T. E. F. Quax, *Viruses*, 2017, **9**, 190.

21. D. Nečas and P. Klapetek, *Open Phys.*, 2012, **10**, 181-188.

CHAPTER 3

CROWDING EFFECTS DURING DNA TRANSLOCATION IN NANOPIPETTES

CONTENTS

Synopsis: This chapter explores the asymmetric dynamics of DNA translocation within the nanopipettes in both out-in ‘’forward’’ and in-out “reverse” directions. Specifically, the process involved running multiple cycles, each with loading and unloading stages. Hence, Nanopipettes were first loaded with DNA at relatively longer time before the bias was reversed to unload the pipette. It has been shown that there is a good agreement with the earlier studies regarding the increased average DNA translocation time duration τ in the reverse direction. However, a new phenomenon has revealed an increased τ in the forward direction which is assumed to be due to retaining translocated DNA molecules near the tip inside the nanopipettes. Further, the occurrence of this new phenomenon, termed as crowding of DNA molecules, and the capacity of nanopipettes were further studied by means of DNA length, current noise, and pore conductance. This could help in understanding the

mechanism of DNA trapping inside a confined space and the influence of surrounding DNA molecules on DNA translocation signals in nanopipettes.

3.1. Introduction

Biological, solid-state, and nanopipettes have all been explored as detection devices from sequencing, to DNA structures, and single molecule sensing.[1, 2] Notably, solid-state nanopores have seen significant advancements over the last two decades, driving their application in diagnostics, small molecule characterization, and single-molecule biophysics.[3, 4] However, utilizing them as storage and release devices at the nanoscale level has remained relatively under-explored, with few researchers noting this this effect.[5-11] The ability to repeatedly capture analytes provides a significant advantage in studying dynamic biological processes and improving the accuracy and reliability of measurements in biochemical assays. Particularly, this principle was first referred as "molecular ping pong".[12] This concept was first demonstrated by Gershow,[6] who showed capture and recapture cycles with varied time durations (T_{delay}) for 4 kbp and 6 kbp DNA molecules. Plesa et al. further expanded this study to include longer DNA molecules to slow the diffusion and study folding dynamics(≥ 48.5 kbp).[7] Both studies were conducted using chip-based nanopores, which are approximately symmetric on both sides. This symmetry results in equal dwell times for both capture and recapture events, facilitating detailed analysis of the molecular interactions and dynamics involved. As introduced in Chapter 1, quartz nanopipettes have emerged as a notable low-noise, cost-effective label- free sensing platformss.[13] Nanopipettes have been extensively explored for their potential in DNA translocation, protein detection, and drug delivery.[13-17] Ying et al. was the first to apply the DNA trapping principle and delivery using nanopipettes

relying on signals provided by fluorescence and optical detection of labelled ssDNA.[5] Further, Ivanov et al. used nanopipettes to deliver DNA by applying time-defined electrical pulses between two separate compartments.[18] Neither of these two explored the dynamics of DNA entering and exiting the nanopipettes. Nanopipettes are conically shaped nanopores, which introduces an asymmetry when the DNA enters or exits the pore.[1, 13, 14] The asymmetry has been demonstrated to affect the duration of DNA molecule exiting the pore, as shown by Bell and further studied by Chen and Zheng.[19-21] This effect highlights the importance of the geometry of nanopores in designing experiments and interpreting results, particularly in applications involving DNA translocation and recapture dynamics. Despite other approaches explored trapping DNA by the employment of dual nanopores in a term referred to a "tug-of-war".[8-11, 22] Previous studies have not further investigated whether the asymmetry of a single nanopipette can trap DNA molecules before being released from the trans compartment when the pipettes are loaded for longer time instead of fast recapture. On the other hand, bare nanopipettes experiments can provide a simpler tool for molecular trapping. Typical analysis involves short time acquisition and "loading" of the pipette. Where a bias is applied until a molecule is detected, before the bias is inverted and the molecule leaves the pore.[6, 7, 19-21] Whilst this offers the opportunity to capture a single molecule multiple times, typically if this is not followed directly after the analyte enters the sensing region of the pore, the analyte will be lost in the bulk of the pipette and is unlikely to be re-translocated. This presents an issue as it is required for the molecule to overcome diffusion barrier before it is captured[23] accompanied with the crucial requirement of sophisticated instrumental setup and experimental design.[24] Further, this approach has not explored whether the molecules inside the pore might affect the translocation characteristics of the following translocating

molecules. Intriguingly, extending the pipette loading time and subsequently evaluating its capacity to retain and release DNA could provide valuable insights into the interactions between molecules within a confined space, potentially advancing our understanding of molecular structures under confinement and open the door toward real time reaction monitoring inside a nanoconfined space in the future.

3.2. Experimental Objectives

The main purpose of this chapter was to analyse the effect of nanopipette's asymmetry on dsDNA translocation characteristics in both directions using a range of DNA lengths (4 kbp to 48.5 kbp). Initially, this study focused on examining the effect of prolonged DNA translocation, where the pipettes were loaded with DNA molecules at extended intervals (1000 s) while varying V_{bias}. Previous studies have employed a rapid recapture strategy and studied only the effect on the trans compartment. For instance, Zheng et al. used a rapid recapture strategy with a nanopipette and DNA carrier to achieve over 100 repeated translocations, focusing on detailed analysis in the trans compartment.[20] Opposed to previous work, as shown in Figure 3.1, the approach represented first loading the pipette with DNA for longer duration (1000 s, Figure 3.1 step 1), to investigate the effect of prolonged nanopipette loading in the forward direction (Figure 3.1, step 2). Afterwards, the bias was reversed to unload the

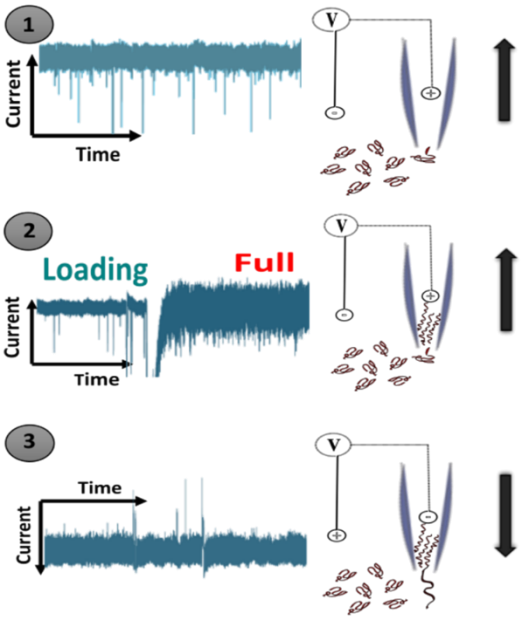

Figure 3.1 Illustration showing the electrochemical translocation cell set up and the principle of DNA loading and unloading using nanopipettes

pipette at the same durations (Figure 3.1, step 3). This could allow the identification of crowding effect while the DNA translocate into the pipette. Consequently, a new translocation state was observed when a DNA length-dependent threshold is exceeded. This newly identified state was characterised by increased dwell time(τ) and broader time distribution. Accordingly, the aim of this work was reshaped to fully understand the emergence of this distinct state. Particularly, DNA translocation characteristics in both directions were analysed in terms of DNA folding statistics, duration, and current magnitudes. Then, the electric properties of the pore during translocation including current noise and pore conductance were further studied to identify potential changes associated with the sensing regions.

3.3. Results and Discussion

3.3.1. DNA Purity /Gel Electrophoresis

Prior to DNA translocation, it was important to assess the quality and purity of DNA standards used in this study. Hence, the length and purity of commercially acquired DNA fragments were verified through gel electrophoresis (refer to Chapter2, section 2.1.1, for further details on gel recipes). Figure 3.2A shows the gel electrophoresis results for 4 kbp, 7 kbp and 10 kbp on a 1% TAE agarose gel. The 1st Lane and the 5th lane contained 1 kb DNA ladder, serving as a reference standard. Meanwhile, the 2nd, 3rd, and the 4th lanes correspond to 4 kbp DNA, 7 kbp and 10 kbp DNA fragments, respectively. Three distinct bands are visible in the lanes corresponding to the loaded DNA fragments. Relative to the DNA ladder, all bands' positions align with their expected migration level, confirming proper DNA sizing. However, migration can also be influenced by factors such as DNA conformation, buffer composition, and gel concentration. Apparently, DNA bands appear thicker in lane 2 and 3 compared to lane 4,

since the DNA samples in those lanes were loaded in higher amounts (500 ng). In Figure 3.2B, gel electrophoresis results for λ DNA sample were analysed under the same condition, in which the 1st lane contained λ DNA sample and the 2nd lane contained the 1 kbp DNA ladder. The band for λ DNA appears as single band higher than the largest band in the DNA ladder, confirming its purity as well as its larger size. Overall, both gel images provide evidence that the observed DNA bands correspond to the accurate DNA lengths in accordance with the DNA ladder. Finaly, to further assess the purity, the concentration of each standard DNA was determined by UV-Vis spectroscopy. Consistently, measurements of DNA standards have demonstrated an absorbance ratio of 260/280 close to 1.8, aligning with the widely accepted threshold for DNA purity.

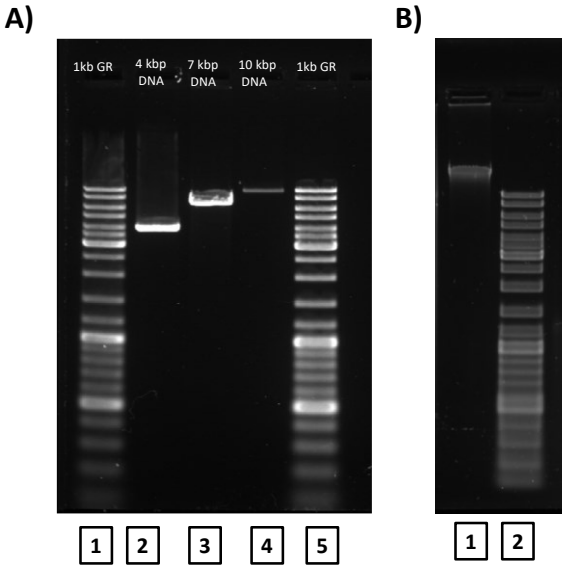

Figure 3.2: Agarose gel electrophoresis results of (A) DNA standard of lengths of 4,7 and 10 kbp (B) λ DNA (1% agarose in 1xTAE, running parameters: 80V for 49 minutes).

3.3.2. Characterization of Nanopipettes

To accurately characterize dsDNA, the nanopore diameter should allow the entry of DNA while maintaining optimal signal to noise ratio (SNR).[1] Specifically, nanopipettes were employed for

DNA translocation studies due to their superior attributes, including lower noise characteristics, exceptional stability, and straightforward fabrication processes as demonstrated in previous research. [25] Ultimately, nanopipettes were characterized by CV as described in Chapter 2, section 2.2.4. Figure 3.3 shows a current-potential (I-V) curve for 4 nanopipettes in 4M LiCl, along with representative optical micrograph of the taper of the nanopipette (bottom inset). All four pipettes show linear (ohmic) response. This linear correlation indicates that the bulk ion concentration dominates ion transport (salt concentration > 0.1 M), while the nanopipettes act as a major resistor with negligible surface charge effect. [25-27] Since the nanopipette geometry is not symmetric, this might affect the ion current flow from both directions.[27, 28] To investigate the effect of pore geometry on ion current flow, the ion current rectification ratio (r) was calculated for all nanopipettes used. The mean value of r= ~0.9± 0.2 was considered acceptable for successful nanopipettes in the line with previous studies (see section 2.2.4 in chapter2).[28-30] in conclusion, r values are in the propensity of an ohmic behaviour with slight deviation due to imperfections in nanopipette geometry.[29]

From I-V curves, the conductance (G_{pore}) of each pore was obtained by calculating the average slope of the curves. Following the measurement of the taper length (L), the size of the pore was estimated using equation 2.2 (Chapter 2, Section 2.2.4). Table 3.1 presents estimated pores size in nm for 12 nanopipettes used in this thesis, along with taper length(L) and G_{pore}.

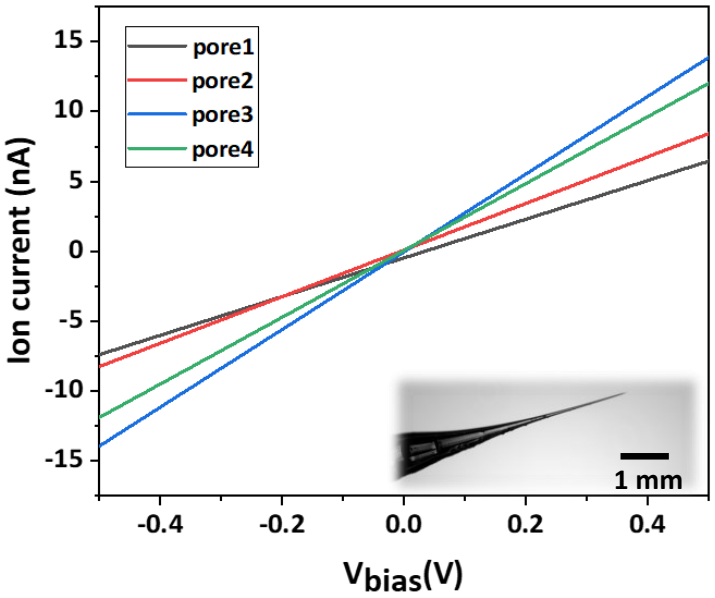

Figure 3.3: I/V_{bias} curve for four typical quartz nanopipette in 4 M LiCl with 1xTE buffer. The conductance (G_{pore}) values were: 14-27.5 nS, as determined from the average of the slopes from the forward and reverse scan between ±0.5 V. bottom inset: optical micrograph showing the taper of the nanopipette (G_{pore}=27.5nS) used for 7kbp experiment (L = 3.17 mm)

Table 3.1: Nanopipettes geometries

Pore #	Est. Pore size (nm)	Taper length (mm)	G_{pore} (nS)	Program #
1	14	3.12	26.76	1
2	20	3.05	48.4	1
3	14	3.17	27.5	1
4	18	3.2	36.9	1
5	12	3.13	31.6	1
6	23	3.1	44.1	1
7	11	3.24	21.5	1
8	13	3.15	26.1	2
9	13	3.14	26.3	2
11	16	3.1	35.7	1
12	13	3.2	24.4	1
	15.2 (±3.7)	**3.14 (±0.054)**	**31.7 (±8.58)**	

Overall, the estimated pore sizes of the nanopipettes were in a range of 11 nm to 23 nm (the average d_p = 15.2 ± 3.7 nm). Similarly, pore conductance (G_{pore}) for the nanopipettes has a variable range of values (21.5 nS - 48.4 nS) with a mean of 31.7 ± 8.58 nS, while taper length shows minimal variation with a mean of 3.14 ± 0.054 mm. A variable range of pore conductance (G_{pore}) values for the nanopipettes were employed based on the specific requirements of each experiment. This flexibility allowed us to tailor the nanopipette conductance to optimize performance and achieve reliable results across different experimental conditions. Taper length values were controlled at approximately at 3.1 mm. Although shorter taper length is preferred, controlling taper length at constant values would later place greater emphasis on pore size as the primary factor on DNA crowding (discussed later).

In nanopores research labs, it is a common practice to measure pore size with Scanning Electron Microscopy (SEM) or Transmission Electron Microscopy (TEM) as these techniques allow direct measurement of pore size with less estimation.[13, 14] Yet, this approach is time-consuming, making it impractical for every nanopipette. Therefore, the pore size was estimated using the equation 2.2 (Chapter 2, Section 2.2.4).[27] Overall, the estimated pore sizes of the nanopipettes used were approximately 11-23 nm. Moreover, prior TEM imaging was conducted in the Albrecht group for pore size measurement, and it has shown good agreement with the pore size values estimated by CV measurement.[15]

4M LiCl 1xTE solution was used as the main electrolyte solution. This choice has been supported by the previous work reported by Kowalczyk.[31] Specifically, changing the electrolyte solution from 1M KCl to 4M LiCl has slowed down the translocation time of kilobase pair DNA by factor of 10.[31] Conceptually, this can enhance temporal resolution which is necessary for

subtle structural differences identification within DNA molecules.[1] Such an effect was further explained due to smaller size of Li^+ accompanied by higher charge density and stronger binding affinity to the DNA backbone in comparison with K^+.[31]

3.3.3. Current Signal Measurement and Clustering in DNA Translocation

After nanopipettes characterisation, DNA translocation experiments were performed using a custom-built high-bandwidth amplifier (methods chapter, section 2.2.5.2). Figure 3.4 shows the AC channel output for three current time traces recorded in real time for 10 seconds per file. The first trace (Figure 3.4A) shows only ionic current measurement for 4M LiCl 1xTE at V_{bias}= -0.7 V, no dips or enhancements observed and only background noise with stable baseline. While Figure 3.4B shows the ionic current measurement at V_{bias}= -0.7 V after the injection of 7 kbp DNA ($C_{DNA, out}$= 600 pM) into the bulk solution, the trace shows downward spikes indicates blockage of the pore. Finally, Figure 3.4C shows a recorded trace after the voltage reversal (V_{bias} = +0.7 V). Upward spikes were observed against the stable current baseline. Theses spikes correspond to transient blockages, which are indicative of DNA molecules moving out of the pore. This occurs when a positive potential is applied to the bulk immersed electrode, reversing the direction of electrophoretic force (Chapter 1, section 1.5).

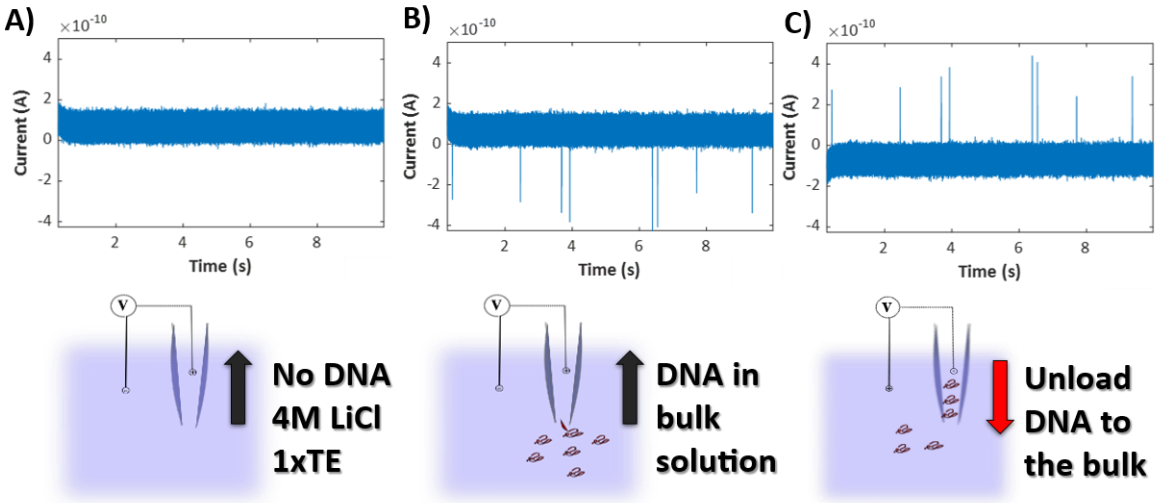

Figure 3.4 AC channel output of recorded current time traces for 10 seconds (G_{pore}=22.2 nS, taper length = 3.156 mm and d_{pore} ~ 14nm) where (A) blank solution of 4M LiCl 1xTE at -0.7V (B) 7kbp DNA was added to the bulk solution ($C_{DNA\ ,out}$ = 600 pM) at -0.7 V, black arrows indicates the direction of electrophoretic force from outside to inside the pipette and (C) recorded trace once the pipette was loaded with DNA and bias reversed to +0.7 V. Red arrow indicates the direction of electrophoretic force from the inside to the outside of the pipette, data were low-pass filtered at 100 kHz using 8-pole analogue filter (Krohn-Hite)

Following sufficient recording time, threshold-based event detection was proceeded to extract meaningful features for each spike (refer to Chapter 2). The AC channel has been corrected to zero-mean in our system. Hence, the background noise can be calculated by means of the standard deviation of the AC channel (σ). Toward this end, the application of a 5σ cutoff was sufficient to capture DNA translocation events out of the background noise.

Figure 3.5A and 3.5B present scatter plots of effective current ΔI_e vs duration (τ) for the events detected in 4kbp DNA experiment at V_{bias}=-0.7 V and V_{bias} = +0.6 V ($C_{DNA,\ out}$ = 300 pM, G_{pore} = 26.7 nS). Both plots reveal two distinct clusters. The first cluster is a well-defined cluster (red coloured) with comparable current magnitudes formed at approximately 0.4 ms and 1ms for V_{bias}= -0.7 V and V_{bias}= +0.6 V datasets, respectively. The emergence of this cluster follows the typical translocation behaviour of standard 4 kbp DNA in 4M LiCl as reported earlier.[26, 32, 33]

Coming back to Figure 3.5A, another sparse cluster (black coloured) is observed for events with shorter τ (<100 μs). Shorter duration events cluster has been previously documented, either due to baseline current fluctuation (background noise), or unsuccessful attempts of DNA translocation through the pore, where DNA molecules collide with the pore instead.[34] In Figure 3.5B, the second cluster is more scattered and irregular, this is more likely due to baseline current fluctuations.

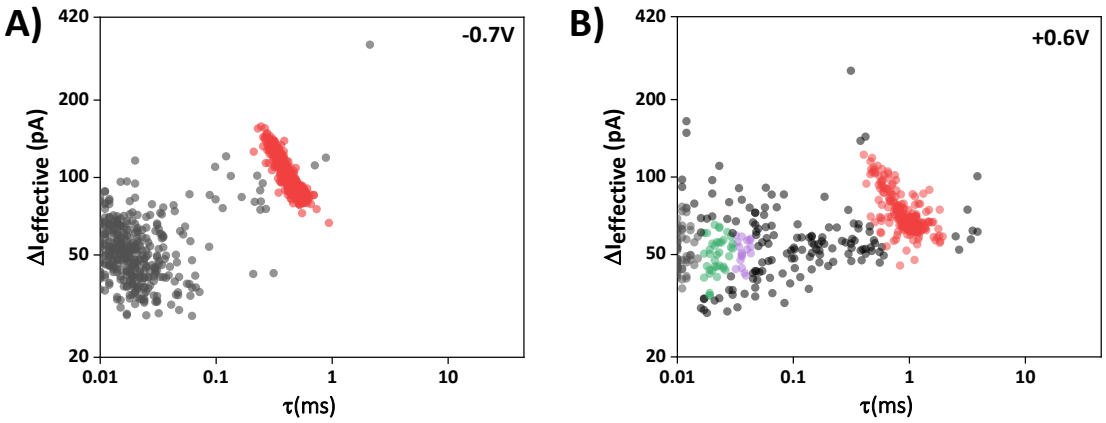

Figure 3.5 Effective current plots of detected events of 4kbp DNA (4M LiCl 1xTE, G_{pore}=26.7 nS). (A) shows a red coloured DNA cluster(-0.7V) that was separated by manual thresholding from background noise (black), whereas in panel (B) the red DNA cluster (+0.6V) was separated by DBSCAN clustering (epsilon=0.1 and minimum number of points "minpts" =20). Note that this not the actual sequence of bias steps for the experiment as this will be further explained in detail below. ($C_{DNA, out}$ = 300 pM), green and purple clusters are artifacts due to DBSCAN clustering algorithm and do not represent meaningful data groups.

After event detection and plotting, it was important to isolate the cluster attributed to DNA translocation out of unrelated events. In Figure 3.5A, since the cluster was well resolved, manual thresholding was sufficient to extract DNA cluster. A clustering algorithm would be more robust, especially, in the presence of fragments and background noise interference. As the case Figure 3.5B, there is a slight overlap between DNA cluster and background noise cluster. Therefore, DBSCAN algorithm, was employed to cluster out DNA translocation cluster

(coloured in red). DBSCAN separates data points based on density, identifying clusters without needing a preset number of clusters. It was chosen for DNA translocation data because of its effective ability to detect distinct clusters out of background noise in translocation datasets. The principle of DBSCAN operates by determining a neighborhood radius (ε) and a minimum number of points (minpts). Consequently, points have sufficient neighbors form a cluster, while isolated points are treated as noise. [26, 32, 33]

As indicated in the introduction chapter (section 1.5.2), when a DNA molecule translocates through relatively larger pore diameter (>8 nm), there is an increased chance for folded DNA translocations. Therefore, it is crucial to distinguish linear events from folded events. Given that linear events are more likely to be a successful DNA translocation. [1, 7, 35, 36]

Further, extraction of linear events was performed using PCA and k-means as detailed in Chapter 2 section 2.2.5.4. To help explain the steps performed for separating both subpopulation, Figure 3.6 (left) presents ΔI_e vs τ scatter plot for 4kbp translocation at V_{bias} = -0.8 V, after the separation of linear events (coloured as red). The eclipse presents the overall DNA cluster extracted by DBSCAN as depicted earlier in Figure 3.5. Inside the margins of the plot are two insets. The top inset presents a scatter plot for the first two principal components (PC1 vs PC2) from PCA analysis of DNA translocation events, illustrating the separation of both events subpopulations into two distinct clusters. Particularly, linear events are red coloured after k-means clustering. Moreover, the bottom inset demonstrates a scree plot displaying the percentage of explained variance, having the first two components (PC1 and PC2) account for ~85% of the overall variance. Finally, the right side of Figure 3.6 shows a representative event for each subpopulation (linear events are red coloured).

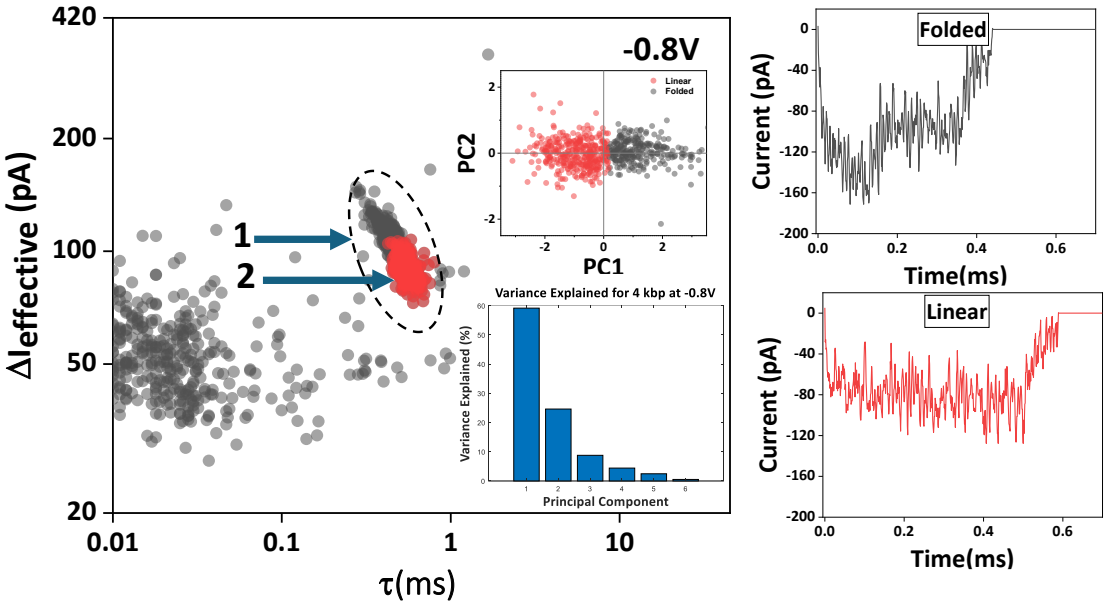

Figure 3.6. Left: ΔI_e vs. τ scatter plot for 4 kbp DNA (G_{pore}= 26.7 nS, $C_{DNA,out}$ = 300 pM at V_{bias}=-0.8 V). The overall DNA-related cluster is surrounded by a black coloured ellipse. Linear translocation events are highlighted in red after the isolation by PCA and k-means clustering. Top inset presents a scatter plot of the first two principal components, while bottom inset shows a Scree plot depicting the variance captured by each principal component, with the first two components explaining 85% of the total variance. Right: two representative events, namely folded (top, black) and linear (bottom, red).

Opposed to previous work employed Gaussian mixed model (GMM)[26], the combination of PCA and K-means as a choice for separating linear events was proposed to enable efficient and unsupervised classification while improving computational efficiency. Despite the slight overlap between PC1 and PC2, clustering by k-means has provided distinct separation for each subpopulation as shown in Figure 3.6 (left bottom and top inset). In contrast, the isolation of linear events based on current levelling as proposed by Bell would be considered a simpler approach.[19] However, the application of multiple of V_{bias} values in our experimental setup will result in variable current responses for DNA translocation events during each bias applied in each experiment. This could be a significant challenge because each individual dataset

requires manual selection of current level cutoffs. Moreover, current levelling-based separation might be influenced by user-dependent factors, introducing human bias into the classification process.

3.3.4. Forward and Reverse Translocation Characteristics for 4 kbp DNA

Starting with shorter DNA length, Figure 3.7 displays scatter plots of effective current ΔI_e vs duration τ for 4 kbp DNA translocation experiment ($C_{DNA,\ out}$ = 300 pM, G_{pore} = 26.7 nS in 4M LiCl 1xTE buffer), where V_{bias} applied sequentially starting from +0.8V to -0.6V (Figure 3.7 A-D). we have called each bias applied as a step, starting with +0.8V (step 0).

In Figure 3.7A, effective current plots at the first step (+0.8V) shows no distinct DNA cluster, given that the direction of electrophoretic force is reversed and no DNA molecules inside the pipette ($C_{DNA,\ in}$ = 0 pM). Figure 3.7B shows subsequent step in which a negative bias (V_{bias} = -0.8 V) was applied. As expected, a distinct DNA cluster with 673 events is formed roughly at τ= ~0.5 ms with ΔI_e values ranges from 100 pA to 200 pA. Further, the cluster has 370 linear events separated and highlighted in red (red, cf. Section 3.3.3). Surprisingly, when the bias was switched to +0.6 V (step 2), the DNA cluster reemerges with 283 events as plotted in Figure 3.7C (τ ~1.1 ms and ΔI_e = 100-200 pA), having 218 events identified as linear (highlighted in blue). Next, Figure 3.7D shows the results when bias again was switched to -0.6 V, resulting in the detection of 464 events (274 linear events, highlighted in green).

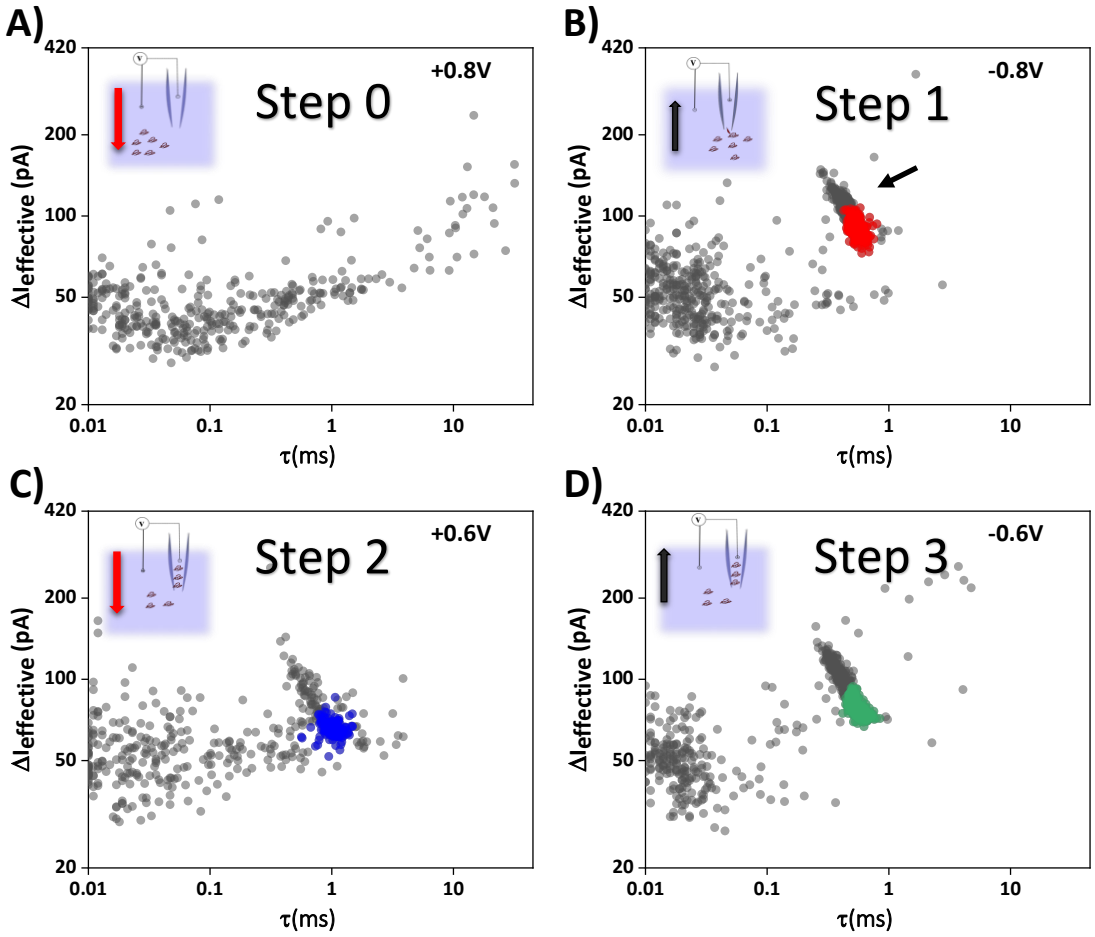

Figure 3.7 Effective current plots of detected events of 4kbp DNA (4M LiCl 1xTE, G_{pore}=26.7 nS). The sequence of bias applied is shown starting from (A)+0.8V to (D) -0.6V, the small inset represents a depiction of the translocation experiment setup with an arrow shows the direction of electrophoretic force of negatively charged species. Step 1 shows no DNA translocation, as there is no DNA present inside the nanopipette initially. linear events are colour-coded ($C_{DNA, out}$ = 300 pm).

The above-mentioned results motivated us to pursue a comprehensive analysis of DNA translocation characteristics in both directions. Translocation frequency is an important metric to evaluate dominant ions transports mechanism. Hence, when the electrophoretic force is dominating ions transport mechanism, a linear correlation is expected for the translocation frequency with the applied bias. [37, 38] Another factor is DNA concentration in the

bulk, as this increases the chance of DNA diffusing into the capture radius (refer to Chapter 1). [37-39]

First, to assess the driving force of DNA translocation it is important to recall back the factors affecting theoretical translocation frequency (f). From equation (1.11, chapter 1), considering a drift dominated regime, f is expected to scale linearly with the applied voltage, nanopore diameter and DNA concentration in the bulk. Experimentally, the translocation frequency (*f*) can be calculated by equation (3.1):

$$f = \frac{N_{evt}}{T} \tag{3.1}$$

Where, N_{evt} is the total number of detected DNA events in the DNA cluster and *T* is the total recording time at specific bias. Specifically, T=900 seconds because the first 10 files of the run were disregarded due to the transition between each step's bias.

Using equation (3.1), $f1$ =0.75s^{-1} for step 1 at V_{bias}=-0.8 V and $f2$=0.3 s^{-1} for step 2 at V_{bias}=+0.6V. Interestingly, f values were comparable at both directions although the concentration of DNA is different inside the pipette compared to the bulk. Hence, having the pipette loaded with DNA from step 1, DNA concentration inside the pipette ($C_{DNA, in}$) can be estimated by the following equation:

$$C_{DNA,in} = \frac{N_{evt}}{N_A * V_{pipette}} \tag{3.2}$$

(N_A is Avogadro number and $V_{pipettes}$ is the volume of electrolyte solution inside) Considering $V_{pipette}$ approximately 7 µL and N_{evt} from step 1 is 673 events. Accordingly, the concentration inside the pipettes ($C_{DNA,in}$) is approximately 0.2 fM.

It is important to mention that this is the overall concentration in the bulk of the pipette. Basically, the concentration inside is much lower compared to the outside, thus, the translocation frequency (*f*) is expected to be too small at the unloading step.[37] However, opposed to what is expected in a drift dominated regime, the experimentally determined f was found to increase by factor 10^6. [37] Such a phenomenon implies that the solution inside the pipette near the tip (confined region) is not well mixed. Based on this, it can be inferred that the DNA molecules enter the pipettes from the loading step, remain trapped near the tip rather than traveling to the vicinity of the pipette. Hence, the asymmetric geometry of the pipette is expected to restrict the diffusion of the DNA molecules toward the bulk of the pipette, increasing DNA local concentration.[19, 20] Another possibility could be attributed to the weaker electric field strength inside compared to the outside, resulting in reduced forces to push the DNA way from the tip (where the electric field is the strongest).

Next steps (step 4 to step 7) were performed subsequently at V_{bias} = ± 0.7 V and V_{bias} = ± 0.5 V (translocation results are presented in Appendix 2). Then, *f* values were calculated at each step and plotted against the applied V_{bias} as shown in Figure 3.8A. Initially, Figure 3.8A shows comparable *f* values as well as linear correlation with the applied V_{bias} in the forward directions (from the linear fit, r^2 = 0.99), the reverse direction has also shown linear fit (r^2 = 0.99, at V_{bias} = +0.5 V, *f* is slightly shifted due to experimental error). This is foreseen in a drift dominated ion transport system.[34, 37] The observation of a linear dependence of the translocation frequency from the trans compartment is interesting. As this supports the earlier hypothesis that the local concentration of DNA in the confined region is increased while DNA diffusion to the bulk of the pipette is restricted.

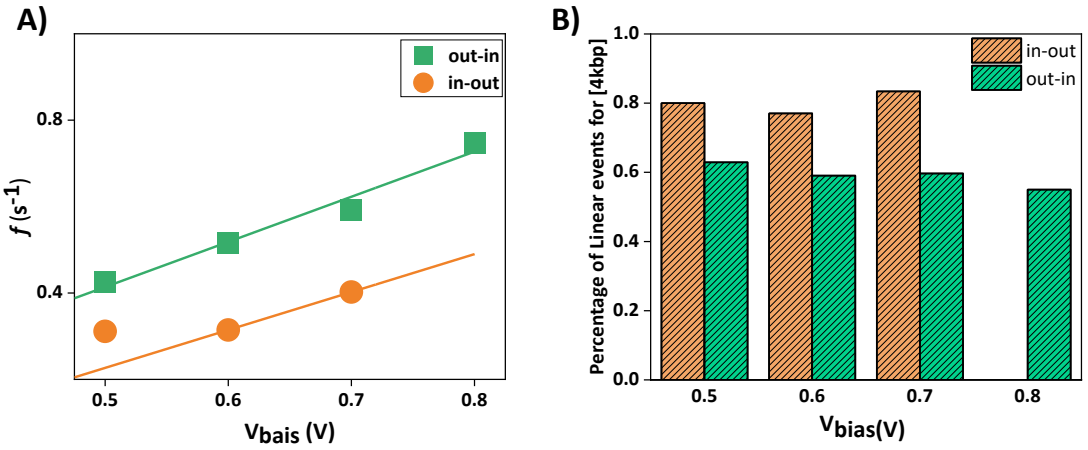

Figure 3. 8 (A) Translocation frequency (f) vs V_{bias} plot for in 4 kpb DNA (B) percentage of linear events and in both in-out (positive polarity) translocation direction and out-in direction (negative polarity), +0.8V data is not shown as it was the initial bias with no DNA inside.

In Figure 3.8B, a histogram plot presents the percentage of linear event extracted for all steps. Approximately, the linear subpopulation ratio in the loading steps was 50%-60% out of total detected events, this is in line with previous studies.[39, 40] Interestingly, a moderate increase (20%-30%) of linear events subpopulation ratio had been observed in the unloading steps. From the histogram plot, we still see a bias dependence in the forward direction. Since the experiment was performed in a sequence of different biases (for example, step 1 was -0.8 V followed by step 2 at +0.6 V), it is hard to conclude whether this moderate effect due to the application of higher bias prior to the next step or due to pore geometry. Ignoring bias sequence at this stage, the confined geometry might restrict the confirmation of DNA molecules when exist the pore. According to Bell's hypothesis, hydrodynamic forces are more significant in the reverse direction, affecting DNA confuguration.[19] Nevertheless, an advantage obtained from the reverse direction at this specific DNA length is increased linear events translocations that can be further utilized to enhance the sensing outcomes of nanopipettes. In the view of that, we have restricted our analysis to linear events in the following sections.

We suggested earlier that the DNA local concentration is high inside the pipette due to retaining DNA molecules near the tip. Thus, we wanted to identify if DNA molecule accumulation by time would modify τ in both directions. Correspondingly, the duration (τ) of linear events vs their indices were plotted for +0.6 V/-0.6 V (Figure 3.9A and 3.9B). As indicated, no transition or drift was observed owing to the stochastic nature of translocation.

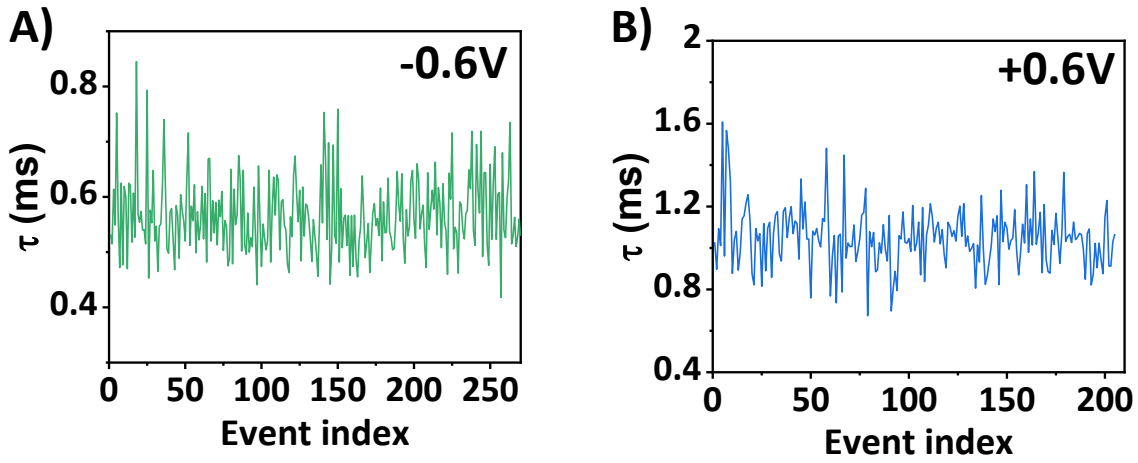

Figure 3.9(A)&(B) linear events durations vs their indices at +0.6V and -0.6V, respectively

Continuing from this point, in drift dominated regime at higher V_{bias}, τ (for linear events) is expected to scale linearly with V_{bias}^{-1}.[41] To perform this analysis, Figure 3.10A presents normal distribution plots of τ values for linear events at two illustrative V_{bias} (-0.6 V and +0.6 V). Based on this, the representative (τ_{max}) was obtained by means of gaussian distribution where $\tau_{max}=x_c$, and x_c represents the Gaussian central value derived from the distribution of τ values. Examining distribution plots in Figure 3.10A, a broader spread is observed in the reverse direction at +0.6 V with longer τ_{max} (>1ms), compared to the out-in direction at V_{bias} = -0.6V. This is an important insight in line with the previous studies, where the duration of the DNA exiting the pore is relatively longer than the duration of the DNA entering the pore.[19, 21]

As the experiment was performed at different biases, presumably, the higher the bias applied the shorter τ_{max} expected for DNA translocation.[34, 41] Figure 3.10B presents a plot of τ_{max} against V_{bias}^{-1}. As shown in the figure, a linear correlation between τ_{max} vs the inverse of bias (V_{bias}^{-1}) noticed in both directions, following the trend documented in literature.[34, 41]

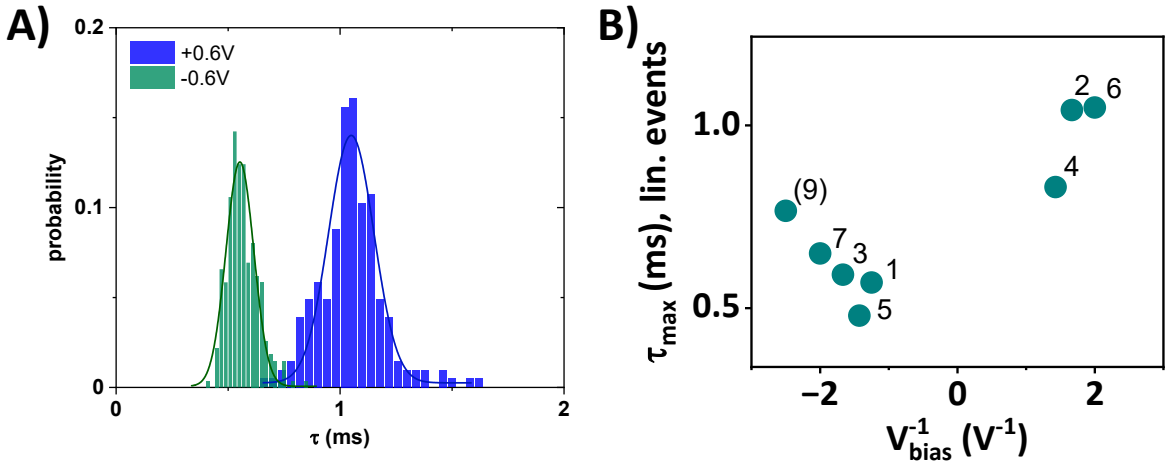

Figure 3.10 dwell time analysis of 4 kbp DNA (G_{pore} = 26.7 nS, $C_{DNA,out}$=300pM) where (A) is the normal distribution of events duration (-0.6V and +0.6V), and (B) is τ_{max} vs V_{bias}^{-1}, which appears to be linear aligning with electrophoretically driven transport.

A higher-than-expected values of translocation frequency was calculated for the reverse direction, this indeed surpass the crucial requirement of instant recapture after DNA translocation.[6, 19, 20] DNA molecules travel slower to the bulk of the pipettes due to the asymmetry this means that the nanochannel can retain a certain number of DNA molecules. Correspondingly, the number of DNA molecules remaining inside the pipette from each step bias can be estimated by a simple equation (equation 3.3):

$$N_{DNA,in} = N_{before} + S.N_{after} \quad (3.3)$$

where $N_{DNA,in}$ is the number of DNA molecules remaining inside the pipette at certain step, $N_{before}$is the number of DNA molecules was inside the pipette at the previous step and N_{after}

is the number of DNA molecules translocated in or out of the pipette. S is a variable based on each step bias polarity, where ($S = +1$) for the loading step and ($S = -1$) for the unloading step.

This equation is step dependent, and all steps were calculated in accordance with the preceding steps. As the number of DNA molecules is identified in each step and using the above-mentioned equation, Figure 3.11 represent a plot of the number of DNA molecules vs step sequence. Moving from step 1 to step 7, the number of remaining DNA molecules inside the pipettes increases as the experiment proceeds (0 molecules to 1200 molecules). The Oscillating behaviour is expected due to the fact of loading and unloading the pipettes in sequence as cycles . However, a careful visualization of the plot would suggest that the pipette is able to retain more DNA molecules with each subsequent step. This interpretation could be further referred to as pipette's capacity versus step bias. Disregarding the fact that some DNA molecules might escape to the bulk of the pipettes (as presented by the lower number of events recorded in the reverse direction), the number of remaining DNA molecules inside is increasing with time.

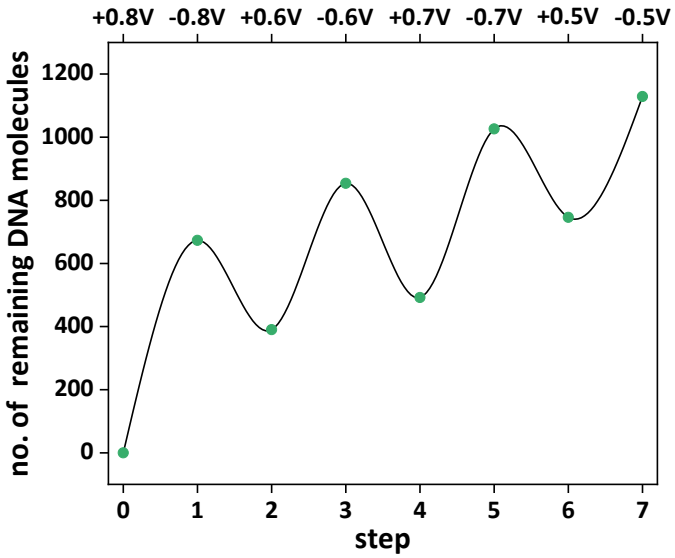

Figure 3.11 number of DNA molecules remaining inside the nanopipette at each step bias of 4kbp DNA (G_{pore} = 26.7 nS, $C_{DNA, out}$=300pM). The sinusoidal curve was plotted for the purpose of reflecting the loading and unloading steps, representing variations in the number of remaining DNA molecules inside the pipette at each step.

In another translocation experiment of 4 kbp DNA ($C_{DNA, out}$=300pM in 4M LiC1xTE, Figure 3.12 A-F), the use of a larger pipette (G_{pore} = 48.4 nS) has shown comparable results. Accordingly, τ scales linearly with V_{bias}^{-1}, and an observed DNA cluster has emerged at positive bias (*f* higher than expected). Figure 3.12F shows an increased number of DNA molecules remaining inside the nanochannel. A simple explanation would be the larger the pipettes, the more DNA molecules it can accommodate. In a drift-dominated regime, larger pores tend to produce stronger fields over larger distance in comparison to smaller pores, thereby increasing the capture radius and capture rate.[37, 38]

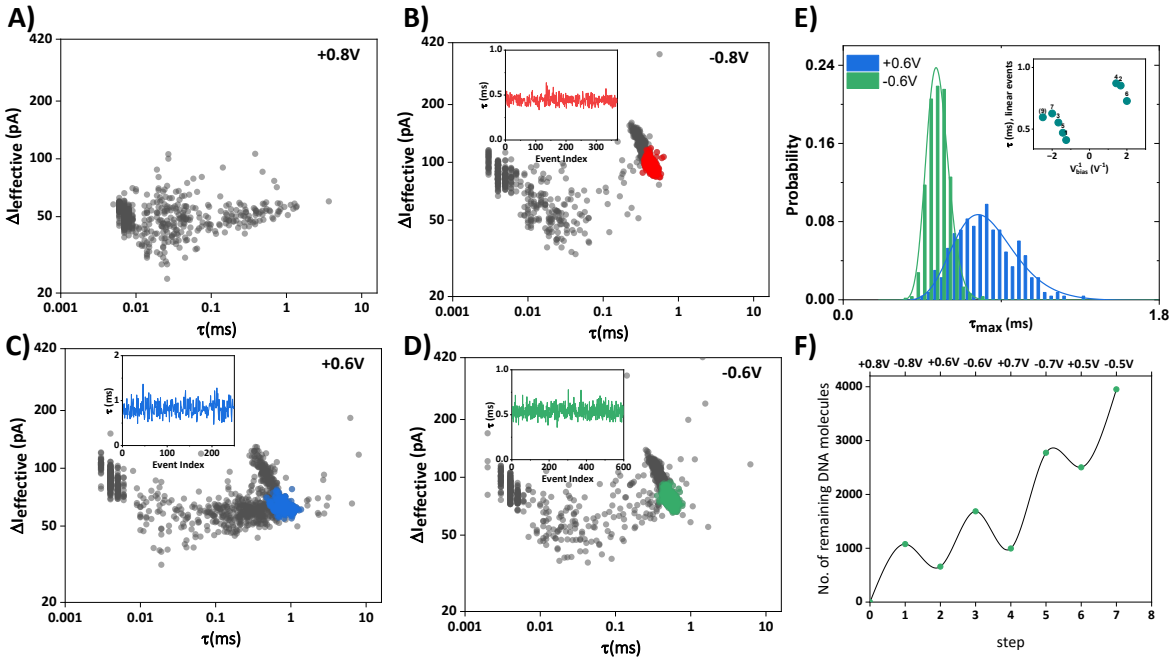

Figure 3.12 (A-D) Effective current plots of detected events of 4kbp DNA (4M LiCl 1xTE, G_{pore}=48.4 nS, $C_{DNA,out}$= 300 pM). The sequence of bias applied is shown starting from (A)+0.8 V to (D) -0.6 V, the small inset represents linear events duration vs their indices. Step 0 shows no DNA translocation, as there is no DNA present inside the nanopipette initially. Linear events are colour-coded (E) Normal distribution of -/+0.6 V, top inset is τ_{max} vs $Vbias^{-1}$ and (F) is the number of remaining DNA molecules at each step bias

3.3.5. Two Clusters Emerge in Longer DNA

After the dynamics explored in 4kbp DNA using relatively smaller and larger nanopipettes, it was more significant to initiate a study on longer DNA molecules. This is to understand how DNA length influences translocation dynamics under similar conditions. The experimental design followed the same principle of loading and unloading cycles as in 4 kbp experiments. However, the step bias sequence was modified to ±0.5V; ±0.8V; ±0.6V, and ±0.7V, respectively. This is to investigate the effect of running time vs sequence, also to help maintain a more rigorous blind approach. Figure 3.13 represents translocation data of 7kbp along with the same analysis conducted with 4kbp.

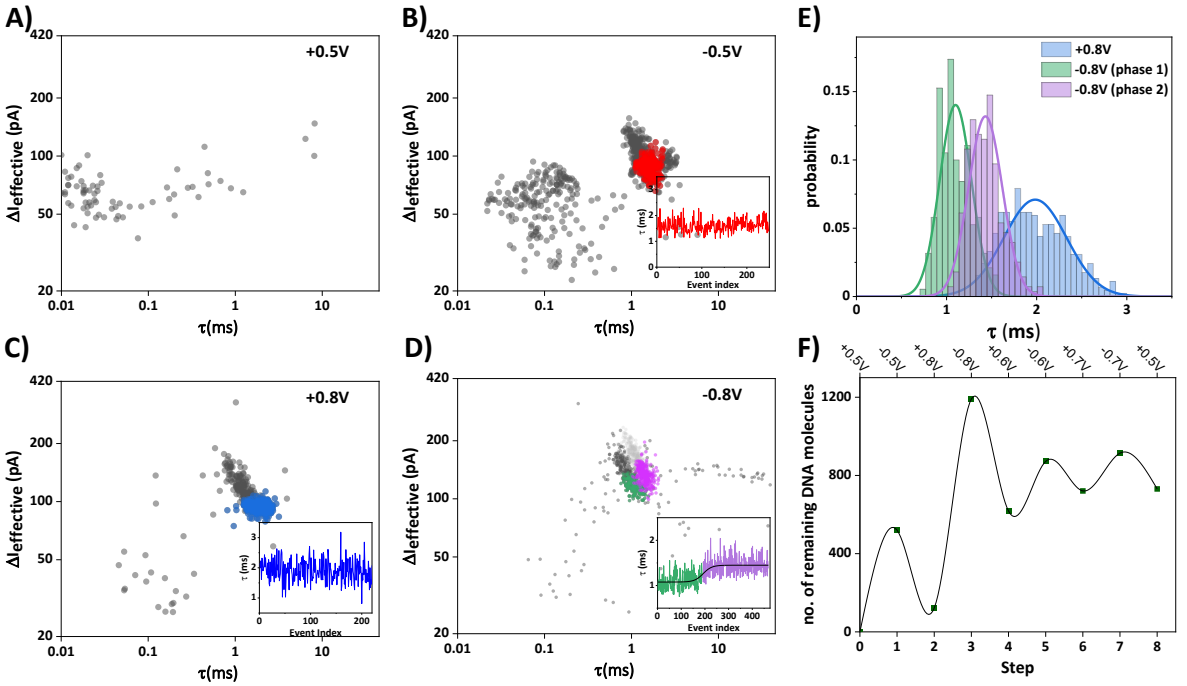

Figure 3.13 (A-D) Effective current plots of detected events of 7 kbp DNA (4M LiCl 1xTE, G_{pore}=27.5 nS, $C_{DNA,out}$ = 600 pM, overall sequence applied: ±0.5 V; ±0.8 V; ±0.6 V; ±0.7 V and additional +0.5 V at the end). The sequence of bias applied is shown starting from (A)+0.5V to (D) -0.8V, the small inset represents linear events duration vs their indices. Step 0 shows no DNA translocation, as there is no DNA present inside the nanopipette initially. Interestingly, Step 3 shows two phases at the same bias, linear events are colour-coded (E) Normal distribution of +0.8V and -0.8V (two phases) and (F) is the number of remaining DNA molecules at each step bias.

In line with previous results, no DNA translocation is expected in step 0 (Figure 3.13A) whereas a DNA cluster (Figure 3.13B) is formed in step 1 (at τ = 1 ms with ΔI_e values of 100-200 pA). Thoroughly, cluster tailing has emerged in step 1 at later stages of the applied bias (the datapoints in the tail were excluded from linear events analysis). Again, when the bias reversed to +0.8V (step 2), a distinct DNA cluster formed (at τ = 2-3 ms, ΔI_e values from 100 pA to 200 pA) after unloading the pipettes at longer τ (τ_{max} = 3 ms, obtained from the blue distribution curve in Figure 3.13E). As observed in 4kbp DNA, there was no initial drift or transition in those early steps, which is exemplified by the absent correlation between linear

events duration against their indices (bottom insets Figure 3.13B and 3.13C). Moving to step 3 at -0.8V (Figure 3.13D), an initial cluster follows comparable properties to DNA clusters of the precedent steps at τ = 1 ms and ΔI_e = 100-200 pA (phase 1). Unusually, after the translocation of certain number of DNA molecules (366 events in "phase1"), a new translocation phase emerges where a DNA cluster (phase 2) started to appear with longer τ ($\tau \sim 1.1$ ms) for linear events.

Figure 3.13D bottom inset shows the start of the drift approximately at index 193. Looking at the green line (phase 1), again no correlation between linear events and their indices. On the other hand, transitioning into phase 2 (magenta) started abruptly and a new steady phase observed with absent correlation with event index after the transition. In Figure 3.13 a gaussian distribution fit of τ is shown for steps 2 (+0.8V) and 3 (-0.8V, both phases). As seen in 4 kbp step 2, τ distribution for the reverse direction is shifted to longer durations, having broader spread (τ_{max} = 2 ms), while for step 3, $\tau_{max,1}$ = 1.1 ms in phase 1, and $\tau_{max,2}$ = 1.4 ms in phase 2.

This apparent shift in τ implies a slowing down of the translocation process for successive molecules. Accordingly, this shift could be postulated for several hypotheses. First, multiple DNA molecules translocation is expected, given that the overall estimated pore diameter was 14 nm. multiple molecules translocation expected to have increased duration and doubled intensity.[39, 42] However, the analysis of current magnitude values has disputed this hypothesis, since effective current values were close in both phases ($\Delta I_{e,phase2} < 2\ \Delta I_{e,\ phase1}$) as demonstrated by the gaussian distribution plot in Figure 3.14.

The second hypothesis suggests that the accumulation of DNA molecules inside near the tip could alter the electrolyte solution conditions near the sensing region in terms of viscosity or electric properties. Concurrently, the remaining DNA molecules may exert additional frictional force on subsequent molecules entering the pipette.

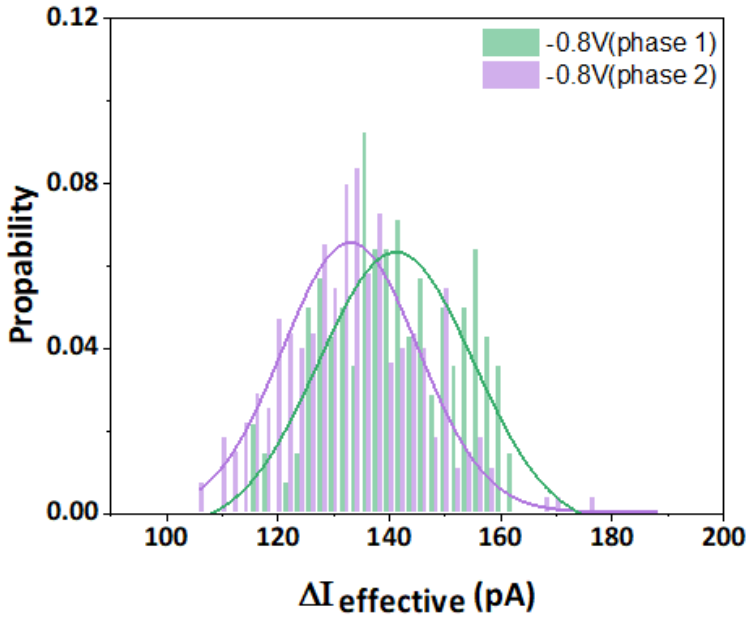

Figure 3.14 normal distribution plots of effective current values of DNA linear events of phase1 and phase 2 at -0.8V, the values were extracted for linear events and colour coded in line with phases shown in Figure 3.13D

In Figure 3.13F, the number of DNA molecules remaining inside the nanopipette were plotted against each step bias. As presented in 4kbp translocation data (Figure 3.9D and Figure 3.10F), the number of remaining DNA molecules inside the pipettes increased in accordance with a drift dominated regime at the first two steps.[37, 38] This continued until it reached a steady point, reducing the transport of DNA molecules into the pipette at step3. In this regard, we estimated *f* values using equation (3.1), and plotted *f* against V_{bias} as displayed in Figure 3.15A. Notably, there was no positive linear correlation between *f* and the applied bias for the out-

in direction (Figure 3.15A), rather f decreased relatively with time (f data for the in-out direction are presumably dependent on the previous step). In Figure 3.15B, where τ_{max} plotted vs V_{bias}^{-1}, τ_{max} was affected when the translocation experiment progressed as represented by the non-linear correlation with V_{bias}^{-1}. Based on previous hypothesis, after the translocation of a certain number of DNA molecules into the pipette, the pipette's nanochannel properties change. At this point, pipette capacity reaches a point where the DNA transport from the bulk is impacted. This is likely due to either residual negative charge buildup or alteration of solution viscosity which impacts the transport of additional molecules. However, increased residual negative charges would be transformed into changes in the electrical properties of the nanochannel, as well as a gradual increase in duration over the course of this effect (outlined in section 3.3.6).[15] Since the transition happened abruptly and then stabilized, increased solution's viscosity and frictional force hypothesis is more convincing. This can be inferred due to gradual "crowding" of DNA molecules near the tip, as escaping into the bulk of the pipette is limited by the slow diffusion. Such an effect can be attributed to the asymmetry of the pipette and the increased length of the DNA molecules ($(R_H \alpha D^{-1})$ in equation (1.9), Chapter 1).Therefore, the remaining DNA molecules would induce frictional force on subsequent DNA molecules slowing down their translocation time (τ).[43]

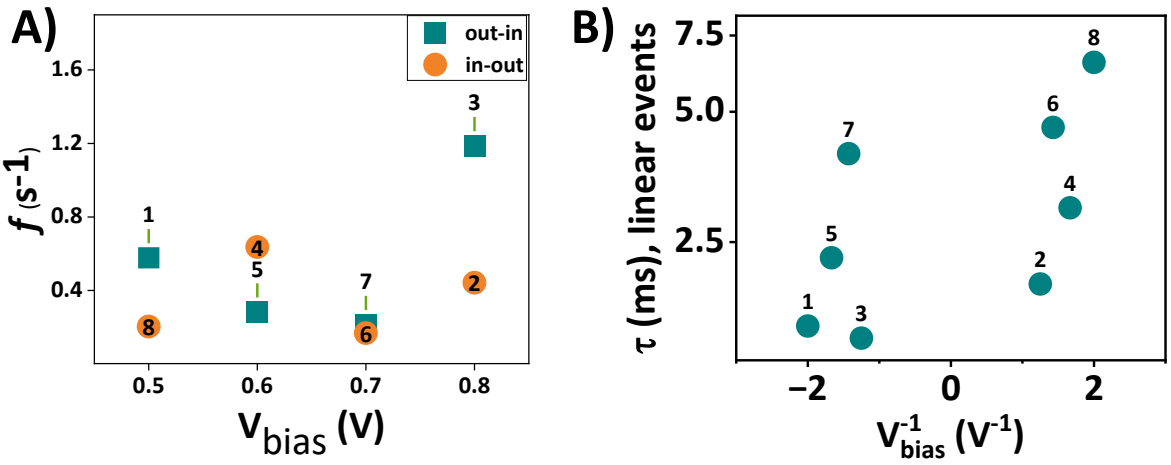

Figure 3.15 (A) Translocation frequency (f) for 7 kbp experiment out-in steps (negative polarity, dark cyan squares) and in-out steps (positive polarity, orange circles), points were labelled with each step number as stated earlier(vertical projections for negative polarity and inside the circles for positive polarity), clearly, no linear dependence in both plots.(B) τ_{max} vs V_{bias}^{-1} for 7kbp translocation experiment, labelled with step sequence for each.

Finally, DNA dynamics were analyzed (phase 2 cluster was excluded from the analysis). It was worth noting that there was less influence of the direction of the translocation on the number of linear DNA events captured. Only 5-10% increase of linear events in the reverse direction as depicted in Figure 3.16. This could be complex to explain as we have DNA length and bias as additional factors affecting DNA confirmation during translocation.[44]

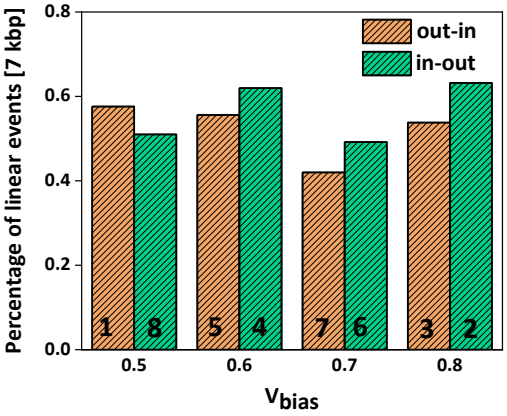

Figure 3.16 Percentage of linear events in 7kpb DNA in both in-out (positive polarity) translocation direction and out-in direction (negative polarity), black numbers represent the sequence of the bias applied

To explore the observed crowding state, the study was expanded to include longer DNA molecules. The same experimental setup with shorter lengths was used to perform the translocation experiment. Hence, 10 kbp DNA was performed as a single experiment ($C_{DNA,out}$ = 600pM, G_{pore}= 36.9 nS). Two experiments were performed using 48.5 kbp DNA at two different concentrations: 1) $C_{DNA,\ out}$ = 300 pM and G_{pore} = 31.6 nS, and 2) $C_{DNA\ ,out}$ = 39 pM, G_{pore}= 44.2 nS. The bias sequence applied in 10kbp DNA experiment was +0.5 V; -0.8 V, ±0.6 V; ±0.7 V, and for 48.5 kbp experiment the sequence was ±0.5 V; ±0.8 V; ±0.6 V; ±0.7 V.

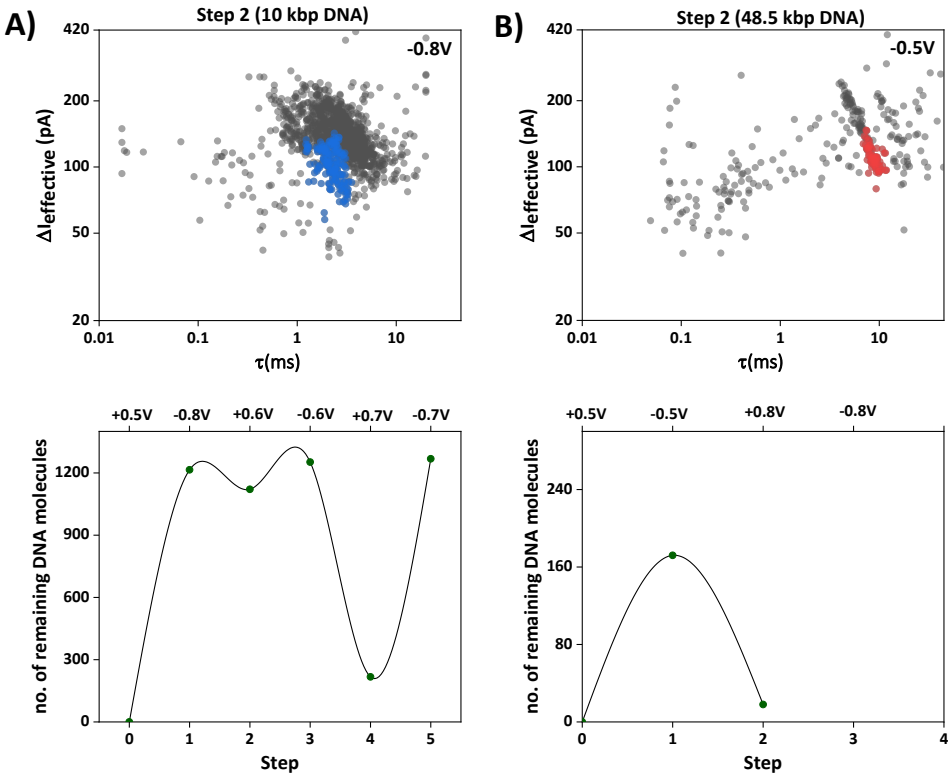

Figure 3.17 Effective current plots at step 2 where the two phases emerge along with the plot of the number of remaining DNA molecules inside after the overall run of (A) 10kbp DNA (4M LiCl 1xTE, G_{pore}=36.9 nS, $C_{DNA,\ out}$ = 600 pM), And (B) 48.5 kbp DNA (4M LiCl 1xTE, G_{pore}=31.6nS, $C_{DNA,out}$ = 300 pM), (the subsequent steps were noisy with inability to find DNA associated cluster, this is further explained next in section 3.3.6), linear events are colour coded

In Figure 3.17A and 3.17B, we show the results of effective current plots at the step where the crowding effect occurred in 10 kbp DNA ($C_{DNA\ ,out}$ = 600 pM, G_{pore}= 36.9 nS) and 48.5 kbp DNA($C_{DNA,\ out}$ = 300 pM, G_{pore} = 31.6 nS), respectively.

Notably, the two cluster phases were also identified with longer DNA lengths at relatively comparable concentrations (τ = ~5 ms for 10 kbp and τ = ~10 ms for 48.5 kbp). Furthermore, crowding effect occurred faster with a smaller number of DNA molecules required for the emergence of phase 2 cluster transition. Following the same trend observed in 7 kbp experiment, phase 1 transitioned into phase 2 reaching a steady state directly after the

transition. From the number of remaining DNA molecules plot presented in Figure 3.17A (bottom) and 3.17B (bottom), the transition occurred after approximately 349 molecules of 10kbp DNA and 124 molecules of 48.5kbp. Subsequently, the pipette capacity was limited for next DNA translocations. The next steps biases in both experiments were not shown because it was harder to find DNA clusters as compromised by pore blockage. This blockage induced a drastic change in noise level, this phenomenon was further investigated next (see section 3.3.6). From the results of the abovementioned experiments, we can conclude that the longer the DNA, the smaller the number of DNA molecules needed for the transition from phase 1 into phase 2. Considering that longer DNA molecules have greater hydrodynamic volume in comparison to smaller lengths (chapter 1 equation (1.9) and equation (1.10)), this will indeed mean that longer DNA molecules will occupy more volume in the pipette when trapped in the nanochannel, reducing the available volume for the next forward DNA translocations.

In this regard we have also analysed the translocation frequency (*f)* values to compare it with literature for related DNA lengths. Accordingly, It was documented that (*f*) scales positively with DNA lengths from 0.5 kbp to 10 kbp DNA in 4M LiCl in conical nanopores, which suggest an entropic factor effect.[38] On the other hand, we have found that translocation frequency reduced with 10kbp (f=0.6s^{-1}, T=2000s, -0.8V) and 48.5kbp (f=0.2s^{-1},-0.5V) at high concentrations in comparison with shorter lengths. Since our lowest applied bias was -0.5V for 48.5 kbp DNA (300 pM), When higher voltage applied the pore tended to clog at the first file hindering obtaining proper cluster to resolve events.[45, 46]

To investigate the effect of concentration regardless the length, the experiment for 48.5 kbp was performed at low concentration ($C_{DNA,out}$ = 39 pM). Figure 3.18 represent effective current plots for 48.5 kbp experiment at low concentration ($C_{DNA.\ out}$=39pM and G_{pore}=44.2nS). The

sequence is shown starting from top-left to bottom-right. Particularly, a single cluster was formed at approximately τ= ~10 ms for the loading steps and at τ= ~40 ms for the unloading steps.

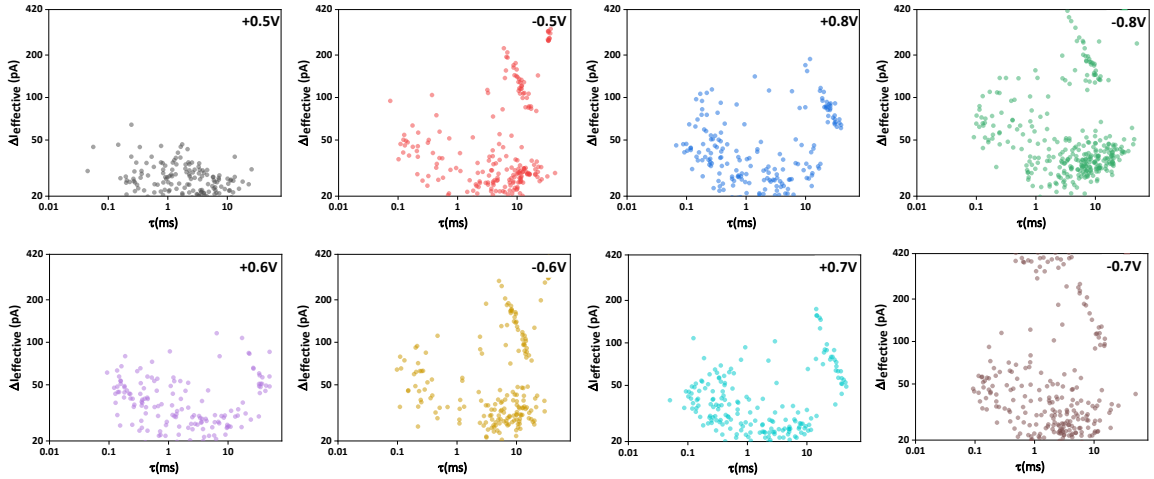

Figure 3.18 Effective current plots of detected events of 48.5kbp DNA (4M LiCl 1xTE, $C_{DNA.\ out}$=39pM and G_{pore}=44.2nS). The sequence of bias applied is shown from left to right, starting from +0.5V (left-top) to -0.7V(right-bottom). Step 1 shows no DNA translocation, as there is no DNA present inside the nanopipette initially.

Interestingly, no phase transition observed in the overall experiment after a prolonged experiment duration. This would suggest that the number of remaining molecules does not reach the cutoff value to cause this effect, and the pipette is still able to accommodate more DNA molecules. In contrast to higher DNA concentration experiment, all applied bias worked despite the small number of events. We have again plotted the number of remaining DNA molecules in the nanochannel against step bias as illustrated in Figure 3.19. we found that 48.5 kbp at low concentration followed a similar pattern with 4kbp data. Furthermore, the number increased in every step without reaching the plateau observed at higher DNA concentration experiment for 48.5 kbp DNA. Hence, it was easier to use remaining DNA molecules versus step plots to define the transition point (the crowding effect).

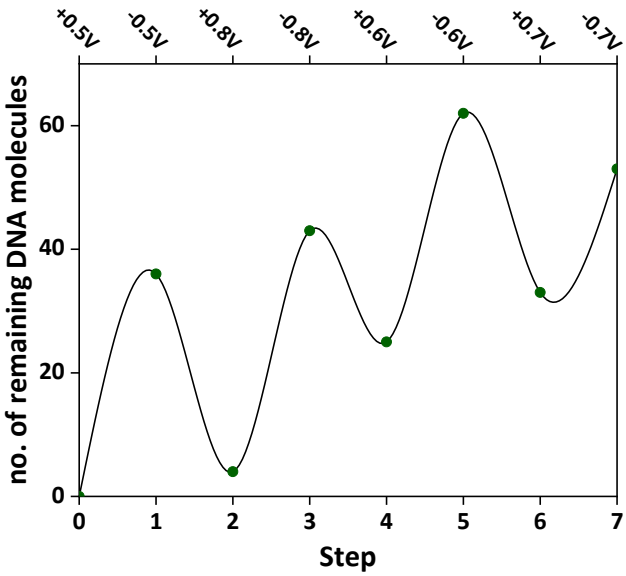

Figure 3.19 Plot of the number of remaining DNA molecules inside after the overall run of 48.5kbp DNA (4M LiCl 1xTE, G_{pore}=44.2nS, $C_{DNA,out}$ = 39pM)

Toward this end, we have evaluated three important factors controlling phase 2 transition while the DNA remain trapped inside the nanochannel, including DNA length, Concentration, and $G_{pore..}$ We observed that longer DNA tend to reach this phase faster at relatively comparable concentrations compared to shorter lengths.

Coming back to DNA configuration, we have a range of DNA lengths and mixed V_{bias} polarities. In this aspect, we wanted to analyse data carefully to develop conclusion about the most controlling factor for DNA configuration. This is to get overall picture about the influence of the direction on linear events ratio. Therefore, we have calculated the average percentage ratio of linear events in the reverse direction to the average percentage of linear events in the forward direction for all DNA lengths. In Figure 3.20A, this ratio plotted against the applied V_{bias}. Interestingly, the applied bias was a dominant factor in controlling the confirmation of the DNA. Further, we have found that the percentage linear events scales positively with V_{bias}

in the reverse direction. Conversely, a negative correlation has been noticed in the forward direction (out-in).[39, 42]

To further contextualize our findings, we performed an additional analysis on the mean ΔI_e Although not central to this study, this aligns with prior reports[21] and provides complementary insights. According to Chen's previous study using nanopipettes, the direction of translocation was observed to affect the level of ionic current slightly at a high salt concentration. This effect is also influenced by both DNA length and the bias applied.[21] As shown in Figure 3.20B, a scatter plot of the mean value of ΔI_e vs V_{bias} , where the red coloured data points present the mean value of ΔI_e for the forward direction and the blue datapoints present values for the reverse direction. For DNA length of 4 kbp to 7 kbp, the level of current response slightly decreases in the reverse direction for each DNA length at the same V_{bias}. This slight decrease in the reverse direction has been recently attributed to the competition between electrophoretic force and hydrodynamic drag, which is more prominent in the reverse direction.[47] However, in 10 kbp the response was slightly increased (48.5kbp data were not shown as we have only two mixed bias values), in contrast to what documented above.

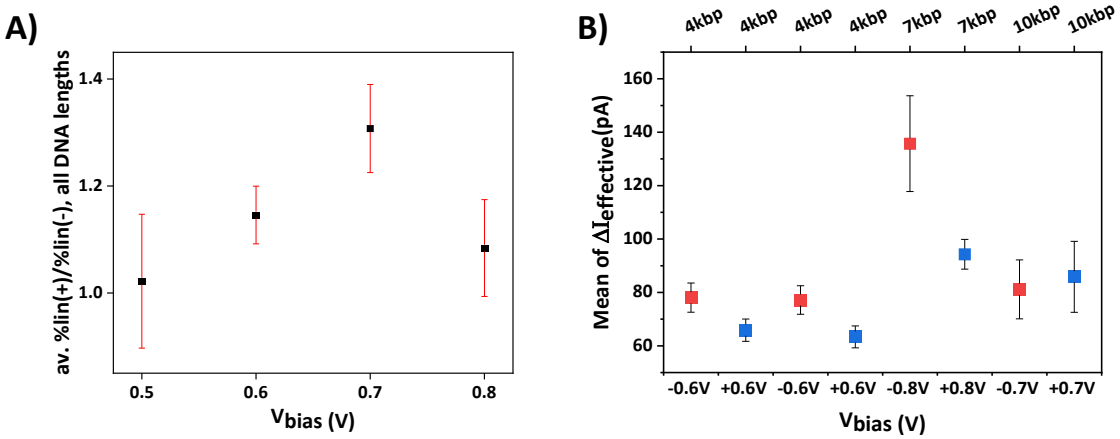

Figure 3.20 (A) Correlation plot of ratio of the average percentage of linear events in the in-out direction vs the average ratio of percentage of linear event in the out-in direction vs V_{bias} applied, the error bars represent standard error of the mean of overall number of measurement (B) the mean effective current of linear events in both directions (red for out-in and blue for in-out) at the same bias for each DNA length, error bars represent standard deviation of the gaussian fit of ionic current level distributions.

3.3.6. Electrical Properties of The Nanopipettes Before and After the Transition

DNA is a charged analyte along with the charged surface of the nanopipettes, there might be a possibility for the DNA to interact with the surface of nanopipettes causing clogging [46]or smoothly block the nanopipette, this could modify the electric and temporal response of the translocation signal for the successive translocating DNA molecules.[48, 49] Specifically, when there's a potential of pore clogging, a change in pore conductance might be traced.[45, 48-51] Accordingly, the observed phase transition could be associated by DNA clogging event or surface interaction rather than crowding near tip. However, pore blockage is a random process which happens at any time during translocation, and in most cases, it is irreversible which impacts pore's operational lifespan.[45] Coming back to the amplifier (see chapter2

section 2.2.5), the signals were recorded with two main outputs: the baseline current: "DC" channel which is filtered at 7 Hz and the "AC" channel where fast fluctuations associated with translocation events is recorded. Since DC channel presents baseline current(I_B), we can calculate back the conductance of the pore (G_{pore})

Further, when the pore is blocked partially or the crowded DNA molecules accumulated near the tip, this might modify the electric properties of the sensing region and impact the steady flow of ions toward the pore. In this case, the level of AC sigma noise would be expected to increase[49] (a zero mean channel will make the AC sigma noise and RMS noise values comparable, where $I_{RMS}=\sqrt{\overline{\Delta I^2(t)}}$, ΔI^2 is the current values deviated from the mean of the trace).

Figure 3.21 shows the analysis of both G_{pore} (top) and AC noise (bottom) for all DNA lengths employed in this work. The columns on the left represent values before the transition (phase 1), while the column on the right present values after the transition into phase 2. The dashed columns represent additional phase for 48.5 kbp DNA, which is discussed later. From the figure, Analysis of both AC noise and G_{pore} for both clusters have revealed no significant change with either minimal decrease or increase in the conductance and AC noise by 0.2%-0.8%, respectively. This supports our hypothesis that cluster transition is likely due to DNA accumulation near the tip, increasing viscous drag and slowing translocation instead of altering the electric response.

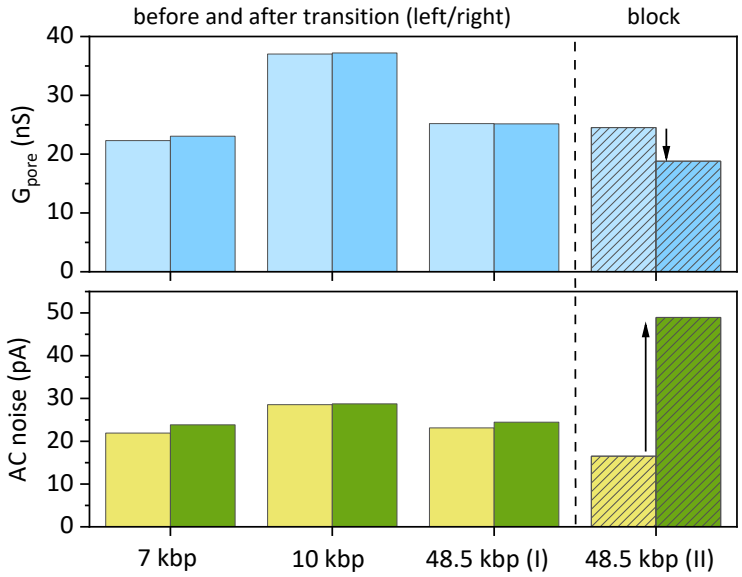

Figure 3.21 Analysis of pore conductance (G_{pore}) of the DC channel in the top and standard deviation(sigma noise) of the AC channel below of the 2 cluster phases before and after the emergence for longer DNA lengths (starting from 7, 10 to 48.5kbp DNA), the dashed line separates subsequent state of an imminent pore block for 48.5kbp (I)).

Another interesting transition also occurred with longer DNA molecules (10kbp and 48.5kbp), where we identify a statistically significant change in both AC noise and DC conductance for 48.5kbp DNA as presented in the last column plot in Figure 3.21 (for ease of presentation 10kbp data for blockage was not presented and further analysis shown below). To explore the characteristics of this new phase transition (phase 3), we have further examined the electrical properties before and after the emergence of this transition.

Figure 3.22 details the analysis of 10 kbp DNA ($C_{DNA,\ out}$ = 300 pM) and Figure 3.23 details the analysis for 48.5 kbp DNA ($C_{DNA,\ out}$ = 300 pM). From Figure 3.22 A and Figure 3.23B, the effective current plot for 10 kbp reveals a well-defined DNA cluster at approximately τ=0.9-1.7 ms, having ΔIe ranges from 80 pA to 150 pA. In 48.5 kbp, DNA cluster emerges at $\tau \approx 4$ ms to 11 ms and $\Delta Ie \approx$ 100 pA to 200 pA (this cluster defines phase 1). Transition into phase 2

cluster starts from τ=2 to 3.2 ms for 10 kbp and from τ =10 to 20 ms for 48.5 kbp at comparable ΔIe values. From panel C in Figure 3.22 and 3.23, we have not observed any changes in G_{pore} [DC conductance] and AC noise at the end of phase 1 in 10 kbp experiment. However, in 48.5 kbp G_{pore} slightly increased by up to 3% and the level of noise doubled to approximately to 22 pA. In line with the analysis illustrated in Figure 3.21 (see above), phase 2 transition did not exhibit any changes in AC noise values or G_{pore}. Delving deeply into the third phase in panel C, we have found that AC channel noise increased 4 times in 10kbp DNA and three times in 48.5kbp DNA, additionally G_{pore} was reduced by 10% for 10kbp and 23% in 48.5kbp. In terms of this point, we have then further analyzed the traces recorded after the transition into phase 3.

Examining the results shown in Figure 3.22B and 3.23B, no DNA translocation detected after the noise level increased. In the plot of both AC noise and DC conductance (Figure 3.22C and 3.23C), phase 3 transition started at file index 544 for 10kbp (after approximately 5440 seconds, considering each file was recorded for 10s and -0.6V step started at file index 533). While for 48.5 kbp, the transition started at file index 132 (after 1320 seconds of total run in which step2 at -0.5V started at file 115). Also, when the traces were analysed after this block an initial spike of noise was detected along a significant change in the DC channel, then the baseline stabilised without additional detected spikes in the trace (at around 98pA for 10kbp and 55pA for 48.5kbp).

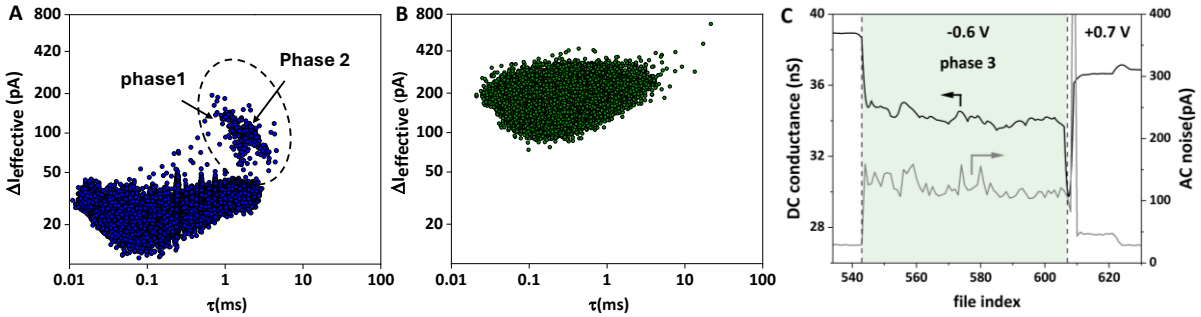

Figure 3.22 (A) and (B) are effective current plots for 10 kbp DNA, for the both phase1(cluster at 1.7 ms) and phase 2(cluster at 2-3.2 ms) translocation cluster, and phase 3 transition of the experiment (V_{bias} = -0.6V, event detection threshold of 1.5σ) (C) pore conductance(G_{pore})and AC channel sigma noise, as a function of file index . green region in C, represents the status when pore was blocked, and no further translocation events could be detected (as shown in B). unhighlighted region presents the values when bias was reversed to +0.7V, where pore conductance recovered, and AC noise stabilised back to the optimal values.

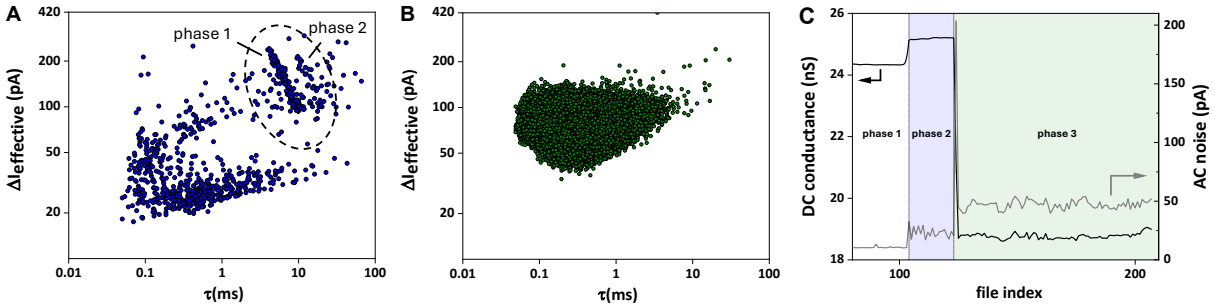

Figure 3.23 (A) and (B) are effective current plots for 48.5 kbp DNA, for the both phase1(τ ~ 4 ms) and phase 2 (τ ~ 10 ms)translocation cluster, and phase 3 transition of the experiment (V_{bias} = -0.5 V, event detection threshold of 3σ) (C) pore conductance (G_{pore}) and AC channel sigma noise, as a function of file index. Phase 2 (highlighted in blue) shows the change for the two clusters of translocation events (as appeared in A). green region in C, represents the status when pore was blocked, and no further translocation events could be detected (as shown in B).

Interpretation of the results shows that this blockage happened earlier in 48.5kbp accompanied by no events detected after the change happened. Importantly, a change in G_{pore} and AC noise is considered an alteration of the sensing region of the nanopipettes thus this could be due to pore block.[49] A clogging event associated with G_{pore} decrease is not the case because the pipette remained conductive with only 10-23% block. Hence, an event cluster

should be formed as a dimed cluster in the noise, however, analysis of the data at lower threshold has no DNA cluster formed. It was also interesting that the pipette remained conductive, and ion current flow continued. This could support the hypothesis of DNA accumulating near the tip and after certain number of molecules the pipette was filled up to the maximum capacity (densely populated) in which no more DNA molecules translocated (remaining number of DNA inside the pipette required for blockage phase was 1121 and 174 molecules for 10kbp and 48.5kbp, respectively). An additional challenge could be that the detected current is probably due leakage of current through quartz.[52] Notably, phase 3 continued during the application of negative polarity (the loading step of the pipettes). Contrariwise, when the V_{bias} reversed to the unload phase, the pipette recovered, and the noise level stabilized back to the original value. Figure 3.24 shows scatter plots of effective current for the reversed bias directly after the blockage phase for both 10kbp (Figure 3.24A) and 48.5kbp (Figure3.24B), a cluster related to DNA formed with a magnitude of 100pA at 4-9 ms for 10kbp and at 10-14 ms for 48.5kbp. The reversibility of the process further disproves that the change in pore conductance was due to pore clogging.

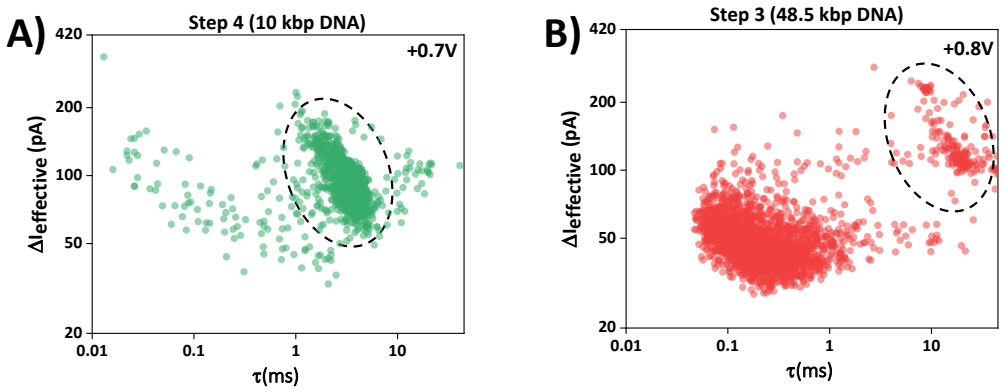

Figure 3.24 Effective current plots of detected events of the reversed bias (unloading phase) after the blockage phase for (A)10kbp DNA at +0.7V(4M LiCl 1xTE, G_{pore}=36.9 nS) and (B) 48.5kbp DNA at V_{bias} = +0.8 V (4M LiCl 1xTE, G_{pore}=31.6 nS)

It was worth noting that for 10kbp DNA the pipette remained working for next step applied V_{bias}. In 48.5kbp, the case was different because the lowest applied V_{bias} was -0.5V and the application of higher V_{bias} for the following steps values has let the pipette reach the blockage phase before event capture and mainly during the switch between applied voltage.

3.3.7. Remaining DNA Molecules and Time Shift

According to the above-mentioned results, phase 2 cluster emergence has no significant change in AC noise or DC conductance. Further, we have also observed that this transition occurs faster with longer DNA molecules. This prompted us to systematically examine the relationship between DNA length and the number of DNA molecules present in the nanopipette. The number of remaining DNA molecules inside the pipette can be determined just prior to time shift, this is to define how many DNA molecules would cause this transition or crowding as explained in section 3.3.5. In Figure 3.25A, the number of remaining DNA molecules pre-transition was plotted against DNA length, and the numerical label represents G_{pore} for each data point in the plot. The datapoint associated with the number of DNA molecules remaining for 4kbp experiments were left as an open circle, instead, the remaining number of DNA molecules from last the step V_{bias} were plotted, since phase 2 never reached experimentally.

Obviously from the plot, the number of DNA molecules decreases vs DNA length. Further observation, the plot follows a reciprocal cubic function which suggests the presence of scaling volume factor in understanding the actual relationship. In this aspect, ignoring that DNA molecules traveling to the vicinity of the nanopipette, the asymmetric pipette can be assumed as a holder with an effective volume of the nanochannel (V_{ch}) and the DNA with certain length

has a volume (V_{DNA}). Thus, we can obtain a relationship between the V_{ch} and V_{DNA} by equation (3.4):

$$N_{Max} = \frac{V_{Ch}}{V_{DNA}} \tag{3.4}$$

(Where N_{Max} represent the maximum number that a pipette can accommodate before the emergence of the transition)

In general assumption for modelling DNA volume in a confined space, we can proportionally assume that V_{DNA} α (DNA Length)^3, therefore N_{Max} is linearly proportional to $1/$(DNA Length)^3.[53, 54] This simplification assumes a finite volume for DNA in solution and is used for order-of-magnitude comparisons. While not physically accurate in terms of shape, this approach provides a reasonable upper-bound estimate for our analysis.

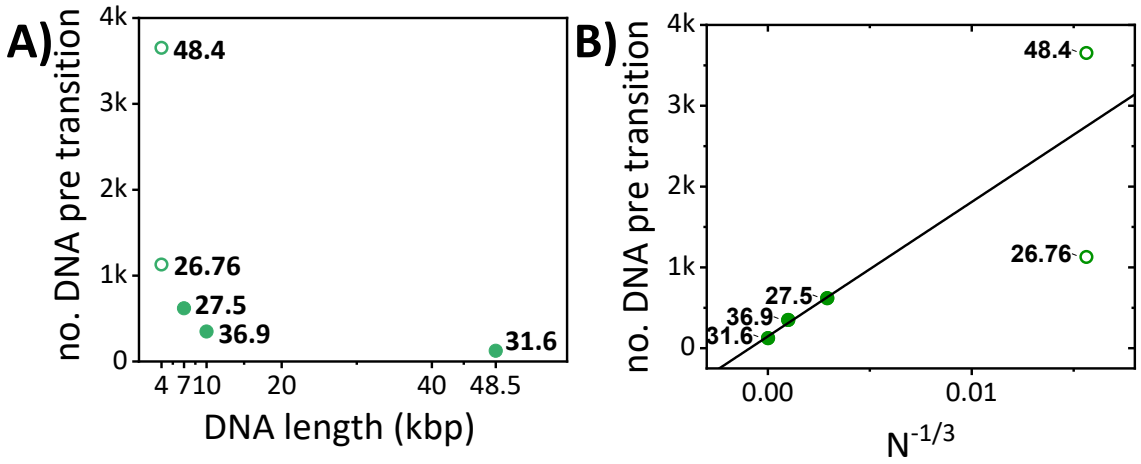

Figure 3.25 plot of number of DNA molecules just pre-transition vs (A)Length of DNA(kbp) and (B) 1/(DNA Length)^3 for all DNA lengths used, Linear relationship indicates DNA volume as factor for filling the asymmetric volume of nanopipettes. Text labels are G_{pore} values in nS for each experiment.

As plotted in Figure 3.25B, a linear relationship is obtained considering the 4kbp experiments which did not reach the actual maximum number experimentally. Based on the model, in the

4kbp experiment with G_{pore}=26.7nS, this number was approximated around 3000 molecules. Since the pipette is larger for the 48.4nS experiment, it is expected to accommodate higher number of molecules to reach that phase (not shown because it is beyond the scale, larger pores tend to have larger capture radius as well increased capture rate).[34]

3.4. Conclusion

In this chapter, the asymmetric nanopipettes were explored systematically for their potential to trap and release DNA molecules. This was demonstrated by loading the pipettes using typical translocation experiment in which DNA molecules is electropherically driven from the outside "cis" compartment to the inside "trans "compartment for an extended duration (100 seconds) before being released for an equivalent period. The study included a different ranges of DNA lengths starting from 4kbp to 48.5kbp, all were injected at picomolar concentration into the bulk solution of 4M LiCl without adding DNA molecules into the trans compartment. Due to geometrical constraint, translocation frequency from the trans side has comparable value to that in the cis side, opposing to what expected. Additionally, folding dynamics tend to relatively increase linear subpopulation of released molecules with less influence on longer DNA molecules. Accordingly, it has been hypothesized that DNA molecules remain trapped near the tip rather than escaping into the vicinity of the nanopipettes. This effect was more prominent in longer DNA≥7kbp supported by small diffusion constant, in which for the first time after loading the pipettes with certain number of molecules, a new translocation state of the cis side characterised by delayed dwell time appearing due to crowding of DNA molecules near the tip. The electrical properties of this state in terms of pore conductance and sigma noise were alike to the precedent molecules, hence, increased frictional force was

suggested. Further, the increased length of DNA was associated with lower number needed for this transition. Interestingly, in 10kbp and 48.5kbp, another translocation state marked by partial decrease in pore conductance and increase in AC noise, where the pipettes were unable to accommodate more DNA molecules, and no events were detected at that state. Recovery from this state was proceeded after the start of the unload step of the pipette. A future study could involve, loading the pipette and pause the translocation rather than instant release in addition to propose a new method to image the structure of the DNA inside the pipettes, this might further explore the properties of DNA inside a confined space. A simpler suggestion since DNA remain trapped near the tip, an analyte could be placed inside and a real time capture can be approached by first translocating a modified DNA with capturing probes for the analyte of interest inside, then compare the signal of both directions once the DNA released.

3.5. References

1. T. Albrecht, *Annu Rev Anal Chem (Palo Alto Calif)*, 2019, **12**, 371-387.
2. Y. L. Ying, Z. L. Hu, S. Zhang, Y. Qing, A. Fragasso, G. Maglia, A. Meller, H. Bayley, C. Dekker and Y. T. Long, *Nat Nanotechnol*, 2022, **17**, 1136-1146.
3. L. Xue, H. Yamazaki, R. Ren, M. Wanunu, A. P. Ivanov and J. B. Edel, *Nat Rev Mater*, 2020, **5**, 931-951.
4. K. Lee, K. B. Park, H. J. Kim, J. S. Yu, H. Chae, H. M. Kim and K. B. Kim, *Adv. Mater.*, 2018, **30**, e1704680.
5. L. M. Ying, A. Bruckbauer, A. M. Rothery, Y. E. Korchev and D. Klenerman, *Anal. Chem.*, 2002, **74**, 1380-1385.
6. M. Gershow and J. A. Golovchenko, *Nat Nanotechnol*, 2007, **2**, 775-779.
7. C. Plesa, L. Cornelissen, M. W. Tuijtel and C. Dekker, *Nanotechnology*, 2013, **24**, 475101.
8. X. Liu, P. Zimny, Y. Zhang, A. Rana, R. Nagel, W. Reisner and W. B. Dunbar, *Small*, 2020, **16**, e1905379.
9. A. Rand, P. Zimny, R. Nagel, C. Telang, J. Mollison, A. Bruns, E. Leff, W. W. Reisner and W. B. Dunbar, *ACS Nano*, 2022, **16**, 5258-5273.
10. Y. Zhang, X. Liu, Y. Zhao, J. K. Yu, W. Reisner and W. B. Dunbar, *Small*, 2018, **14**, e1801890.
11. X. Liu, Y. Zhang, R. Nagel, W. Reisner and W. B. Dunbar, *Small*, 2019, **15**, e1901704.
12. D. Stein, *Nat Nanotechnol*, 2007, **2**, 741-742.
13. J. Stanley and N. Pourmand, *APL Materials*, 2020, **8**.
14. D. Perry, D. Momotenko, R. A. Lazenby, M. Kang and P. R. Unwin, *Anal. Chem.*, 2016, **88**, 5523-5530.
15. A. Y. Y. Loh, C. H. Burgess, D. A. Tanase, G. Ferrari, M. A. McLachlan, A. E. G. Cass and T. Albrecht, *Anal. Chem.*, 2018, **90**, 14063-14071.
16. R. J. Yu, Y. L. Ying, Y. X. Hu, R. Gao and Y. T. Long, *Anal. Chem.*, 2017, **89**, 8203-8206.
17. F. Bošković, C. Maffeo, G. Patiño-Guillén, R. Tivony, A. Aksimentiev and U. F. Keyser, *bioRxiv*, 2023, 2023.2009. 2012.557357.
18. A. P. Ivanov, P. Actis, P. Jonsson, D. Klenerman, Y. Korchev and J. B. Edel, *ACS Nano*,

2015, **9**, 3587-3595.

19. N. A. W. Bell, K. Chen, S. Ghosal, M. Ricci and U. F. Keyser, *Nat Commun*, 2017, **8**, 380.
20. F. Zheng, M. Alawami, J. Zhu, C. M. Platnich, J. Sha, U. F. Keyser and K. Chen, *Nano Lett.*, 2023, **23**, 11145-11151.
21. K. Chen, N. A. W. Bell, J. Kong, Y. Tian and U. F. Keyser, *Biophys. J.*, 2017, **112**, 674-682.
22. A. Choudhary, H. Joshi, H. Y. Chou, K. Sarthak, J. Wilson, C. Maffeo and A. Aksimentiev, *ACS Nano*, 2020, **14**, 15566-15576.
23. M. Muthukumar, *J. Chem. Phys.*, 2010, **132**, 195101.
24. Z. Mi, X. Zhao, X. Chen, Y. Wang, X. Shan, K. Chen and X. Lu, *ACS Nano*, 2024, **18**, 23243-23252.
25. R. L. Fraccari, M. Carminati, G. Piantanida, T. Leontidou, G. Ferrari and T. Albrecht, *Faraday Discuss.*, 2016, **193**, 459-470.
26. R. L. Fraccari, P. Ciccarella, A. Bahrami, M. Carminati, G. Ferrari and T. Albrecht, *Nanoscale*, 2016, **8**, 7604-7611.
27. L. J. Steinbock, A. Lucas, O. Otto and U. F. Keyser, *Electrophoresis*, 2012, **33**, 3480-3487.
28. Z. Siwy, E. Heins, C. C. Harrell, P. Kohli and C. R. Martin, *J. Am. Chem. Soc.*, 2004, **126**, 10850-10851.
29. P. Actis, A. C. Mak and N. Pourmand, *Bioanal Rev*, 2010, **1**, 177-185.
30. R. M. Kuanaeva, A. N. Vaneev, P. V. Gorelkin and A. S. Erofeev, *Biosensors (Basel)*, 2024, **14**.
31. S. W. Kowalczyk, D. B. Wells, A. Aksimentiev and C. Dekker, *Nano Lett.*, 2012, **12**, 1038-1044.
32. B. Lu, F. Albertorio, D. P. Hoogerheide and J. A. Golovchenko, *Biophys. J.*, 2011, **101**, 70-79.
33. C. Plesa and C. Dekker, *Nanotechnology*, 2015, **26**, 084003.
34. M. Wanunu, J. Sutin, B. McNally, A. Chow and A. Meller, *Biophys. J.*, 2008, **95**, 4716-4725.
35. N. A. Bell and U. F. Keyser, *Nat Nanotechnol*, 2016, **11**, 645-651.
36. C. Plesa, N. Van Loo, P. Ketterer, H. Dietz and C. Dekker, *Nano Lett.*, 2015, **15**, 732-737.
37. M. Wanunu, W. Morrison, Y. Rabin, A. Y. Grosberg and A. Meller, *Nat Nanotechnol*,

2010, **5**, 160-165.

38. N. A. W. Bell, M. Muthukumar and U. F. Keyser, *Phys. Rev. Lett.*, 2016, **93**, 022401.
39. L. J. Steinbock, O. Otto, C. Chimerel, J. Gornall and U. F. Keyser, *Nano Lett.*, 2010, **10**, 2493-2497.
40. N. Ermann, N. Hanikel, V. Wang, K. Chen, N. E. Weckman and U. F. Keyser, *J. Chem. Phys.*, 2018, **149**, 163311.
41. D. Y. Ling and X. S. Ling, *J Phys Condens Matter*, 2013, **25**, 375102.
42. M. Mihovilovic, N. Hagerty and D. Stein, *Phys. Rev. Lett.*, 2013, **110**, 028102.
43. D. Fologea, J. Uplinger, B. Thomas, D. S. McNabb and J. Li, *Nano Lett.*, 2005, **5**, 1734-1737.
44. A. J. Storm, J. H. Chen, H. W. Zandbergen and C. Dekker, *Phys Rev E*, 2005, **71**, 051903.
45. G. Ando, C. Hyun, J. Li and T. Mitsui, *ACS Nano*, 2012, **6**, 10090-10097.
46. O. M. Eggenberger, C. Ying and M. Mayer, *Nanoscale*, 2019, **11**, 19636-19657.
47. F. Zheng and Q. Han, *Analyst*, 2024, **149**, 5131-5138.
48. Z. X. Liang, Z. P. Tang, J. Li, R. Hu, D. P. Yu and Q. Zhao, *Nanoscale*, 2015, **7**, 10752-10759.
49. S. F. Liang, F. B. Xiang, Z. F. Tang, R. Nouri, X. D. He, M. Dong and W. H. Guan, *Nanotech Precis Eng*, 2020, **3**, 9-17.
50. S. Li, S. Zeng, C. Wen, L. Barbe, M. Tenje, Z. Zhang, K. Hjort and S. L. Zhang, *J. Phys. Chem. B*, 2020, **124**, 11573-11583.
51. S. Gravelle, R. R. Netz and L. Bocquet, *Nano Lett.*, 2019, **19**, 7265-7272.
52. J. Velmurugan, D. P. Zhan and M. V. Mirkin, *Nat Chem*, 2010, **2**, 498-502.
53. D. E. Charlton, H. Nikjoo and J. L. Humm, *Int. J. Radiat Biol.*, 1989, **56**, 1-19.
54. A. E. Nkodo, J. M. Garnier, B. Tinland, H. Ren, C. Desruisseaux, L. C. McCormick, G. Drouin and G. W. Slater, *Electrophoresis*, 2001, **22**, 2424-2432.

CHAPTER 4
NANOPIPETTES AS AN INCUBATION TOOL

CONTENTS

Synopsis: This chapter explores a new analyte capture method that utilizes the benefit of accumulation of DNA molecules inside the nanochannel to capture weakly charged analytes.[1] Hence, an analyte of interest is incubated inside the nanopipette at lower concentration and then captured by reverse translocation. The process involved incubating streptavidin functionalised gold nanoparticles inside the pipettes. Then, biotin functionalised 10 kbp DNA in the bulk solution was translocated into the pore and allowed to incubate with the streptavidin functionalised AuNP inside the nanopipette. Finally, the reverse translocation events of the functionalised 10 kbp DNA were analysed to assess successful capture of the target.

4.1. Experimental Objective

This chapter explores the application of nanopipettes as a multifunctional tool, leveraging their confined geometry to serve as a reservoir for nanoparticles, as well as an incubation platform to facilitate binding with functionalized 10 kbp DNA. Streptavidin-functionalized gold nanoparticles (SA-AuNP) were loaded into the nanopipettes, benefiting from the crowding phenomena discussed in Chapter 3. Thus, crowding was hypothesized to enhance interactions and capture through concentrating on the nanoparticles in the confined region of nanopipette. To assess the capture mechanism, biotinylated 10 kbp carrier DNA was introduced into the bulk solution surrounding the nanopipettes. Then, the carrier DNA was electrophonically driven into the nanopipettes, enabling interactions with the SA-AuNP. The aim of this chapter is to evaluate a new analyte capture method benefiting from the asymmetry of the nanopipettes in promoting capturing analytes through spatial crowding. The biotinylation of the DNA enables efficient binding to streptavidin-functionalized gold nanoparticles. This binding will be further studied next in an in-situ capture strategy involving DNA translocation through nanopipettes. Further, biotin and streptavidin functionalization ensure precise and controlled interaction between the DNA and the gold nanoparticles. Additionally, the nanoparticles were chosen as a model analyte due to their tailored physicochemical properties, which minimize the risk of sticking and clogging to the nanopipettes' surface.

4.2. Results and Discussion

4.2.1. Design and Characterization of 5'-Biotinlyted DNA

Biotinylated 10 kb DNA fragment was generated through PCR amplification, employing lambda DNA as a PCR template. Further, the reverse primer was functionalized at the end with a 5' biotin group that is necessary for binding with streptavidin. The length of DNA fragment was verified through agarose gel electrophoresis (refer to Chapter 2.1.1). Figure 4.1A presents gel electrophoresis results of PCR reaction, lane 1 contained a 1 kb DNA ladder, while lane 2 and lane3 displayed the PCR product and lambda DNA, respectively. The PCR product band position in lane 2 aligns approximately with the expected migration level of 10 kbp DNA, confirming the successful amplification of PCR construct.

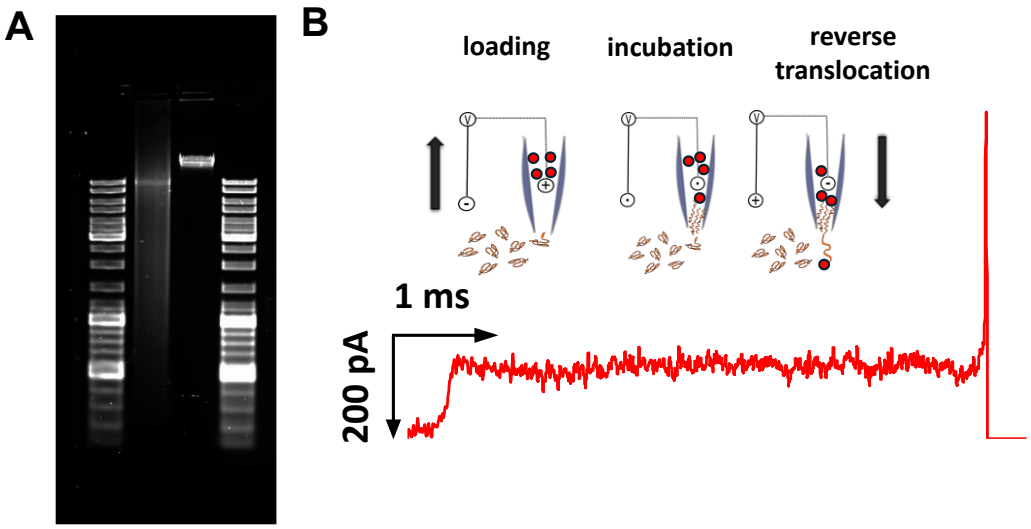

Figure 4.1 (A) agarose gel electrophoresis of PCR amplification of the biotin-functionalized 10kbp construct. (1% agarose in 1xTAE, running parameters: 80V for 49 minutes) DNA samples are loaded from left to right where Lane 1, Lane 2 and Lane 3 presents 1kb DNA ladder, 10kbp DNA fragment and Lambda DNA, respectively (B) depiction of capture strategy showing a SA-AuNP nanoparticles loaded into the nanopipette while functionalised DNA injected to the bulk solution.

Finally, the biotinylated 10 kb DNA construct's concentration was determined using UV-Vis spectroscopy and found to be 600 ng/µL using (refer to Chapter2, Section 2.1.2).

Figure 4.1B illustrates the capture strategy, in which biotin is essential for the subsequent binding of the PCR fragment to streptavidin-functionalized gold nanoparticles (SA-AuNP). To further illustrate the strategy, the figure first shows the loading step (left, top). SA-AuNP nanoparticles were preloaded inside the nanopipette, while functionalized DNA injected to the bulk solution outside. Once the bias was switched on (negative bias), DNA starts to enter the nanopipette. The next step incorporates incubation with the particles after sufficiently loading the pipette with functionalized DNA (incubation, depicted in the middle). After incubation, bias was reversed, and bound construct exits the pore (reverse translocation, right). At the bottom of Figure 4.1B, a representation of linear event obtained after reverse translocation. Hence, an intense subevent spike occurs approximately at the end of the carrier, confirming successful binding. The schematic in Figure 4.1B summarizes the detection principle for smaller particles assuming nanopipette's pore diameter < 20 nm. Since 40 nm gold nanoparticles are larger than the pore size, passing through the smaller pore is physically constrained. Accordingly, DNA-particle bound construct in case of large particles, the interaction with the pore was found to produce complex and intense current responses, as further discussed in the following sections.

4.2.2. Reverse Translocation of 40 nm SA-AuNP

Prior to introducing functionalized DNA into the bulk solution, it was essential to assess the translocation characteristics of functionalised nanoparticles in nanopipettes. Thus, a control experiment for 40 nm SA-AuNP was conducted to evaluate the electrical properties of the

translocation signals. Hence, 40 nm SA-AuNP were loaded into a nanopipette ($C_{SA\text{-}AuNP,\ in}$ = 500 pM, G_{pore} = 26.1 nS in 4M LiCl 1xTE buffer) while the bulk solution consisted of only 4M LiCl 1xTE buffer. Figure 4.2 displays the translocation results of 40 nm streptavidin-functionalized spherical gold nanoparticles at V_{bias} = ±0.7 V. The top inset depicts nanoparticle movement with directional arrows indicating the assumed translocation direction. The schematic suggests that negatively charged DNA molecules migrate in the opposite direction to the applied electric field. Panel A shows clustered data after the application of V_{bias} = +0.7 V, presenting effective current (top) and maximum current (bottom) as functions of τ. Panel B displays corresponding results at V_{bias} = -0.7 V. Finally, bottom insets in both Panel A and Panel B shows representative current-time traces, providing additional context regarding the signal characteristics under each applied V_{bias}.

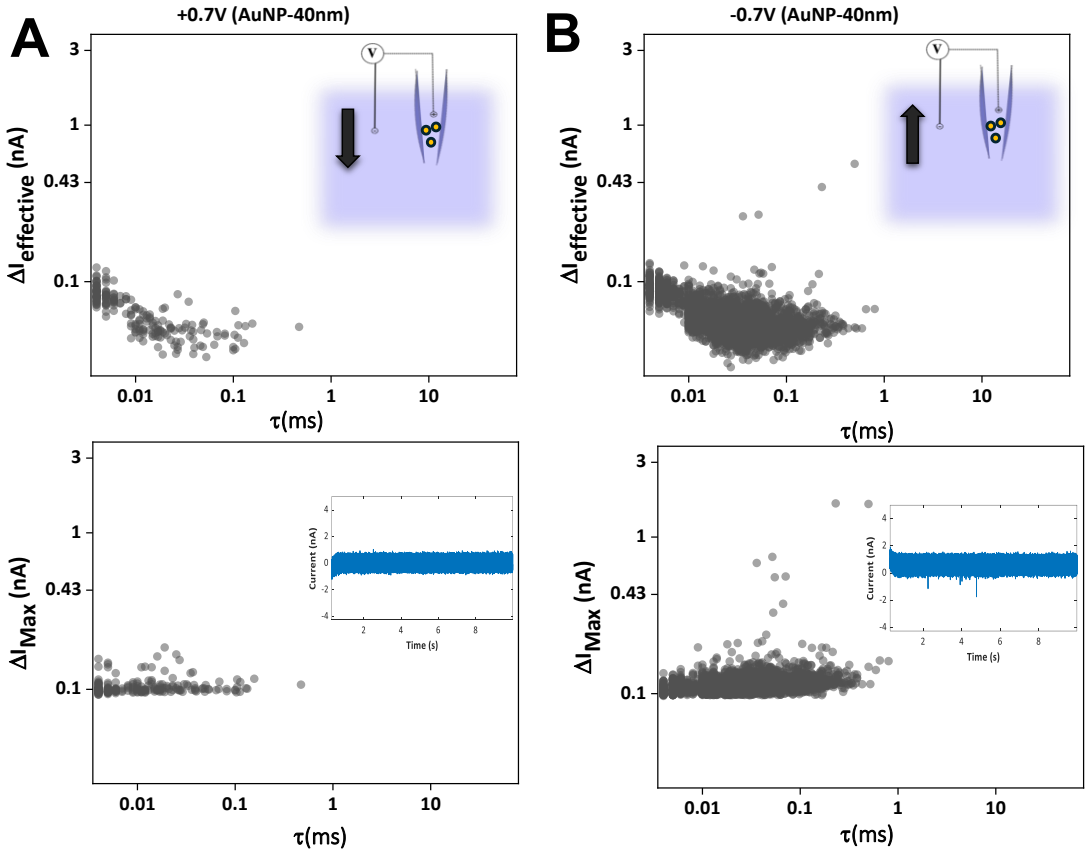

Figure 4.2 Translocation results of 40 nm SA-AuNP ($C_{SA\text{-}AuNP,\ in}$ = 500 pM, G_{pore} = 26.1 nS in 4M LiCl 1xTE buffer), (A) effective current ΔI_e vs τ plots (top) and maximum current ΔI_{Max} vs τ plot (bottom) of detected events at V_{bias} = +0.7 V. (B) effective current ΔI_e vs τ plots (top) and maximum current ΔI_{Max} vs τ plot (bottom) of detected events at V_{bias} = - 0.7 V. Top insets are depictions of experimental setup with a directional arrow representing the direction of electrophoretic force assuming DNA movement. Bottom insets are representative current time traces (10 s) recorded at each applied V_{bias}.

At V_{bias} = +0.7 V, effective current $\Delta I_e vs\ \tau$ plots does not show a distinct cluster, while the mostly irregular and scattered cluster represents background noise (Panel A, top). This suggests the absence of significant translocation events. In contrast, at V_{bias} = -0.7 V, $\Delta I_e vs\ \tau$ plot shows low-intensity cluster with ΔI_e ranging from 30 pA to 100 pA, at τ = 10-150 µs. Across both conditions, the lack of intense and consistent signals highlights the challenges in detecting free nanoparticle translocation through the nanopipette. It is important when study

nanoparticles to consider all forces governs the movement of particles. Accordingly, there is a contribution of both electrophoretic force (EP) and electroosmotic flow (EOF). In our case, the 40 nm SA-AuNP possess a weak positive charge (+10 mV, as provided by the manufacturer). Thus, the direction of both forces will be additive. Since the particles loaded inside the pipettes, at positive bias (+0.7 V), the particles are expected to move away from the sensing region. This translated by low number of detected events as shown in $\Delta I_e vs\ \tau$ plot (Panel A, top). However, at negative bias (-0.7 V), the particles are expected to move toward the sensing region (Panel B, top). This has been observed as diffuse cluster with low intensity ($\Delta I_e =$ 30 pA to100 pA) at relatively short durations (τ = 10-150 μs). Since higher voltage applied, the combination of both EP and EOF force might increase the translocation velocity limiting their detection by our instrumental setup.[2-5] For instance, the translocation rate for a DNA single base pair in this system is approximately 0.28 μs/base pair, implying a translocation time of ~3.74 μs for a 40 nm diameter nanoparticle. Such short durations, coupled with our limited amplifier bandwidth, likely limit the resolution of signals lasting less than 50 μs.[6, 7] Another aspect worth discussing is that there appears to be a mismatch between pore size and SA-AuNP nanoparticle diameter. In this experiment, SA-AuNP has diameter of 40 nm which exceeds the estimated pore diameter of the nanopipette used (13 nm). Thus, translocation of the particles through the pipette is physically not possible. Large nanoparticles will become trapped near the pore and away from the sensing region. Consequently, incomplete translocation of the particles could result in weak current responses accompanied by a short duration.[5, 8] Given that the sample has a heterogeneous size distribution, low intensity short lived pulses may also be generated by smaller particles, thereby further complicating accurate detection.[9]

4.2.3. Detection of 40 nm SA-AuNP Inside by Reverse Translocation

As mentioned in Chapter 1, the design of a carrier DNA with modified probes has been proposed as an elegant solution for capturing rapidly translocating analytes. This approach has been demonstrated to be effective in capturing smaller analytes, eliminating the requirement of precise tailoring of the pore geometry to the analyte's size.[10] Using the same pipette from the control experiment of the free nanoparticles (G_{pore}=26.1 nS, estimated d_p = 13 nm), we proceeded to the incubation experiment. Thus, biotinylated 10kbp DNA was introduced into the bulk solution ($C_{DNA,\ out}$ =300 pM, in 4M LiCl 10%TE buffer), to capture the 40 nm SA-AuNP already placed inside the nanopipettes ($C_{SA\text{-}AuNP,\ in}$ = 500 pM, in 4M LiCl 10%TE buffer). The experimental procedure involved a series of applied V_{bias} to drive DNA translocation into the pipette and allow incubation with the functionalized nanoparticles (cf. Chapter 2, section 2.2.5).

Figure 4.3, Panels A and B, display the translocation data obtained during both directions at V_{bias} = ±0.8 V. The top plots show ΔI_e vs τ, reflecting the overall detected events. The effective current plots reveal the formation of DNA clusters (blue and green) at approximately 1 ms for V_{bias} = -0.8 V and at 3-10 ms for V_{bias} = +0.8 V (colored as blue at V_{bias} = -0.8 V and as green at V_{bias} = +0.8 V). Both color coded clusters were extracted using DBSCAN (cf. Chapter 2, Section 2.2.5.3 and Chapter 3, section 3.3.3). The bottom plots present the maximum current plots (ΔI_{Max} vs τ) for DNA related cluster. To assess successful capture of functionalised nanoparticles, we need first to analyse the translocation results at each step, loading at V_{bias} = -0.8 V and unloading step at V_{bias} = +0.8 V. Hence, successful loading step can be confirmed first by the formation of DNA cluster at V_{bias} = -0.8 V. From panel A (top), a well-defined DNA

related cluster emerges at τ = 1-2 ms with a broad range of ΔI_e values ranging from 100 pA 450 pA, confirming DNA entry.

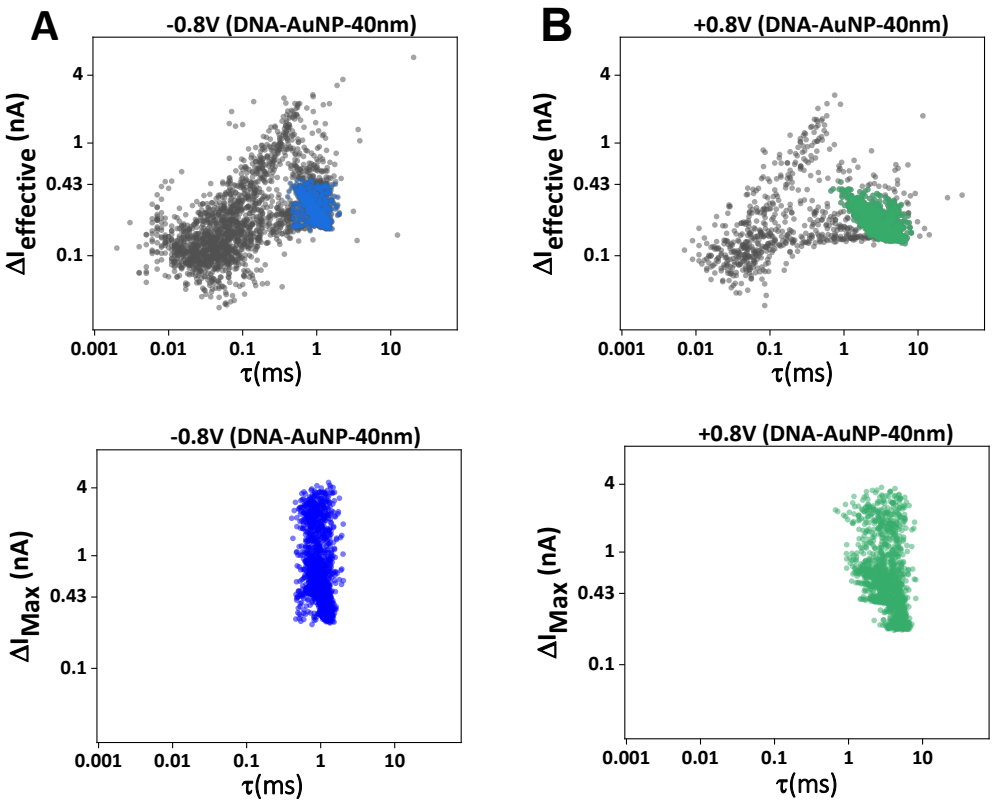

Figure 4.3 Translocation results of 10 kbp DNA incubation with 40 nm SA-AuNP experiment ($C_{DNA,\ out}$ = 300 pM, $C_{SA\text{-}AuNP,\ in}$ = 500 pM, G_{pore} = 26.1 nS in 4M LiCl 1xTE buffer), (A) effective current ΔI_e vs τ plots of overall events detected (top) and maximum current ΔI_{Max} vs τ plot for DNA related cluster (bottom) at V_{bias} = -0.8 V (B) effective current ΔI_e vs τ plots of overall events detected (top) and maximum current ΔI_{Max} vs τ plot for DNA related cluster (bottom) at V_{bias} = +0.8 V.

After the incubation with the preloaded nanoparticles inside, a new cluster related to DNA translocation should be formed after the bias was reversed to V_{bias} = +0.8 V. As shown in Panel B (top), DNA related cluster emerges at τ = 3-10 ms with ΔI_e = 100-430 pA. ΔI_{Max} vs τ scatter plots (shown in bottom plots), demonstrate the occurrence of intense current spikes, ranging from 0.5 nA to 4 nA.

For further comparison, Figure 4.4 shows plots of maximum current response normalized to the open pore [DC] current ($\Delta I_{Max}/I_{DC}$ (%)) for unmodified 10 kbp DNA (Panel A) and the incubation experiment (Panel B) at V_{bias} = +0.7 V. To account for different pipettes used, we have normalized the current response to the [DC] open pore current (cf. Chapter 2, section 2.2.5). This normalized percentage is referred to as "average blockages." Hence, the average blockage in the unmodified 10 kbp DNA experiment was less than 2%. Conversely, the average blockages in the incubation experiment have a broader range (1%-25%). In Panel C, a gaussian distribution fit was applied to the log values of the average blockages for both experiments. For unmodified 10 kbp DNA (red), the lognormal distribution has log-space values: μ=−0.0838 and σ =0.11, corresponding to μ = 0.824 ± 0.238, when transformed to normal values. Further, the distribution shows a positive skew, where most DNA-induced blockages cluster around 0.824%. This is likely corresponding to linear DNA translocation, assuming DNA has an average diameter ~1-2 nm. On the other hand, the lognormal distribution fit of the average blockages in the incubation experiment demonstrates positive skewness with a longer tail to the right (μ=−0.429 and σ = 0.0672). This tailing is associated with either a strong pore interaction or blockages by relatively larger molecules, likely a DNA-particle complex translocation.

Dwell time (τ) histograms for both experiments were also fitted into lognormal distribution (Panel D). The mean τ (μ = 5.01 ms) for the incubation experiment is longer than the mean τ of unmodified DNA (μ = 3.16 ms). Increased dwell time in the incubation experiment could be attributed to the smaller size of the pipette used in the experiment.[8] In terms of spreading, the nanoparticle incubation experiment exhibits narrower fit ($\sigma = 0.13$ and relative standard deviation (RSD(%))=13.6) in comparison with unmodified DNA experiment ($\sigma = 0.17$, RSD(%)=18.73). Accordingly, the translocation process was more uniform in the incubation

experiment compared with the unmodified DNA. The distribution of dwell time of unmodified DNA is broader, implying less uniform process and a higher crowded state assumed.[1, 11] In section 4.2.2, we have discussed that the EOF force of particles are in the direction of the electric field. Which may likely reduce the effect of crowding and enhance the translocation process, assuming similarity of other factors including thermal motion and geometrical constraints.[1]

In conclusion, ΔI_{Max} values in particle experiments significantly exceed the anticipated responses typically associated with standard DNA translocation events under similar experimental conditions. Suggesting relatively successful AuNP capture, particularly after unloading DNA.

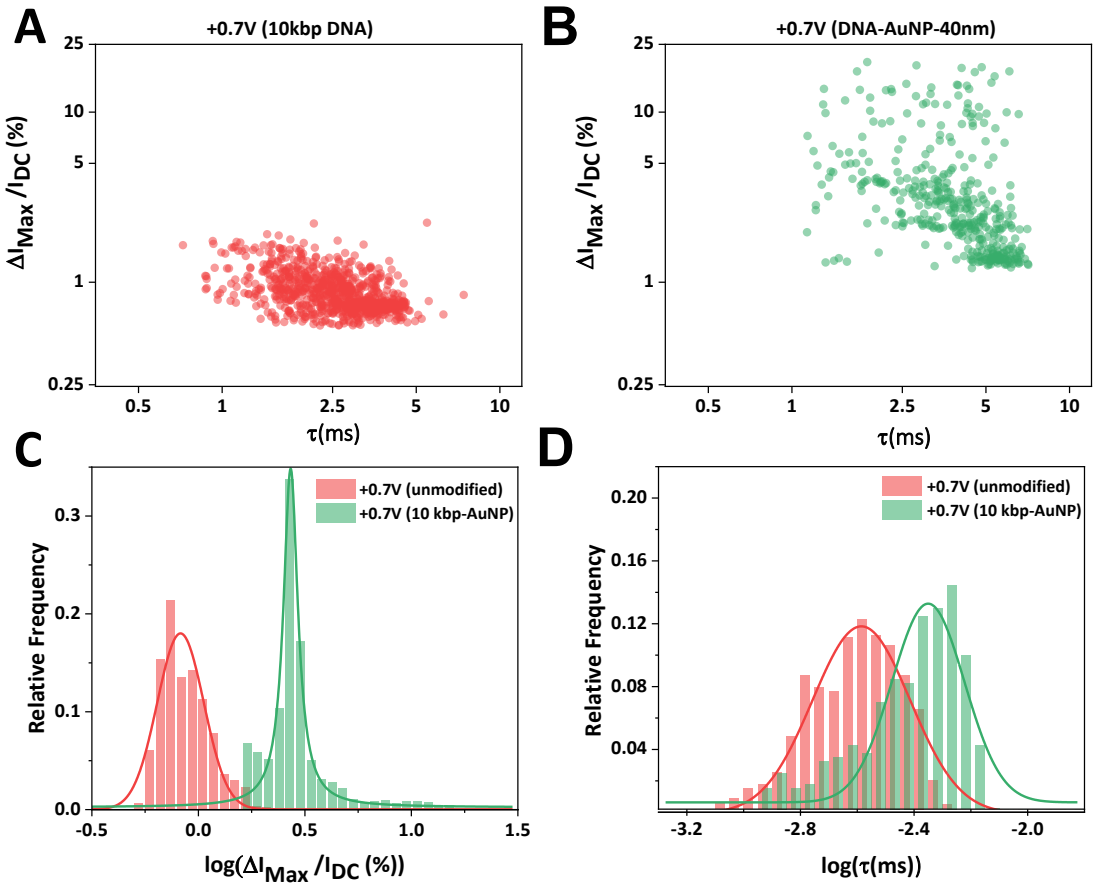

Figure 4.4 $\Delta I_{Max}/I_{DC}$ (%) vs τ plot for (A) unmodified 10 kbp DNA ($C_{DNA,\ out}$= 600 pM, G_{pore} = 36.9 nS) and (B) the incubation experiment ($C_{DNA,\ out}$ = 300 pM, $C_{SA\text{-}AuNP,\ in}$ = 500 pM, G_{pore} = 26.1 nS) at V_{bias} = +0.7 V.(C) histograms of log values of $\Delta I_{Max}/I_{DC}$ (%), along with gaussian fit (D) lognormal distribution of τ (red histograms: unmodified 10 kbp and green histograms: incubation experiment)

The above-mentioned results show only maximum current blockage without identifying their position or location along the biotinylated DNA. Hence, some of the spikes might be DNA-particle co-translocation or simply blockage of the pore. The biotin modification was localised at the 5' end of the carrier. Thus, in case of a linear DNA event bound to the particle, subevents spikes would predominantly occur at either the beginning or the ending of the bound event in an equally statistical ratio.[12-17] To confirm that the particles are bound to the carrier, the

position of subevent spikes should be identified through the application of proper subevent detection strategy.

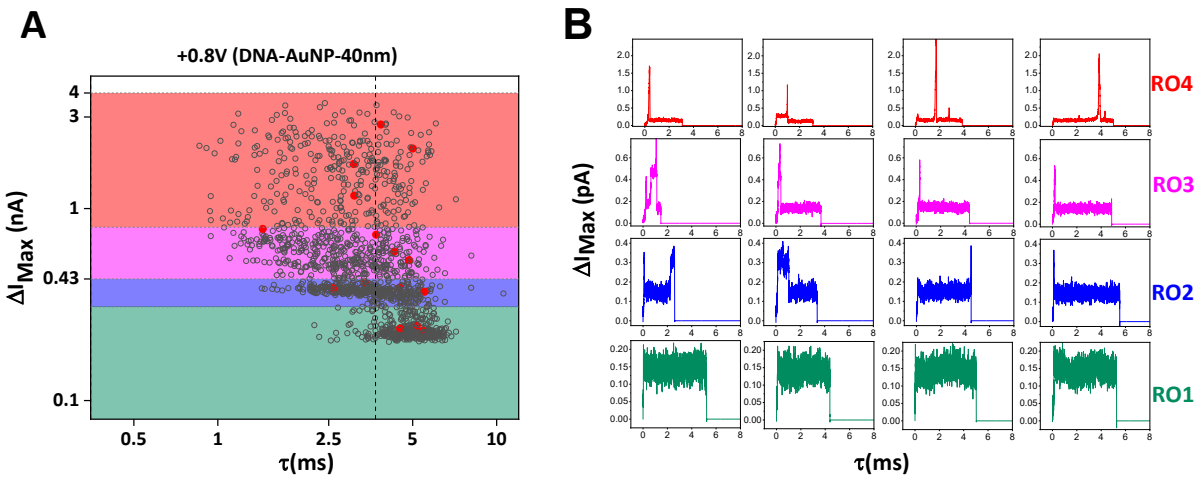

Figure 4.5 Translocation results of biotinylated 10 kbp DNA incubation with 40 nm SA-AuNP experiment ($C_{DNA,\ out}$ = 300 pM, $C_{SA\text{-}AuNP,\ in}$ = 500 pM, G_{pore} = 26.1 nS in 4M LiCl 1xTE buffer, V_{bias} = +0.8 V) (A) Scatter plot of ΔI_{Max} vs τ., coloured regions represent each Region-of-Interest (RoI1 To RoI5) used to separate events, based on ΔI_{Max} values from lower (green) to higher (red) (B) representative events from each RoI, coloured by means of region of interest as indicated ΔI_{Max} vs τ plot.

To test the success of particles capture, we have chosen to analyse +0.8 V dataset since it was the first reverse bias applied after loading the pipette with DNA. Hence, to carefully analyse the events, we first visualised the ΔI_{Max} x vs τ scatter plot for events recorded at V_{bias} = +0.8 V. The plot was divided it into four main Regions-of-Interest (RoIs) based on maximum current thresholding. This was performed to achieve a high-level classification of event types as shown in Figure 4.5A (RoIs are highlighted):

1) RoI1 (dark green) included the events has ΔI_{Max} < 300 pA.

2) RoI2 (blue) included the events that has ΔI_{Max} > 300 pA and < 430 pA

3) RoI3(magenta) included the events with ΔI_{Max} > 430 pA and < 800 pA.

4) Finally, RoI4 (red) included the events that has ΔI_{Max} > 800 pA and < 4.30 nA

Figure 4.5B presents randomly selected event shapes from each RoI and presented as the red-coloured datapoints (Figure 4.5A), arranged from shorter to longer τ values (left to right).

Accordingly, we have clustered the data based on maximum current threshold discussed above. Figure 4.6A shows a histogram of the log values of ΔI_{Max}. Whereas in Figure 4.6B gaussian fits were applied to the log values of τ histograms for each RoI (RoI1, RoI2, RoI3, RoI4).

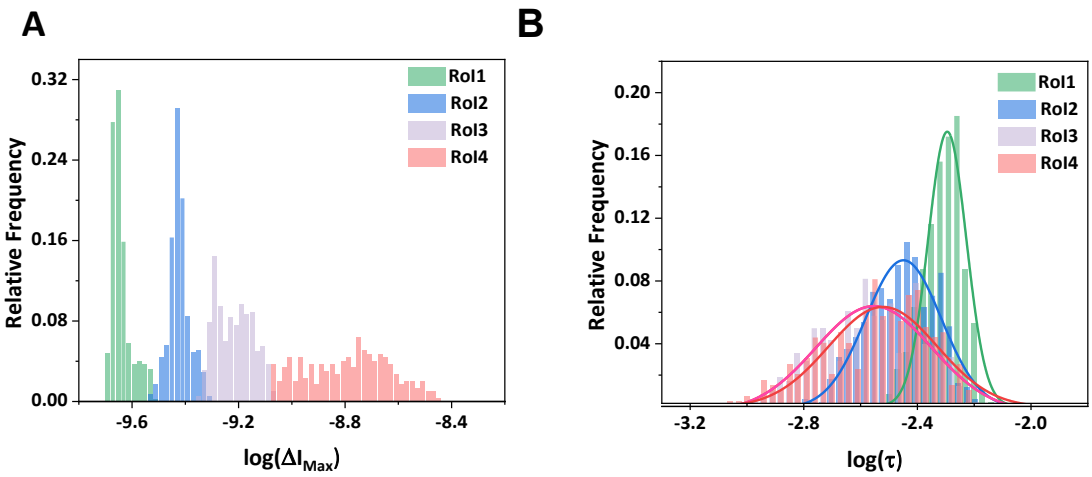

Figure 4.6 (A) histograms of log values of ΔI_{Max}, and (B) histograms with gaussian fit for log values of τ for each Region-of-Interest RoI (RoI1:green, RoI2:blue, RoI3:magenta, RoI4:red)

Looking more closely at these regions in Figure 4.5A and Figure 4.6B, RoI1 (green) consists of events occurring at relatively long τ and low ΔI_{Max} (~220pA). The dwell time distribution for RoI1 has a narrow distribution (σ = 0.068) and shifted toward longer duration (μ ~ 5 ms). Additionally, RoI1 events does not show significant sub-structure or a spike. Thus, RoI1 region is predominantly associated with the translocation of unbound linear DNA.

Moving to RoI2 (blue), we observe a cluster of events with a wider τ (σ = 0.13) and narrow ΔI_{max} distribution (320–430pA converted to antilog). The event shapes in this region suggest

two distinct behaviours: longer τ events correspond to linear DNA translocations with short, spike-like sub-events, whereas shorter τ events are more likely folded DNA translocation.

Notably, these folded DNA events display an initial excursion to approximately twice the ΔI_{max} value of linear DNA translocation, with a significant subevent duration (cf. Chapter 3, Section 3.3.3).

Shifting to RoI3 (magenta), we again see a broader distribution of events dwell time (σ = 0.20). This indicates mixed DNA confirmational entry into either linear or folded form, in line with what observed in RoI2. However, we distinguished the presence of short, intense sub-event feature, in both linear and folded events with ΔI_{Max}values typically lower than 800 pA.

Finally, RoI4 (red) forms a more diffuse cluster of events that resemble those in RoI3, but with even a higher intensity sub-event (ΔI_{Max}> 1nA). Overall, the intensity of the events in RoI4 suggests increased intensity, possibly indicating pipette blockage by larger particles.

Nevertheless, a more detailed analysis of the position and duration of the maximum current spikes is necessary to confirm the success of the analyte capture strategy. To investigate the origin of intense spikes (>570 pA, higher than the level of folded DNA), a sub-event detection strategy was employed. In this regard, linear events are more representative, as we discussed in Chapter 1 and Chapter 3.

Since we have a complex translocation behavior presented by RoIs, we decided to extract linear events using a more standardized approach. Accordingly, a Gaussian fit was applied to the overall τ distribution. Then, the mean (τ_{m1}) and standard deviation (σ_{m1}) were calculated. This provides a statistical framework to identify relevant events. Afterwards, linear events were selected if their dwell time exceeded (τ_{m1}), ensuring the inclusion of longer duration,

more distinct signals likely corresponding to linear DNA translocation. Figure 4.7A presents Gaussian fit results where we obtained $\tau_{m1} \pm \sigma_{m1}$ = 3.68 ± 1.61 ms.

In Figure 4.7B, the green dashed vertical line at τ=3.68 ms in ΔI_{max} vs τ scatter present the cutoff value used to extract linear events. At this point, we identified a total of 755 events across all Regions-of-Interest (RoIs) and subjected them to subevent analysis

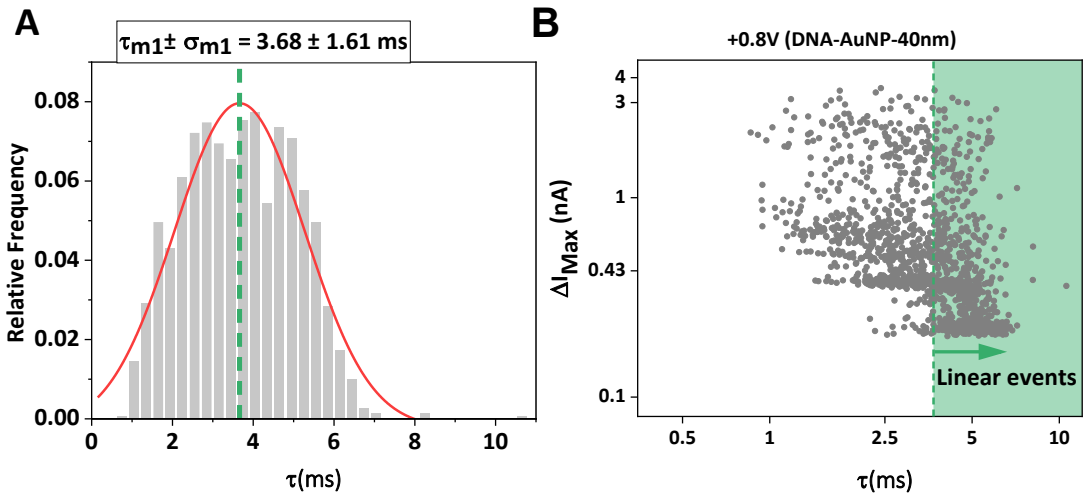

Figure 4.7 Linear event extraction from the translocation results of the incubation experiment ($C_{DNA, out}$ = 300 pM, $C_{SA-AuNP, in}$ = 500 pM, G_{pore} = 26.1 nS in 4M LiCl 1xTE buffer, V_{bias} = +0.8 V) (A) Histograms with gaussian fit for values of τ for overall DNA cluster, green dashed line present the mean (τ_{m1}) obtained from the fit (B) Scatter plot of ΔI_{Max} vs τ, region highlighted in green represents linear events selected for subevent detection($\tau > \tau_{m1}$)

After the extraction of linear events, we proceeded to subevent detection. As described in the methods (Chapter 2, Section 2.2.5.5), the search for subevent features was based on choosing a cutoff for ΔI_{Max} values of the events. Hence, choosing sub-event cutoff was articulated based on the median of ΔI_{Max} values for the filtered events ($\Delta I_{Max,median}$ = 320 pA at V_{bias} = +0.8 V). As shown in Figure 4.8 after applying a 320 pA cutoff, we classified 377 events as having no subevent (unbound). In contrast, 378 events were categorised as bound, among

which 307 exhibited a single subevent, 67 displayed two subevents, and 4 events contained three subevents (see Figure 4.10 panel A and B, respectively).

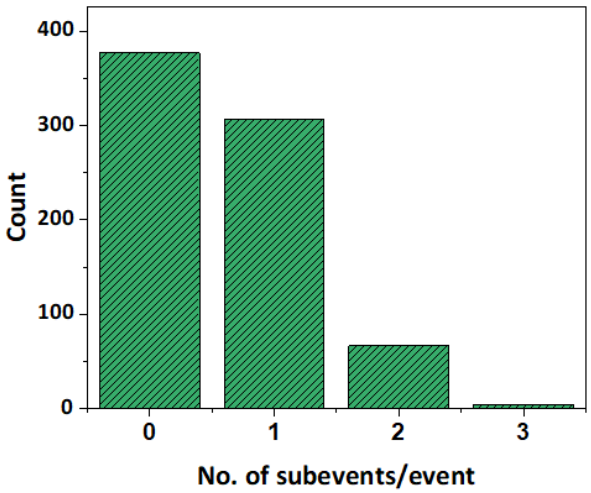

Figure 4.8 column plot shows events count categorized by the number of subevents detected using a ΔI_{Max} cutoff of 320 pA at V_{bias} = +0.8 for extracted linear events.

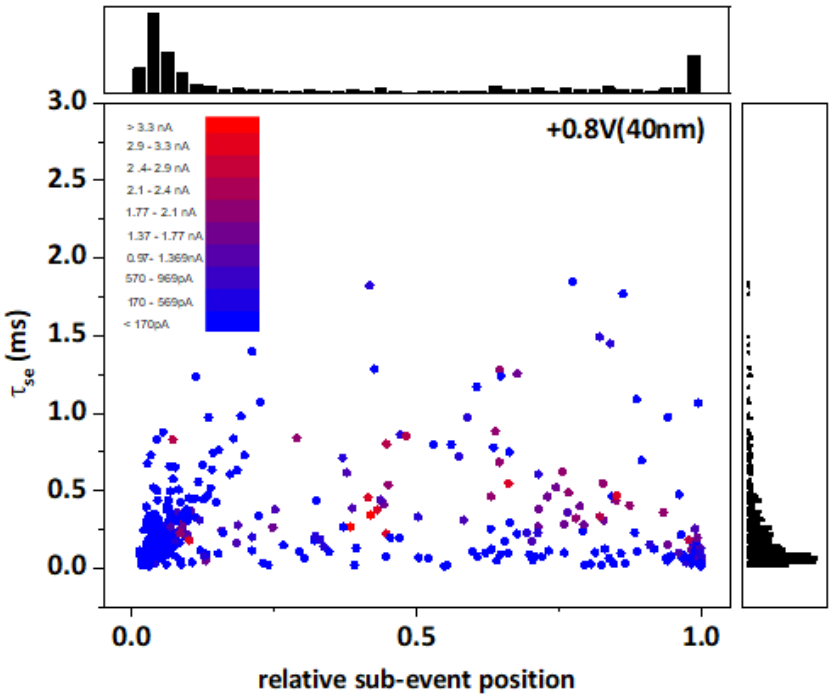

Figure 4.9 subevent analysis of linear events extracted from the incubation experiment, A Scatter plot of subevent duration (τ_{se}) against the relative subevent position within an event (0: event start, 1: event end) is shown. Subevents are colour coded based on the magnitude of sub-event starting from < 170 pA (blue) to > 3.3 nA (red). Corresponding histograms for relative sub-event position and τ_{se} are shown on the top and right sides.

Figure 4.9 displays the results of subevent analysis of linear events. Further, the results are shown as a scatter plot of the duration of subevent (τ_{se}) versus the relative sub-event position within an event. The sub-event magnitudes are color-coded (blue to red) from < 170 to > 3.3 nA, respectively. The corresponding histograms for relative sub-event position and τ_{se} are shown on the top and right sides.

Therefore, we observe that most of the sub-events are concentrated near the start or the end of an event (337 out of 453 subevents within the first or last 15%). Further, these subevents (at first or last 15% of the event) exhibit both relatively short durations (mode = 89 µs) and low current magnitudes (< 570 pA). Obviously, from the statistical distribution of both subevent duration and the relative position, we can conclude that particles are successfully bound to the biotinylated 10 kbp DNA. Accordingly, our new incubation method demonstrated successful binding and detection of nanoparticles. As presented, blockage events (>1 nA, red), where more diffuse and scattered along the events. However, blockages events were statistically less significant compared to the bound events (< 800 pA).

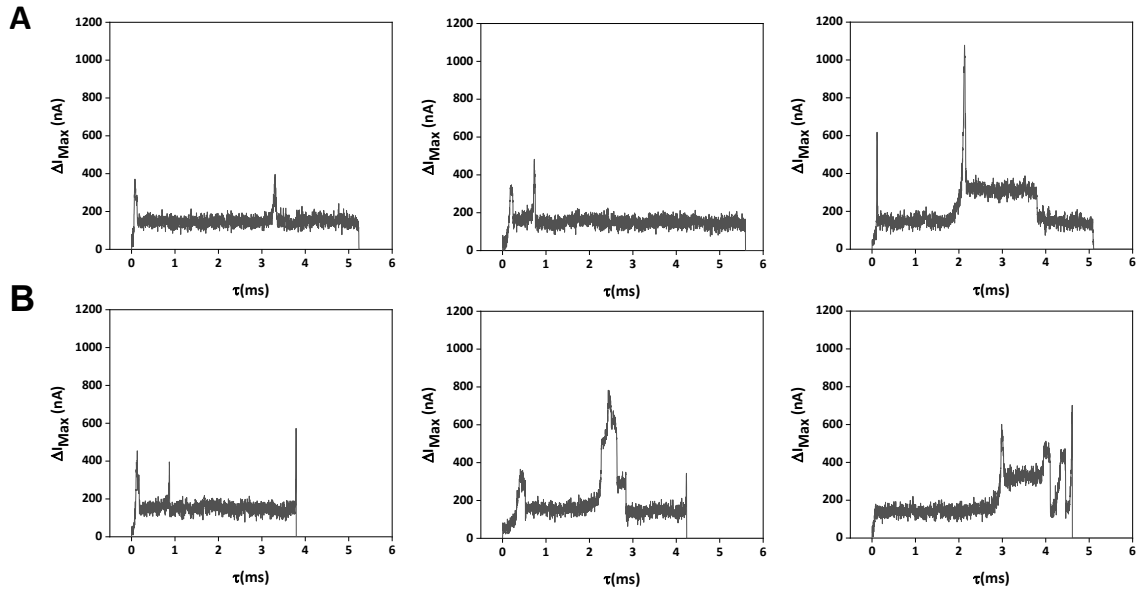

Figure 4.10. Current-time traces of selected events possess (A) two subevents and (B) three subevents

4.2.4. Detection of 5 nm SA-AuNP Inside by Reverse Translocation

After we demonstrated successful capture of 40 nm particles using incubation strategy, we sought to evaluate the same approach using streptavidin-functionalized nanoparticles with an average diameter of 5 nm (according to manufacturer specifications). The average diameter is significantly smaller than 40 nm. In this experiment, we used a higher concentration of the particles inside the pipette ($C_{SA\text{-}AuNP,\ in}$ = 2.7 nM) to compensate for the lower size effect. To expand on this point, we detailed in chapter 3 that the volume of the nanochannel (V_{ch}) needs to be filled with a certain number of molecules before reaching the crowded state (cf. section 3.3.7). Since the concentration of 40 nm particles experiment was sufficient for their capture by functionalized DNA ($C_{SA\text{-}AuNP,\ in}$ = 500 pM). The concentration of smaller particles was increased by a factor of 4.5, bearing in mind that the hydrodynamic radius of 5 nm particles is

approximately ~9 nm in 4 M LiCl.[18, 19] Based on this, we can achieve a capturing efficiency comparable to that of 40 nm particles.

The experimental procedure followed the same experimental setup with the 40 nm SA-AuNP. Hence, we loaded the pipette with biotinylated 10 kbp DNA ($C_{DNA,\ out}$ = 300 pM) at V_{bias} = -0.8 V. Then, DNA-particle incubation was allowed for 20 minutes before reversing the bias at V_{bias} = +0.8 V. To ensure consistent results, we used a nanopipette with a comparable size (G_{pore} = 26.3 nS).

Figure 4.11A presents the ΔI_{Max} *vs* $\tau_{scatter}$ plot for events detected at V_{bias} = +0.8. As follows, we can observe a similar response to that observed with 40 nm particles. In this manner, we have separated the events into RoIs (1-4), following the approach used in the 40 nm particles incubation experiment. In Figure 4.11B, randomly selected events from each RoI are shown, which are arranged from shorter to longer τ values (left to right).

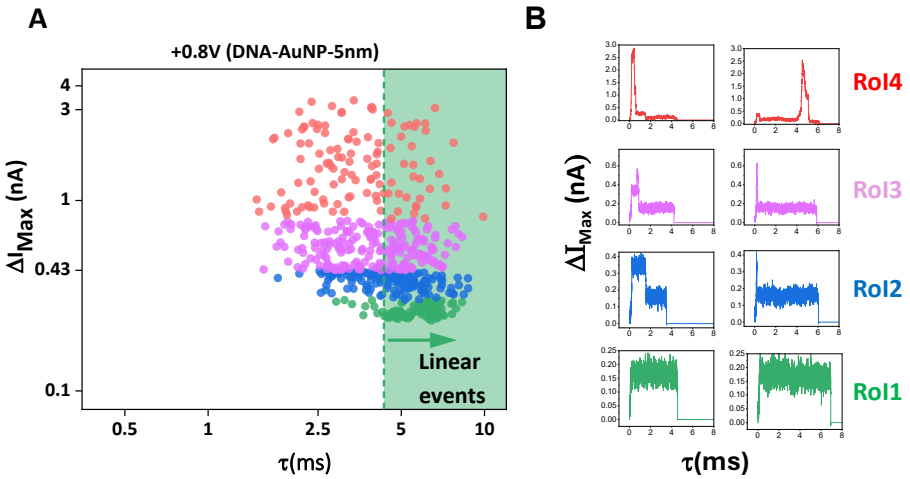

Figure 4.11 Translocation results of biotinylated 10 kbp DNA incubation with 5 nm SA-AuNP experiment ($C_{DNA,\ out}$ = 300 pM, $C_{SA\text{-}AuNP,\ in}$ = 2.7 nM, G_{pore} = 26.3 nS in 4 M LiCl 1xTE buffer, V_{bias} = +0.8 V) (A) Scatter plot of ΔI_{Max} vs τ., coloured datapoints represent each Region-of-Interest (RoI1 To RoI5) used to separate events based on ΔI_{Max} values, starting from lower (green) to higher (red) (B) representative events from each RoI, coloured by means of region of interest as indicated ΔI_{Max} vs τ plot.

Despite using the same concentration of DNA in the bulk solution (300 pM), there were fewer events detected overall. Accordingly, we monitored the current time traces for any significant changes in the AC channel noise. Thus, in 5 nm particles experiments, we have found an increased level of noise accompanied by a decrease in pore conductance. Thus, instead of analyzing the total 100 files, we have partially removed noisy traces from the analysis.

Afterwards, we have clustered the events based on RoIs as described in Section 4.2.3. Figure 4.12A shows a histogram of the log values of ΔI_{Max} and Figure 4.12B presents the gaussian fits applied to the log values of τ histograms for each RoI (RoI1, RoI2, RoI3 and RoI4).

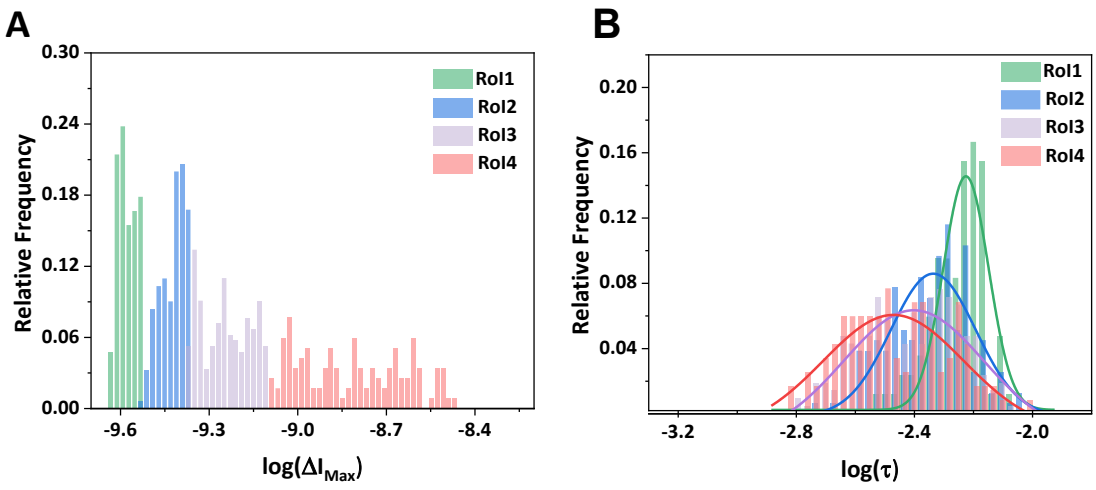

Figure 4.12 (A) histograms of log values of ΔI_{Max}, and (B) histograms with gaussian fit for log values of τ for each Region-of-Interest RoI (RoI1:green, RoI2:blue, RoI3:magenta, RoI4:red)

Like what was observed in the 40 nm particle datasets (V_{bias} = +0.8 V), RoI1 (green) consists of events occurring at relatively long τ and low ΔI_{Max} (~220 pA). Further, τ distribution for RoI1 is narrower and occurs at longer duration ($\mu_1 \pm \sigma_1$ = -2.22 ± 0.077, in log scale). Thus, RoI1 region is predominantly associated with the translocation of unbound linear DNA.

On the other hand, τ distribution of RoI2 is wider (σ_2 = 0.14), showing a narrow ΔI_{max} distribution (<430 pA). This suggests a mixed behaviour of DNA translocation (linear and

folded), as discussed earlier no subevent significant spikes observed in folded events (higher level of current indicates folding where $\Delta I_{max,\ Folded} \approx 2\ \Delta I_{max,\ Linear}$). Whereas RoI2's linear events show a subevent like feature. Further, RoI3 shows wider τ distribution compared to that of RoI2 ($\sigma_3 = 0.23$). However, ΔI_{max} values (>430 pA and less than 800 pA) are significantly higher with increased emergence of subevent spikes. RoI4 region (> 800 pA) shows similar τ distribution but ΔI_{max} is more diffused with complex events shape, indicating pore blockage. As the particles are smaller in size, they probably translocate fully outside the pore. Thus, the emergence of RoI4 region is surprising. However, smaller size (5 nm) comes at risk of increased aggregation and impacted stability.[20] Hence, high intensity events (> 800 pA) indicates the presence of particle aggregation. As a result, aggregated particles could potentially block the pore, complicating event detection.

As discussed earlier in section 4.2.3, extraction of linear events was approached based on the determination of τ_{m1} from a gaussian fit, where $\tau_{linear} > \tau_{m1}$. Figure 4.13A presents Gaussian fit results where we obtained $\tau_{m1} \pm \sigma_{m1} = 4.33 \pm 1.85$ ms. In Figure 4.13B, the green highlighted region at $\tau = 4.33$ ms in ΔI_{max} vs τ scatter presents the region of selected linear events.

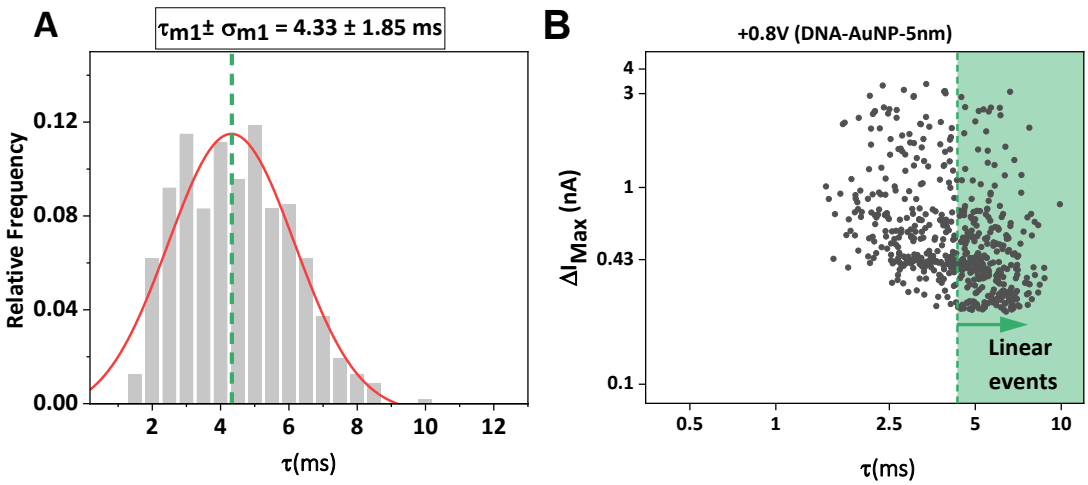

Figure 4.13 Linear event extraction from the translocation results of the incubation experiment with 5 nm particles ($C_{DNA,\ out}$ = 300 pM, $C_{SA\text{-}AuNP,\ in}$ = 2.7 nM, G_{pore} = 26.3 nS in 4M LiCl 1xTE buffer, V_{bias} = +0.8 V) (A) Histograms with gaussian fit for values of τ for overall DNA cluster, green dashed line present the mean (τ_{m1}) obtained from the fit. (B) Scatter plot of ΔI_{Max}vs τ, green highlighted region presents linear events selected for subevent detection($\tau > \tau_{m1}$)

Given that G_{pore} values were comparable in both 40 nm and 5 nm incubation experiments (26.3 nS and 26.1 nS, respectively). Electrical responses are expected to be similar in both experiments. Therefore, a subevent analysis strategy was conducted using the abovementioned approach (section 4.2.3). Specifically, we applied a cutoff of 320 pA ($\Delta I_{Max,median}$ = 360 pA, less events were recorded). The cut off applied was lower than $\Delta I_{Max,median}$, because we visualized that there was a misclassification of bound events at higher cutoff (>350 pA). Through this approach, we identified 282 linear events (Figure 4.13B) and subjected them to subevent analysis. Among these 282 events, 248 subevents were detected: 107 events exhibited no subevents (unbound), while 115 events contained a single subevent, 47 had two subevents, and 13 displayed three subevents (Figure 4.14A).

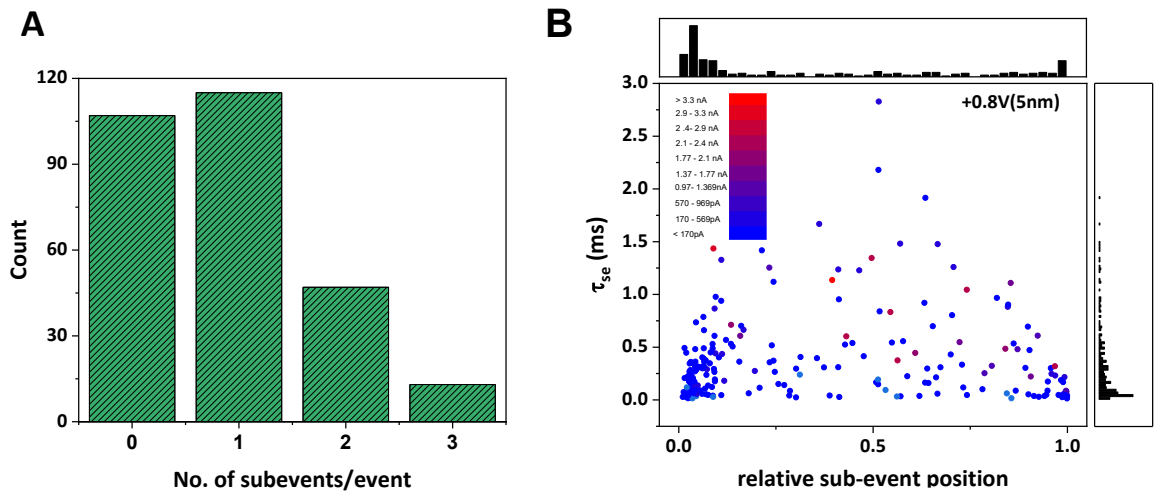

Figure 4.14 subevent analysis of linear events extracted from the incubation experiment with 5 nm SA-AuNP (A) column plot shows events count categorized by the number of subevents detected using a ΔI_{Max} cutoff of 320 pA at V_{bias} = +0.8. (B) Scatter plot of subevent duration (τ_{se}) against the relative subevent position within an event (0: event start, 1: event end) is shown. Subevents are colour coded based on the magnitude of sub-event starting from < 170 pA (blue) to > 3.3 nA (red). Corresponding histograms for relative sub-event position and τ_{se} are shown on the top and right sides.

For the analysis of the relative subevent position is shown in Figure 4.14B, the subevent duration (τ_{se}) is plotted as a function of the relative subevent position within an event. Probability histograms of τ_{se} and relative subevent position, displayed on the right and top sides of Figure 4.14B, respectively. Further, the analysis reveals that most subevents (150 out of 248) are concentrated near the start or end of an event (within the first or last 15% of the event). Compared to the 40 nm SA-AuNP dataset, these 150 subevents exhibited similar magnitudes (< 570 pA), but their durations were relatively shorter (mode = 36 µs). Given that the 5 nm particles share a similar spherical shape and polymer coating with the 40 nm particles (despite a smaller diameter), thus, they may translocate more rapidly, leading to shorter subevent durations. However, the difference in subevent duration (τ_{se}) between both experiments could also be attributed to variations in the pipettes used across experiments.

Subevents exceeding 1 nA (red) are randomly scattered within the event as shown in the scatter plot are likely associated with particle aggregation. These responses (> 1nA) are blockages events similar to what observed in the 40 nm experiment.

Toward this end, our findings demonstrate that the proposed approach can capture both 40 nm and 5 nm particles. However, direct size comparison based on event magnitude alone proved challenging, despite the shorter duration observed for smaller particles. Previous studies have shown that the current response produced by a translocating particle follows a linear relationship proportional to the cubic power of its diameter.[5] To better distinguish between the two particle sizes while minimising the impact of pore size variability, we analysed the relative subevent intensity, defined as the ratio of subevent current maxima to the associated event median ($\Delta I_{Max,subevent,} \,/\, \Delta I_{Median,event}$).

Figure 4.15A presents histograms of relative intensity values for both particle sizes. A closer examination of these distributions reveals a distinct subset of subevents for 5 nm particles with a relative subevent intensity of approximately 1.7, this value is absent in the 40 nm dataset. In contrast, the relative intensity for 40 nm particles starts at 2.2, with a broader overlapping range between 2.5 and 25. This overlap is expected, as previously discussed, given that larger 40 nm particles can block the pore, whereas 5 nm particles may aggregate, leading to mixed distributions. To rigorously assess the statistical significance of these differences, we applied the Mann-Whitney U test to compare the two groups.

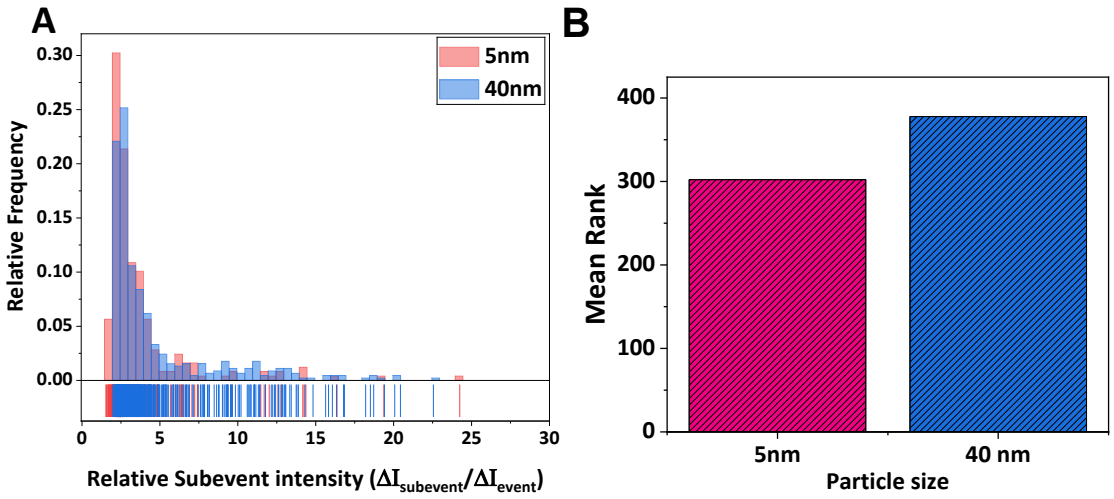

Figure 4.15 comparison between 5 nm and 40 nm SA-AuNP incubation experiments in terms of (A) relative subevent intensity, defined as the ratio of subevent current maxima to event median ($\Delta I_{Max,subevent,} / \Delta I_{Median,event}$) and(B) mean rank distributions obtained through Mann–Whitney U test

The results of the Mann-Whitney test indicate a significant difference between the relative subevent intensities of 40 nm and 5 nm particles (U=68270, Z=4.72Z and p < 0.001).[21] Specifically, the median relative intensity for 40 nm particles is higher than that of 5 nm particles, supporting the hypothesis that larger particles produce greater relative subevent intensities. Moreover, the rank distributions suggest a systematic shift, with 40 nm particles exhibiting a higher mean rank compared to 5 nm particles as shown in Figure 4.15B. These findings reinforce the capability of our approach to differentiate between particle sizes based on their translocation characteristics, despite the complexities introduced by aggregation and pore blockage effects. As discussed, we have observed a heterogeneity in the experimental results. Accordingly, this could be attributed to variability in nanoparticles interactions, pore sizes, and stochastic blockages. This complex behaviour is expected in nanopore experiments, suggesting the need for statistical interpretation instead of relying on individual events alone.

4.3. Conclusion

Based on conclusions and results from Chapter 3, Chapter 4 shows the results of proof of concept of study, where the analyte of interest incubated inside the nanopipette to allow spatial trapping. Specifically, the nanopipette was filled with a solution containing streptavidin-modified, 40 nm gold nanoparticles. Then, biotinylated 10 kbp DNA was translocated from the outside to the inside of the nanopipette and allowed for incubation. After 20 minutes, 10 kbp DNA was reverse translocated out of the pipette. The analysis of the events distributions allowed for the identification of various types of events associated with bare DNA and DNA nanoparticle complexes. Further, from subevent analysis we have shown that a statistically significant subevent feature relies at either the end or the beginning of the event structure belongs to bound DNA-NP bound structure. This indeed demonstrated successful capture of the analyte. Accordingly, we explored this approach to capture smaller size nanoparticles (5 nm streptavidin-functionalised gold nanoparticles). Again, we identified similar binding patterns in terms of events intensity distribution. However, subevent feature demonstrated shorter translocation duration. To inspect differences in subevent intensities, we initiated specific analysis to compare both samples. Hence, we found that smaller particles exhibit a distribution of subevents with lower intensity compared to the larger particles, demonstrating size comparison. Accordingly, the idea of analyte incubation and detection in the nanopipette proposes a new bioanalytical approach. This approach can be further applied to biomolecules implicated in pathological conditions such as proteins or RNA. Additionally, targets incubated inside the nanopipette could be suspended with a physiologically relevant solution, enabling real clinical samples to be integrated in the future.

4.4. References

1. R. A. Al-Waqfi, C. Khan, O. J. Irving, L. Matthews and T. Albrecht, *arXiv preprint arXiv:2501.14347*, 2025.
2. W. J. Lan, D. A. Holden, B. Zhang and H. S. White, *Anal. Chem.*, 2011, **83**, 3840-3847.
3. D. Branton, D. W. Deamer, A. Marziali, H. Bayley, S. A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, S. B. Jovanovich, P. S. Krstic, S. Lindsay, X. S. Ling, C. H. Mastrangelo, A. Meller, J. S. Oliver, Y. V. Pershin, J. M. Ramsey, R. Riehn, G. V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin and J. A. Schloss, *Nat. Biotechnol.*, 2008, **26**, 1146-1153.
4. M. Firnkes, D. Pedone, J. Knezevic, M. Döblinger and U. Rant, *Nano Lett.*, 2010, **10**, 2162-2167.
5. L. Luo, S. R. German, W. J. Lan, D. A. Holden, T. L. Mega and H. S. White, *Annu Rev Anal Chem (Palo Alto Calif)*, 2014, **7**, 513-535.
6. R. L. Fraccari, M. Carminati, G. Piantanida, T. Leontidou, G. Ferrari and T. Albrecht, *Faraday Discuss.*, 2016, **193**, 459-470.
7. R. L. Fraccari, P. Ciccarella, A. Bahrami, M. Carminati, G. Ferrari and T. Albrecht, *Nanoscale*, 2016, **8**, 7604-7611.
8. M. Wanunu, J. Sutin, B. McNally, A. Chow and A. Meller, *Biophys. J.*, 2008, **95**, 4716-4725.
9. J. M. Teulon, C. Godon, L. Chantalat, C. Moriscot, J. Cambedouzou, M. Odorico, J. Ravaux, R. Podor, A. Gerdil, A. Habert, N. Herlin-Boime, S. W. Chen and J. L. Pellequer, *Nanomaterials (Basel)*, 2018, **9**.
10. Y. Xing, A. Rottensteiner, J. Ciccone and S. Howorka, *Angew. Chem. Int. Ed. Engl.*, 2023, **62**, e202303103.
11. N. A. W. Bell, K. Chen, S. Ghosal, M. Ricci and U. F. Keyser, *Nat Commun*, 2017, **8**, 380.
12. M. Mihovilovic, N. Hagerty and D. Stein, *Phys. Rev. Lett.*, 2013, **110**, 028102.
13. N. A. Bell and U. F. Keyser, *Nat Nanotechnol*, 2016, **11**, 645-651.
14. T. Albrecht, *Annu Rev Anal Chem (Palo Alto Calif)*, 2019, **12**, 371-387.
15. C. Plesa, N. Van Loo, P. Ketterer, H. Dietz and C. Dekker, *Nano Lett.*, 2015, **15**, 732-737.
16. C. Plesa, L. Cornelissen, M. W. Tuijtel and C. Dekker, *Nanotechnology*, 2013, **24**, 475101.

17. L. J. Steinbock, O. Otto, C. Chimerel, J. Gornall and U. F. Keyser, *Nano Lett.*, 2010, **10**, 2493-2497.

18. S. W. Kowalczyk, D. B. Wells, A. Aksimentiev and C. Dekker, *Nano Lett.*, 2012, **12**, 1038-1044.

19. G. U. Nienhaus, P. Maffre and K. Nienhaus, in *Methods Enzymol.*, ed. S. Y. Tetin, Academic Press, 2013, vol. 519, pp. 115-137.

20. W. Zhang, *Adv Exp Med Biol*, 2014, **811**, 19-43.

21. H. B. Mann and D. R. Whitney, *Ann. Math. Stat.*, 1947, 50-60.

CHAPTER 5
RNA ASSAY-ON-A-STRING FOR BLADDER CANCER DETECTION

CONTENTS

Synopsis: this chapter evaluates a carrier based nanopore biosensing strategy for the capture of long non-coding RNA (lncRNA) larger than 2000 bases without fragmentation or amplification. Specifically, urothelial cancer associated 1(UCA-1) lncRNA was chosen as a model target since it has been strongly linked to the emergence of bladder cancer. The approach first incorporated the design and the characterisation of an M13 carrier DNA with two complementary capture probes for detecting UCA-1 lncRNA (2.3kb). Then, the translocation characteristics of the carrier, free RNA and carrier-RNA hybrid were analysed by nanopipettes. Finally, electrical signals of each were compared. Accordingly, the challenges and the potential of this strategy were further discussed.

5.1. Introduction

Ribonucleic acids (RNA) are essential biomolecules for protein synthesis, as well as regulators for gene expression. Thus, RNA is either coding messenger RNA (mRNA) or non-coding RNAs. There are three non-coding RNAs: ribosomal (rRNA), transfer (tRNA), and the regulatory micro and long non-coding RNAs (miRNA & lncRNA).[1] Dysregulation and mutations within these biomolecules have been associated with the emergence of various pathological conditions, such as oncogenesis.[2-4] The rapid proliferation of tumour cells could initially release circulating cell-free biomarkers (Cf). Among these biomarkers are miRNA and lncRNA which can be expressed at low concentrations in various biological fluids including urine, saliva and serum based on the tumour specific location.[5, 6] Hence, the early detection of miRNA and lncRNA directly from biological fluids offer a minimally invasive alternative to tissue biopsies and enable the early detection of malignancies.[4, 7, 8] UCA1 is a lncRNA with 2314 bases that has emerged as a significant biomarker in bladder cancer, with studies demonstrating increased sensitivity and specificity in urine samples for bladder cancer detection.[9, 10] Furthermore, UCA1 has been implicated in chemoresistance to certain anti-tumour agents like cisplatin. Thus, the detection of UCA1 at low levels could enhance the early detection strategies and inform treatment responses in clinical oncology.[9, 11-14] Conventional detection methods of RNA biomarkers are based on enzymatic transcription and amplification such as RT-qPCR and microarrays.[4] Despite their cost, these techniques are known to have limited sensitivity and specificity for low-abundance transcripts like UCA1.[4] RNA sequencing provides high precision transcriptomic analysis; however, complicated data analysis and cost limits its capability in point of care detection.[15, 16] Taking this into account, amplification-free biosensing

technologies, including optical surface plasmon resonance (SPR) and electrochemical biosensors (ES) have been investigated as alternative platforms. While SPR provides real-time detection via refractive index shifts, its efficacy is compromised at low RNA concentrations.[17-22] Conversely, ES-based approaches offer femtomolar sensitivity but suffer from fabrication complexity, limiting their translational potential in clinical applications.[17-22] In combination with molecular probes such as DNA carriers, solid state nanopores has progressed as a promising single molecule method for analyte detection (introduced earlier in chapter 1 and chapter 4).[23, 24] Despite extensive research work on miRNA and short fragments, [25-27] there is less work conducted on long RNA molecules. For instance, the Keyser group has contributed significantly to the area of RNA research by detecting RNA isoforms in long RNA molecules. Their strategy involved folding long RNA molecules by carefully designed DNA strands with a 3D structural motif termed structural colour at multiple positions within the RNA molecules. The location of these structural colours was then analysed by nanopores to distinguish different RNA isoforms.[28] Further, the group has developed an RNA detection method based on DNA nanoswitches by detecting both closed state (target-bound) and open state (unbound) along with their hybridised position.[29] They have also explored other biophysical aspect about RNA and RNA/DNA hybrids.[30] Although their work demonstrated significant advance, complementary DNA strands requires redesign as overall for each long RNA structure, which might be expensive. Collectively, these advancements highlight the promise of solid-state nanopores as a powerful tool for RNA biomarker detection. Apart from carrier-based detection, other studies have also explored the potential of using polymer-electrolyte solid-state nanopores to study long RNA molecules conformations at physiological conditions, focusing in enhancing signal to noise ratio in terms of intensity.[31] As outlined in Chapter 1 and

Chapter 4 regarding the translocation properties of a linear DNA,[32] DNA molecules can be hybridised with specific binding sites for analyte capture that can be further detected by nanopores. [33, 34] Previous work by our group has demonstrated the potential of carrier-based method to capture short DNA fragments for tuberculosis diagnostics. Specifically, a DNA carrier was modified with short ss DNA probes (overhangs) that are 88 nucleotides in length. And once the target hybridises, the carrier will form a protrusion that can be detected by nanopipettes (Figure 5.1).[13]

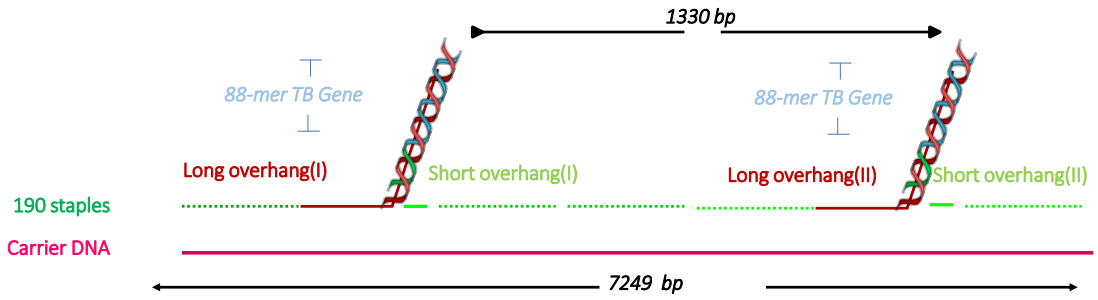

Figure 5.1 : representative diagram of M13mp18 carrier with overhang hybridisation strategy designed by Loh[13], two protrusions in the middle of the carrier represents the short ds DNA after hybridising with the target (RV1910c gene).[13]

Benefiting from this strategy, this design can be adapted for the detection of lncRNA. Simply, the 1300 bp spaced overhangs can be modified to capture lncRNA at each specific end of the sequence. Such a simple strategy could capture lncRNA without the need of extra processing steps like fragmentation/amplification or requiring expensive complex strand design strategy for detecting the UCA-1 lncRNA.[35, 36]

5.2. Experimental objective

The objective of this chapter was to evaluate a nanopore-based approach for the capture of UCA1 lncRNA which is a potential biomarker for bladder cancer detection. This involved designing a hybridised DNA nanostructure with two complementary overhangs, targeting both the 88-mer end and the first specific start sequences of in vitro-transcribed UCA1 lncRNA. Primarily, the study aimed to assess the ability of nanopipettes to capture intact lncRNA without requiring fragmentation or the design of multiple oligonucleotides hybridising probes, instead employing a universal carrier with dual capturing probes.

5.3. Results and Discussion

5.3.1. Analysis of Probe Selection and Design

Based on previous work from the group, this work was conducted to advance carrier-based detection method in a new application for targeting lncRNA. Mainly, this strategy was initially inspired from the combined application of nanopore sensing and DNA origami.[32, 37] The first example of this application in solid state nanopores was demonstrated successfully by Plesa in 2015 for detecting DNA protrusions.[32] This approach was later expanded upon by the Keyser group through DNA barcoding for analyte detection.[37] To design our construct, M13 bacteriophage was employed as a single-stranded DNA (ssDNA) scaffold (refer to Chapter 2, Section 2.1.3.4), an inexpensive and widely used template for constructing DNA origami structures. The process initially involved linearisation of circular M13mp18 scaffold using two restriction enzymes (EcoRI and BamHI-HF) as reported by Bell et al.[14] After linearisation, DNA scaffold was hybridised with a set of complementary staple oligos, where specific oligos are

extended with extra nucleotides to serve as capture probes. Further, the final dsDNA construct has two capture probes (long overhangs). Accordingly, two long overhangs, both complementary to the scaffold, with one matching the 88-mer of the first specific sequence of UCA-1 lncRNA while the other matching the last 88-mer end. To prevent steric hinderance and interaction, both long overhangs were spaced at approximately 1330 bp (452 nm) and were placed in the middle of the carrier.[13] Since their current response will be smaller, positioning in the middle helps reduce the misattribution of the initial excursion and DNA folding at the beginning of the carrier.[38] For further stability and enhanced sensing, short overhangs were added in which they have complementary regions with both the scaffold and the long overhangs as shown in Figure 5.1.[13, 14]

The choice of the two target sequences was based on specificity and stability. Given that UCA1 is 2314 bp as a full transcript (**NR_015379.3**, last accessed on January 2024).[39] The first 391 bases of UCA1 contain a palindromic repeat at position 1911.. Such repeats are often found in long non-coding RNAs, potentially contributing to their role in gene regulation.[40] The capture strategy was meant to reduce the impact of the secondary structure of lncRNA. Hence, stretch the RNA once it is bound to the capture probes (Figure 5.2). Accordingly, the capture probes were designed to target nucleotide positions 392 to 479 within a specific region of the UCA1 transcript (Region 1), as well as the final 88 nucleotides at the 3' end of the transcript (Region2).

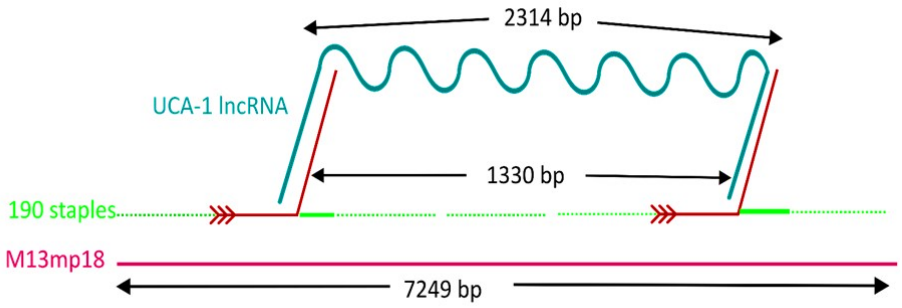

Figure 5.2: representative diagram of the proposed analyte strategy to capture UCA1; M13mp18 ssDNA carrier strand (pink), ss overhangs with capturing probes sequences (red) to UCA1 (blue), 190 staples strands to complement the ssDNA carrier(green).

5.3.2. Target Binding Efficiency

Estimation of the optimal concentration of the capture probes to bind efficiently with the lowest concentration of the target is necessary to implement the next steps in the assay design. Thus, the target's limit of detection was calculated based on the thermodynamic equilibrium calculation.[41] This consideration would help in determining the initial target concentration required to bind with the capture probes efficiently as the probe concentration varied. First, Gibbs free energy ($\Delta G_{binding\ site}$) was estimated using OligcCal online tool for each 88 nts sequence, applying equation (5.1): [42]

$$\Delta G_{binding\ site} = -RT\ ln\ K_d \tag{5.1}$$

K_d is the dissociation constant, R is gas constant and T is kelvin temperature.

Since we have two independent binding sites and one target, we calculated the total binding energy using equation (5.2): [41, 43]

$$\Delta G_{Total} = \Delta G_{binding\ site\ 1} + \Delta G_{binding\ site\ 2} \tag{5.2}$$

Where ΔG_{Total} is the total free energy for both binding sites while neglecting cooperativity.[43] This yielded a total free energy change ΔG_{Total} = -1110 kJ/mol (estimated $\Delta G_{binding\ site\ 1}$= -600 kJ/mol and $\Delta G_{binding\ site\ 2}$ =-510 kJ/mol).

Accordingly, concentrations can be determined from the dissociation constant $K_d$which can be calculated by equation (5.3): [13, 41, 43]

$$K_d = \frac{([UCA1]-[B])([UB]-[B])}{[B]} \quad (5.3)$$

Where $[UCA1]$ and $[UB]$are the initial target and probe concentration, respectively. While $[B]$is the probe-target complex concentration.

Finally, the probe-target complex concentration $[B]$ was determined using equation (5.4): [13]

$$[B] = \frac{([UCA1] + [UB] + K_d) - \sqrt{([UCA1] + [UB] + K_d)^2 + 4[UCA1][UB]}}{2} \quad (5.4)$$

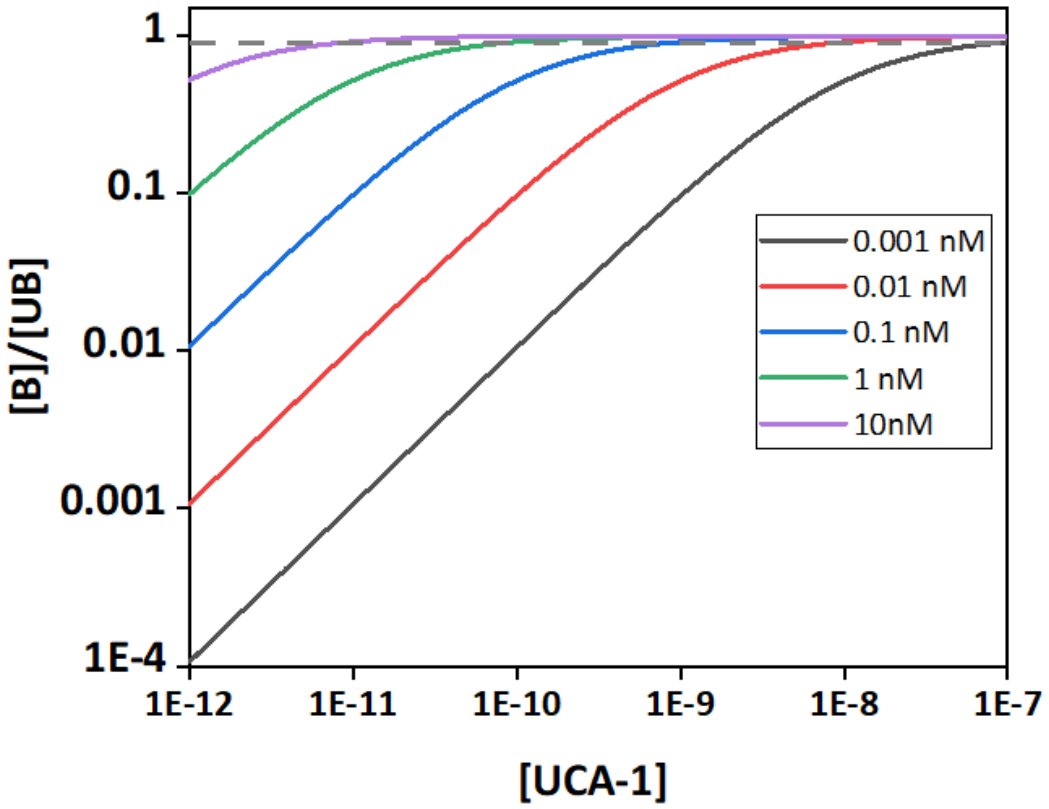

Figure 5.3: the concentration of the bound probe [B] relative to the initial probe concentration [UB] plotted as a function of the initial Target concentration [UCA-1]. [B]/[UB] at 0.9 in this case, indicates the maximum saturation, where 90% of binding sites were fully occupied. At this point, we can determine the limit of detection of the initial target concentration (as low as 10 pM).The initial probe concentration [UB] was varied to determine target concentration conditions, illustrating limit of detection.

Figure 5.3 represents the relative concentrations of [B] to the initial free probe or unbound carrier concentrations [UB] as function of free target [UCA1] concentration. In our main calculation we have varied probe concentrations from 1pM to 10nM as plotted in figure 5.4. at the point where [B]/[UB] fraction equals 0.9 as shown as dashed line in the plot, 90% of binding sites were fully occupied. Accordingly, we can determine the limit of detection of the initial target concentration [UCA1] as low as 10 pM in case we used 10 nM of the unbound probe [UB]. Thus, higher probe concentrations (1-10nM) shift the equilibrium toward more

stable bound probes. In accordance with pushing limit of detection to lower concentrations, we have decided to initially maintain an initial probe concentration of 5-10nM as final carrier DNA concentration.

5.3.3. UCA1 and Carrier DNA with Capture Probes Characterisation

To incorporate the detection of UCA-1 lncRNA, it has been sought out to synthesise the target to model the detection. First, the UCA-1 lncRNA sequence was accessed from the National Centre for Biotechnology Information (NCBI) database with the accession number NR_015379, as it was reported as the most recent and the most abundant transcript.[39] Due to the challenges associated with synthesizing a 2314 bp sequence via PCR, the UCA-1 sequence was purchased as a plasmid insert containing an SP6 promoter at the 5'-end after linearisation. To pursue in vitro transcription (IVT), UCA1 plasmid was first linearized using SFil RE enzyme which has two RE sites at the end and the beginning of the sequence (refer to Chapter 2, Section 2.1.4). Accordingly, gel electrophoresis was performed to confirm successful linearization, as shown in Figure 5.4A. The gel imaging results show the presence of an intense band in lane 2, migrating slightly above the 2 kb marker in comparison with the 1 kb DNA ladder in lane 1 confirming successful RE digestion. However, the remaining undigested plasmid could potentially lead to the generation of unwanted transcripts during IVT. Accordingly, this would result in the production of non-specific RNA transcripts larger than the expected transcript interfering with our carrier-based capture strategy discussed earlier (cf. section 5.3.1).

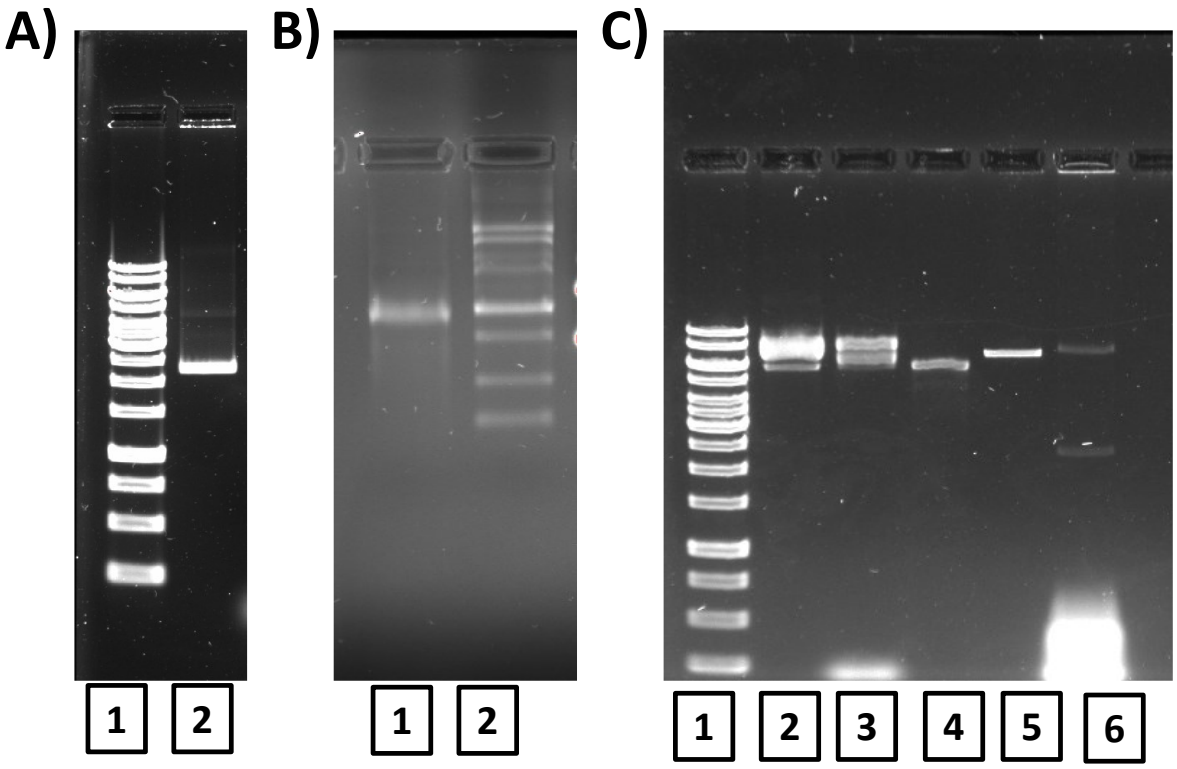

Figure 5.4 Agarose gel electrophoresis results of (A) linearised UCA1 2.3 kbp DNA obtained through RE digestion of the acquired plasmid; lane 1 contains Thermo 1 kb ladder, while lane 2 contains the digested plasmid before purification (1% agarose in 1× TAE, running parameters: 80V for 49 minutes) (B) 1% agarose 0.5× TBE gel run at 80V for 60 minutes to confirm the generation of UCA1 transcript, lane 1 is the generated RNA transcript and lane 2 is NEB ssRNA ladder C) M13 DNA construct with capturing probe , lane 1 is Thermo 1kb GeneRuler, lane2 and 3 are circular ss M13mp18 scaffold and ds M13mp18 RF, whereas lane 4 and 5 represent linearised single stranded ss M13mp18 plasmid and ds M13mp18 RF, finally, lane 6 denotes for unbound capturing probe mixed with linearised UCA1 plasmid (1% agarose in 1xTAE, running parameters: 80V for 49 minutes)

To remove the residual undigested plasmid, the band of interest was further purified by gel extraction. Furthermore, using UV-Vis spectrophotometry revealed final concentration of linearised dsDNA of ~ 15 ng/µL.

Following plasmid linearization, in vitro transcription (IVT) was performed as described in Chapter 2 (Section 2.1.4). The integrity and the size of the transcribed UCA-1 lncRNA (final concentration =236 ng/µl) were assessed using gel electrophoresis (1% agarose 0.5x TBE

native gel). As shown in Figure 5.4B, lane 1 presented a distinct band at approximately 3000 nt, while lane 2 contained the RNA ladder for reference. The observed band size aligns with the expected length of the transcribed UCA-1 lncRNA. It is worth noting, that native gel electrophoresis was used for RNA characterization rather than employing denaturing gel electrophoresis. This choice was partly due to the unavailability of denaturing gel electrophoresis within our research group and considerations for minimizing environmental contamination. Additionally, native gels are widely employed to characterize RNA transcripts >1000 nt.[44] Although native gels provide useful information about sizes of RNA or DNA, the presence of undenatured secondary structures within RNA and ssDNA can drastically affect migration patterns through the gel. On the other hand, longer transcripts migration is less affected in which the case is more prominent with shorter transcripts (<200 nt). Therefore, native gels can be valuable for characterising lncRNA, supported by the large size of the transcript (>2 kilobases).[44]

After confirming the size of the target transcript, successful formation of hybridised M13 carrier DNA was further assessed by gel electrophoresis, as shown in Figure 5.4C. Accordingly, the 2nd and the 3rd lanes present the undigested plasmids of ssDNA M13mp18 and ds M13mp18 RF, respectively. Both lanes are showing 3 multiple bands that are slightly merged. These bands are usually expected due to the different migration patterns of complex formed conformations associated with the plasmids.[45] Correspondingly, the supercoiled form is more twisted and expected to migrate faster on the gel. In the 2nd lane, the lowest band (between 6 kb and 5kb), present the supercoiled circular ssDNA that migrates faster within the gel than the lowest band in the 3rd lane for circular dsDNA. For plasmids in general, the nicked or 'relaxed' conformation is considered the slowest. Interestingly, linear confirmation exhibits a

migration pattern similar to the nicked confirmation, with a slightly faster migration rate. This behaviour explains the submerged bands on the top for both the 2nd and the 3rd lanes.[45]

Moving to the next lanes in Figure 5.4C, the 4th and 5th are the linearised ssDNA M13mp18 and ds M13mp18 RF, respectively. The presence of a single band in each lane indicates successful linearisation of both plasmids, with respect to the faster migration ratio of the linearised ss M13 DNA due to its lower molecular weight. Lane 6 displays the results of a mixed DNA sample vial containing both the hybridised M13 carrier DNA and UCA1 linearised plasmid. In fact, the upper band which belongs to the hybridised carrier appears slightly higher in contrast to linear ds M13mp18 RF in 5th lane, confirming the successful hybridisation of the carrier. Moreover, the lower band at 2kb, which corresponds to UCA1 plasmid, is at the same level as presented in the 2nd lane of Figure 5.4A. The presence of a low thick band at approximately below 200 bp is attributed to the excess overhangs and staples due to the nature of the assay. In this manner, it was difficult to completely remove the excess staples/overhangs even after purification with Amicon filters as previously reported by Loh et al.[13] Accordingly, Using a specialised clean up kit like Zymo-a-select a size clean up kit © has substantially removed these fragment as shown in the 2nd lane of Figure 5.5 where the band below 200 bp in the 2nd lane is absent. However, the use of this cleanup kit has significantly lowered the yield of the hybridised carrier below 10%. Arguably, Ultra High-performance liquid chromatography (UHPLC) or column-based purification was previously suggested.[46] Nevertheless, this comes at very high cost and further complicates carrier synthesis steps. Interestingly, the presence of short fragments band was found not to affect the translocation results, supported by the fact that the 7.2 kbp carrier has a distinct dwell time, which is significantly slower in comparison with these short strands that exhibit faster translocation

rates.[13] In conclusion, the gel electrophoresis results indicate that the bands formed at each step for designing DNA construct were aligned with the correct size against the reference ladders.

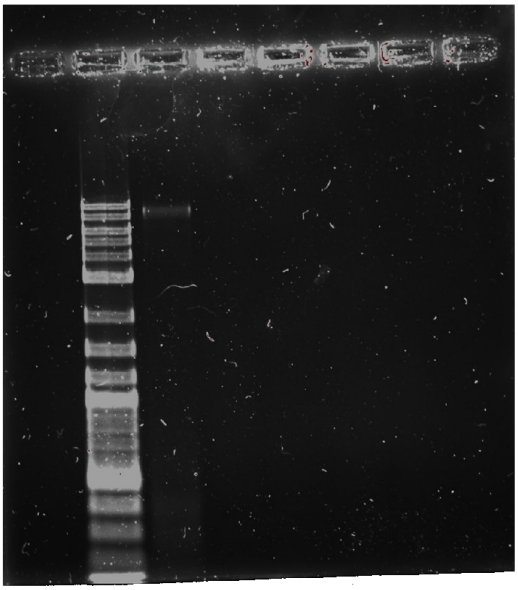

Figure 5.5 : 1% agarose gel electrophoresis of purified M13 construct with capturing probe using Zymo-a-select a size clean up kit. Lane 1 is NEB 1kbp ladder and lane 2 is M13 purified construct with Zymo-a-select a size clean-up kit, the light band at approximately 8 kbp.

Since gel images only provides information on DNA sample size rather than structure, single entity imaging is further devised to characterise the carrier. Thus, AFM imaging was performed to inspect the positioning of the capture probes (two long overhangs) within the carrier. Earlier work by the group has demonstrated good agreement between the predicted positions of the capture overhangs and those observed in AFM images of the carrier. In the line with this, AFM imaging of the carrier was replicated to ensure that the complementary target sequence does not interfere with the position of the capture probes within the carrier. Furthermore, the imaging adopted the same sample preparation, used AFM probes and imaging parameters (refer to Chapter 2 Section 2.3 for further details).

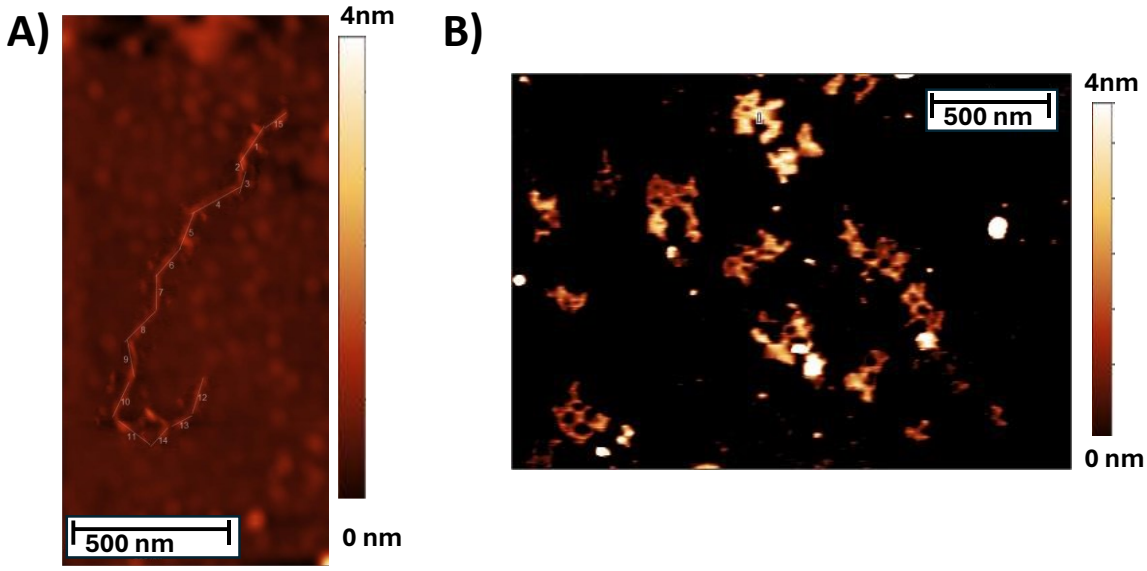

Figure 5.6 Representative AFM micrograph of (A) unbound M13 construct and (B) UCA1 in-vitro transcribed lncRNA using tapping mode while imaging in air.

Briefly, the buffer recipe used for DNA imaging included adding EDTA to deactivate nucleases and stabilise the DNA, while HEPES buffer was added to enhance DNA molecule spreading, adhesion, and to control pH.[47] To facilitate adsorption to the surface, Mg^{2+} was employed as a divalent cation. Mica was chosen as a standard substrate for DNA imaging due to the smoothness of its atomic layer, which does not interfere with the cantilever's tip while imaging DNA.[48] Figure 5.6A shows an AFM micrograph of M13 construct with ss overhangs. Since we have only one image we proceeded analysis on it. Accordingly, the DNA construct height was estimated around 1.8 nm ± 0.3 nm using measurement tool in Gwyddion software,[49] the tool was applied at different positions within the carrier shown in the AFM micrograph. The length of the carrier was also estimated to be approximately 1640 ± 3 nm. Based on those numbers, the imaged carrier appears shorter than calculated full length for 7222 (2455 nm). This behaviour has been previously reported due to DNA confirmational changes introduced by using Mg^{2+} as an intercalating agent aids in DNA adsorption.[13, 50] Multiple protrusion were also

observed in the image with an approximated length of 59 ± 10.3 nm. As we were not sure where is the accurate position of these protrusions. Based on previous work by the group and according to estimated length, the length of the measured protrusion was longer than the reported average length by Loh et al. which incorporated successful characterization of the carrier. However, they discussed the potential of ssDNA folding of short overhangs which does not correlate with the predicted length.[13] Further, explaining the variable lengths of ss overhangs would be difficult due to sequence based property, which introduce an extra probability for the formation of complex secondary structures.[50] Despite that, the obtained image can be used as supplemental evidence for the formation of DNA construct (faint curved line) with the desired capturing probes (resembled as protrusions in the carrier within the image).

Moving to Figure 5.6B, an AFM micrograph was also obtained for the IVT-transcribed target (UCA1 lncRNA) using a different buffer recipe (40 mM HEPES, 1xTE buffer and 10 mM $NiCl_2$ as imaging buffer).[51] The micrograph shows an irregular dispersion of aggregates or folded structures on the mica surface. Statistically, the average height of these structures was estimated around 1.6 ± 0.15 nm, while the average diameter was estimated by taking the longest distance between the two starting points of the structure (d_{max} = 458 ± 38.54nm). Given that the approximate length of a single nucleotide in ssDNA is 0.68nm, and considering the length of a single nucleotide is similar to the length of nucleotide in RNA, the predicted total length of UCA1 lncRNA (2314 nt) is 1573.5 nm excluding folded state.[52] However, the irregular structures in the AFM micrograph appear much shorter than the predicted length of RNA. Irregular structures and shorter lengths can stem from multiple aspects. First, imaging in air might introduce a condensation effect which in turn force RNA molecules to form

aggregates. Additionally, the choice of Ni^{2+} which was added to speed up and strength the adsorption of UCA1 lncRNA to the mica surface, since RNA relatively has less stability in comparison with DNA.[51] However, this increases the risk of condensation and aggregation. Second, as previously discussed with ssDNA, RNA might manifest different complex secondary structures including loops, hairpins or pseudoknots that appear like blobs or aggregates while imaged in air.[53-55] As stated above, these irregular structures appear much shorter than RNA, indicating a high possibility of secondary structure formation.

To investigate the potential for the formation and the stability of a complex secondary structure, RNAfold was used as an online implementation for predicting RNA secondary structure.[56, 57] The main functionality provided by this tool is the estimation of the minimum free energy (MFE) for the most stable secondary structure based on base pairing potential, partitioning function (provides further information of possible RNA structure by means of thermodynamic ensemble) and mean centroidal structures (represent overall shape of the structure in terms of consensus structure).

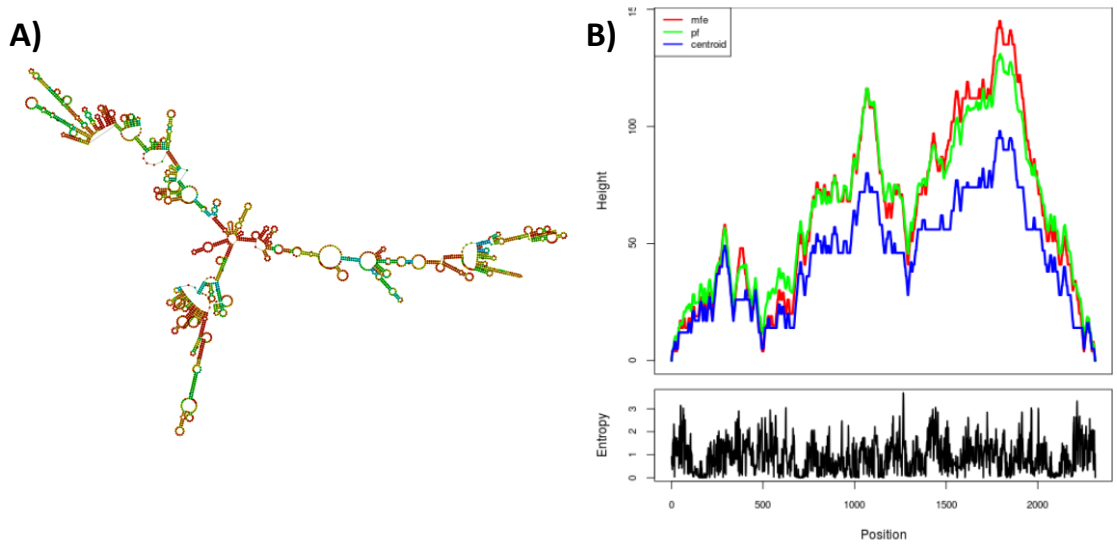

Figure 5.7 (A) predicted secondary structure of UCA1 target by RNAfold (B) mountain plot (MP) generated for the predicted structures by means MFE energy structure (MFE, red), base paring partitioning (pf, green) and centroidal structure(blue) ; position refers to the start and end of UCA1 sequence of bases while height refers to the energy of the formed structure. Plotted lines have three features: small sub-peaks represent the formation of hairpins, flat lines are loops and slopes are helices produced by base pairing, entropy refers to stability.

Figure 5.7A, provides a visual representation of the most probable secondary structure. Looking at the secondary structure presented in the figure, UCA1 lncRNA appears to possess a strong secondary structure. This is further supported by the formed complex features within the structure including loops, harpins and stable helices. Figure 5.7B shows the mountain plot (MP) which represent the energy and shape of those formed features within the RNA structure as function of position. According to RNAfold manual, an increasing/decreasing slope indicates the start of helix formation/closure. If this slope then approaches a small sub-peak, a hairpin structure is expected at the indicated regions. Finally, flat lines indicate the absence of base pairing, which mostly correlate to the formation of a loop. The higher the energy (height) is, the more stable the feature is at that position. [56, 57] In conjunction with height, Positional entropy further shows how the likelihood of base pair formation at a certain segment within RNA sequence. [56, 57] Supposedly, region~ 250-500nts at the beginning of the sequence and

region~1500-2000nts represent strong secondary structure's segments with lower entropy. It can also be noticed from the MP plot that there is an increased frequency of subpeaks (hairpins) and slopes (helices) compared to the frequency of flat lines (ss loops). Consequently, the secondary structure of UCA1 lncRNA is associated with higher stability since ss loops are usually less stable and favourable binding sites for nucleases. Accordingly, the irregular appearance of the UCA1 lncRNA transcript in the AFM micrograph can be explained by the high probability of strong secondary structure formation as provided by the computational analysis.

5.3.4. Carrier and Target Translocation Experiments

To define successful or unsuccessful capture of UCA-1 lncRNA by the proposed method, it was crucial to characterise the hybridised carrier and the free target by nanopore DNA translocation prior to the binding experiment. Accordingly, separate experiments for each were conducted using the same experimental setup described in chapter 2 section 2.2.5 and chapter 3 Section 3.3.3. Figure 5.8 presents effective current plot (ΔI_e vs τ) plots for the hybridised M13mp18 carrier ($C_{DNA,\ out}$ = 116 pM, G_{pore} = 21.5 nS in 4M LiCl 1xTE buffer at V_{bias} = ±0.5V). Accordingly, Panel A does not show any DNA cluster at V_{bias} = +0.5 V, rather only events with scattered τ low ΔI_e(< 50 pA). As stated in Chapter 3 (section 3.3.3 and 3.3.4), this is the background noise resulted from the baseline current, where no DNA was added inside the pipette initially at V_{bias} = +0.5 V. On the other hand, in Panel B when the bias was reversed to V_{bias} = -0.5 V, a DNA related cluster emerges at approximately τ = ~1.2ms with ΔI_e ranges from 100 pA to 200 pA. This cluster is related to the M13 carrier DNA translocation having a

translocation duration comparable to 7 kbp DNA cluster observed in chapter 3 (see chapter 3, Section 3.3.5).

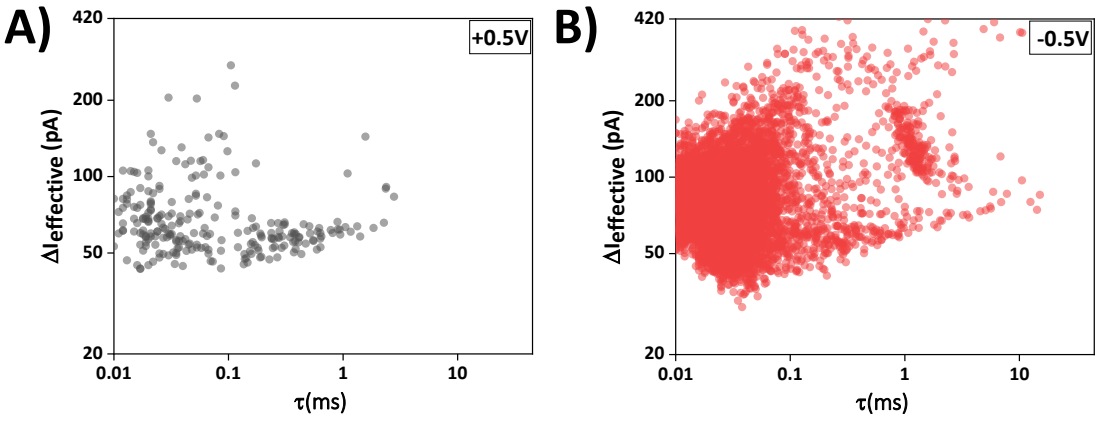

Figure 5.8 Scatter plots of ΔI_e vs τ for the hybridised M13 carrier DNA ($c_{DNA,\ out}$ = 116pM and G_{pore}=21.5nS in 4M LiCl 1xTE) at A) V_{bias} = +0.5 V and B) at V_{bias} = -0.5 V.

At shorter τ (0.15-0.1ms), we can observe a dispersed and sparse cluster of events with comparable ΔI_e values (100 - 200 pA). The observation of these events at shorter τ can be due to the translocation of leftover short DNA fragments into the pipette along with the carrier. In the line with gel electrophoresis results discussed in section 5.3.3, it was difficult to fully remove excess overhangs and staple strands. Thus, it is expected to observe their translocation at shorter duration as documented in the past by our research group and the Keyser group.[13, 37]

As we described in Chapter 3, we performed the experiment in an unload (at positive bias) and load fashion (At negative bias). DNA reverse translocation was observed in the next applied positive bias (V_{bias} = +0.8 V) due to the accumulation of DNA inside the pipette. Figure 5.9 shows effective current plots at V_{bias} = ± 0.8 V. Obviously, once the pipettes was unloaded at V_{bias} = +0.8 V, a DNA related cluster emerged approximately at ~2ms (Figure 5.9A). Again,

DNA-related cluster emerges at approximately ~ 1 ms was formed after the bias was reversed at V_{bias} = -0.8 V (Figure 5.9B).

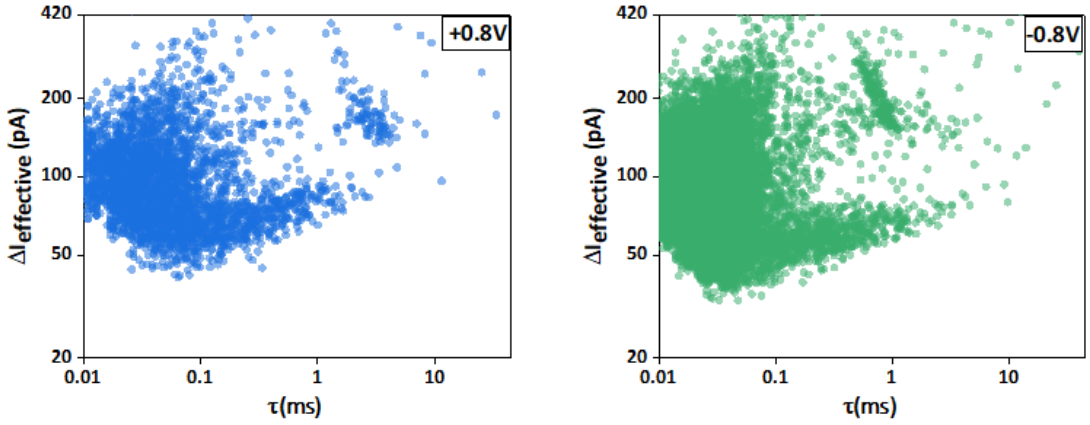

Figure 5.9 Scatter plots of ΔI_evs τ for the hybridised M13 carrier DNA ($c_{DNA,\ out}$ = 116pM and G_{pore}=21.5nS in 4M LiCl 1xTE) at A) V_{bias} = +0.8 V and B) at V_{bias} = -0.8 V.

After the data was plotted, pulling DNA-related cluster using DBSCAN at -0.5 V has further revealed 168 total number of events. We then proceeded to extract linear events using PCA and k-means (cf. Chapter 2, Section 2.2.5 and Chapter 3, section 3.3.3). This resulted in the selection of 78 events as linear, which were later subjected to subevent analysis. Figure 5.10 presents ΔI_evs τ at V_{bias} = -0.5 V, where data points belong to extracted linear events are highlighted in red.

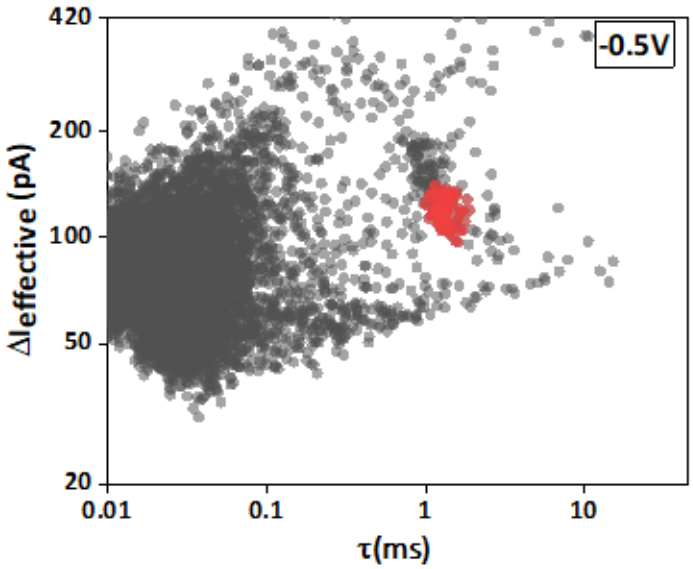

Figure 5.10 Scatter plots of ΔI_e vs τ for the hybridised M13 carrier DNA ($c_{DNA,\ out}$ = 116pM and G_{pore}=21.5nS in 4M LiCl 1xTE V_{bias} = -0.5 V), extracted linear events are highlighted in red.

Due to the absence of full reference translocation data for the target and the bound carrier experiments in the reverse direction, we have chosen to further perform subevent detection for the data recorded at negative bias and further explore the reverse direction in a future work beyond this thesis.

Specifically, the carrier has been modified with two capture probes (overhangs) at 0.3 and 0.7 the length of the carrier DNA, respectively. Both overhangs were single stranded. Thus, the approach to identify their normalised position within the carrier was adapted from the previous work by Loh et al.[13] Eventually, choosing sub-event cutoff was articulated based on the median of the ΔI_{Max} values for the filtered events ($\Delta I_{Max,\ median}$ = 200 pA at V_{bias} = -0.5 V). This criterion identified 142 subevents, with 65 events exhibiting a detected subevent feature. Figure 5.11A illustrates the count of events categorised by the number of subevents in each individual event. Notably, 29 out of 65 events have two subevents, while 14 events have a single subevent, and 23 events contain three or more subevents (cf. 2.2.5.5 and 4.4.2).

From the carrier design strategy, it is assumed that each event has two subevents with comparable magnitudes and positioned at relative position = 0.3 and 0.7. Figure 5.11B shows the sub-event duration τ_{se} versus the relative sub-event position within an event (0: event start, 1: event end), with sub-event magnitude color-coded from < 52 pA (blue) to > 336 pA (red). Histograms for relative sub-event position are shown on the top of the plot (black). From the scatter shown, we have two clustered subevents distributions. The first one is centred near the start an event (15% of total subevents), with relatively short duration that increases to longer durations (50 µs -110µs) and subevent magnitudes values ranging from 50pA to 194pA. Conforming to, positioning near the beginning is usually attributed to DNA initial excursion which is an expected behaviour of DNA during the translocation process through the pipette.[13, 37, 38] Further, Figure 5.12 presents current time traces of sample of events with detected excursion at the beginning of the carrier highlighted in grey. It is worth mentioning that these traces also show additional subevents spikes in the middle of the carrier. Accordingly, an event detected with three subevents is likely presenting a DNA molecule with excursion and two capture probes (explained next).

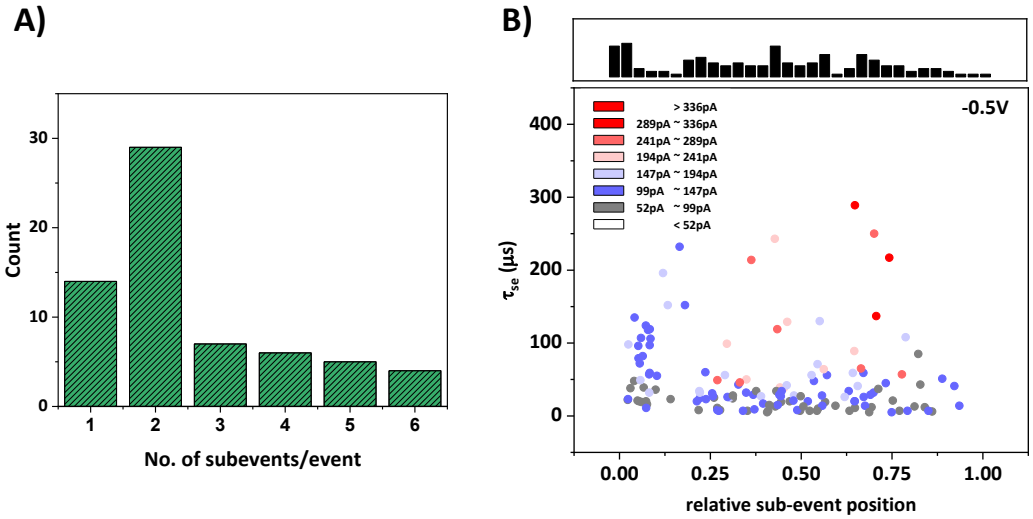

Figure 5.11 M13 unbound carrier (A) events count categorized by the number of subevents detected using a ΔI_{Max} cutoff of 320 pA at V_{bias} = +0.8 for extracted linear events (B)subevent duration τ_{se} vs. relative subevent position within an event (0: event start, 1: event end). Colour code: sub-event magnitude, from < 52 pA (white) to > 363 pA (red). Top side: histograms of the relative subevent position.

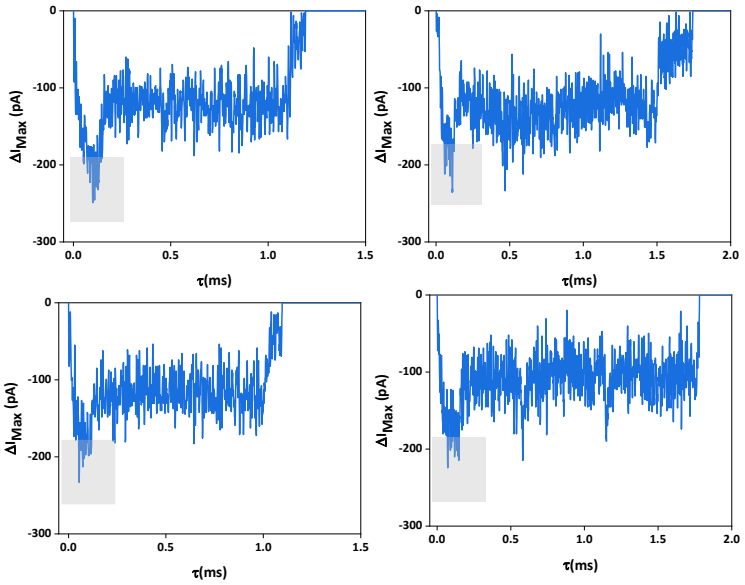

Figure 5.12 : Current-time traces for sample events with DNA excursion sub-event feature (highlighted in grey) for the unbound M13 carrier.

The other subevent distribution in Figure 5.11B lies within the region of interest, where the proposed modifications located specifically, at 0.3 and at 0.7 of the carrier length. Accordingly, the scatter at this region stands for bottom population which has very short duration (20µs-50µs) and ranging from low to mid subevent magnitude (50pA-289pA). As shown in the histograms in Figure 5.11B, if we excluded the subevents at the beginning (belongs to DNA excursion), there is an increased occurrence of subevents that lies between 30%-70% of the carrier length. Hence, this is the region associated with the position of the two capture probes. As a result, this provides evidence that the hybridisation of the carrier with the two ss capture probes was successful.

Sample events with variable intensity are presented in Figure 5.13. The variability in intensity is expected due to the flexibility of the ssDNA capture probes which might adopt different conformations/secondary structures that interacts differently with the pore while entering the pipette.[13] The increase in subevents intensity was plotted in sequence as shown in the top six plotted current time traces in figure 5.13.

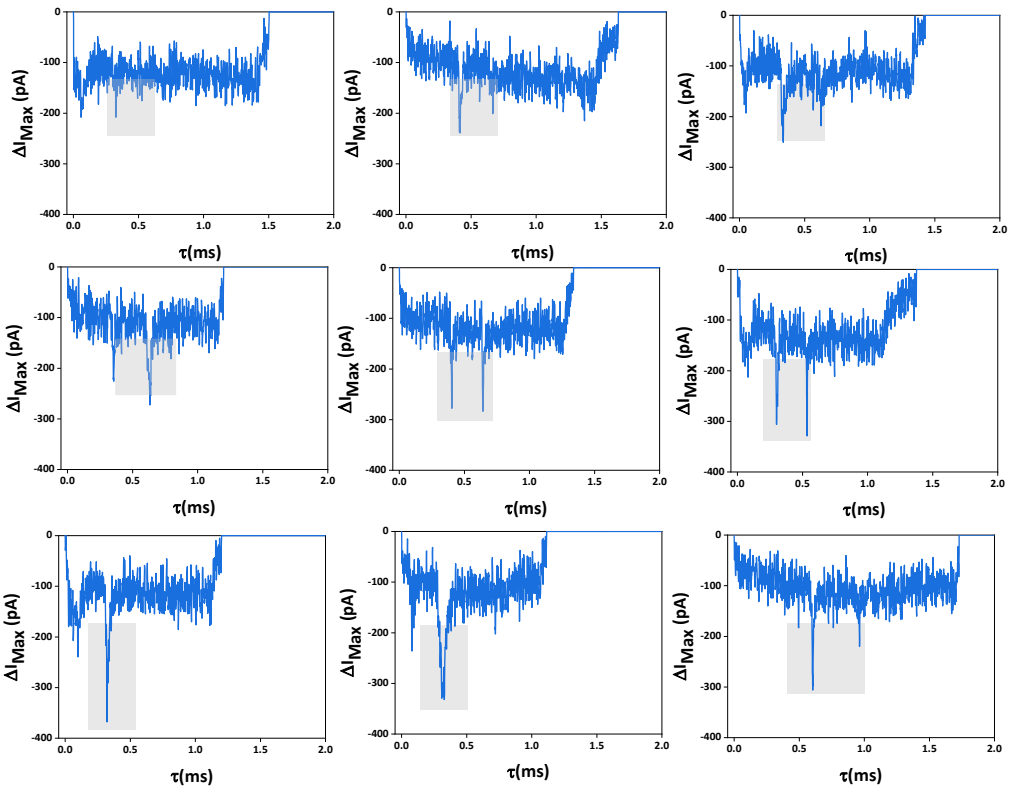

Figure 5.13 Current-time traces for sample events with two subevents feature at the middle of the unbound M13 carrier, both subevents are highlighted in grey within each plot and all plots are shown in sequence based on intensity of their detected subevents from top left to bottom right (please refer to the main text)

Additionally, the bottom three traces in Figure 5.13 shows an intense subevent feature with slightly longer durations (τ_{se} < 110 µs). These subevents exhibiting magnitudes > 2 ΔI_e of an DNA event (assuming the current level of DNA event at 100 kHz is around 180 pA–200 pA) is mostly due to the presence of DNA knots, as the frequency of these events is low. [58, 59]

Finally, a small population of subevents with durations ≥ 200 µs were also investigated and presented in Figure 5.14, these events are more likely due to strong interaction with the pore. Since, the size of the pore is approximately 11 nm, while the length of ssDNA overhang 71 nm (considering unfolded state). It could be assumed that these subevents is an indication of the

translocation of a unravelled tertiary structure, which is large enough to interact with the pore, leading to an increased frictional force, eventually increases τ_e.[13, 60] To investigate DNA-interaction with the pore, charge of subevents (q_{se}) vs their normalised duration (τ_{se}/τ) was plotted as displayed in Figure 5.15. From the scatter plot, a positive correlation is observed between the subevent charge and the subevent duration for intense subevents (at $\tau_{se}/\tau > 0.1$). Accordingly, intense subevents (> 350 pA) with long τ_{se} (≥ 200 µs) indicating a strong interaction with the pore during the translocation of the hybridised carrier.[13] However, these subevents are rare. Hence, we can conclude that the subevent analysis provides strong evidence of the successful formation of the hybridised carrier with the two capture probes.

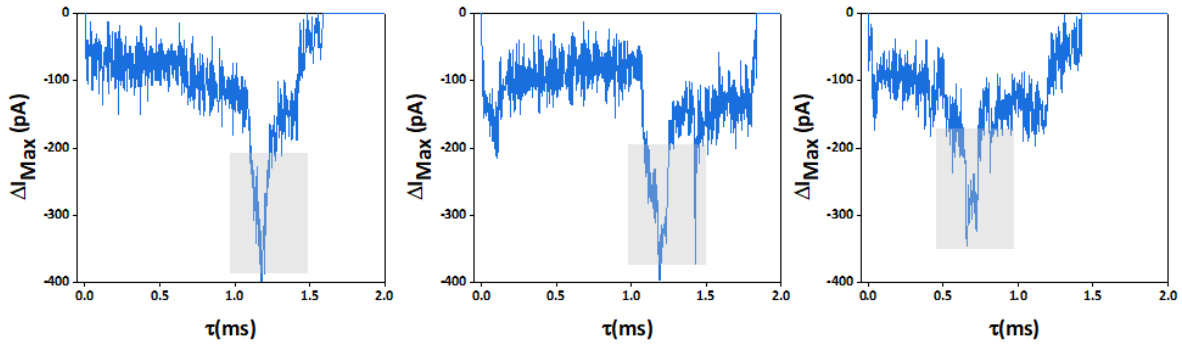

Figure 5.14 Current-time traces for events with long duration (τ_{se}) sub-events feature (highlighted in grey) for the unbound M13 carrier.

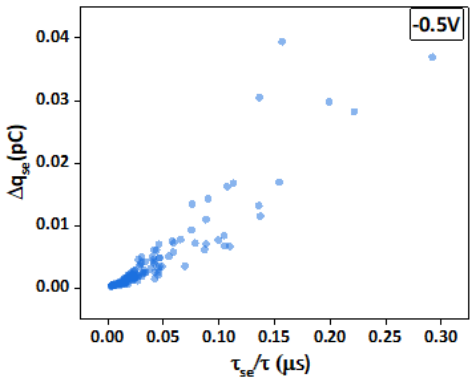

Figure 5.15 scatter plot of subevent charge (Δq_{se}) vs τ_{se}/τ for detected subevents of the carrier at V_{bias} = -0.5 V.

After characterizing the carrier, the translocation experiment of the in-vitro transcribed UCA1 lncRNA was performed and analysed (cf. Chapter 2, Section 2.2.5). Accordingly, Figure 5.16 represents effective currents (ΔI_e) vs τ plot at V_{bias} = -0.5 V ($c_{RNA,\ out}$ = 100 pM and G_{pore} = 35.7 nS in 4M LiCl 1xTE).

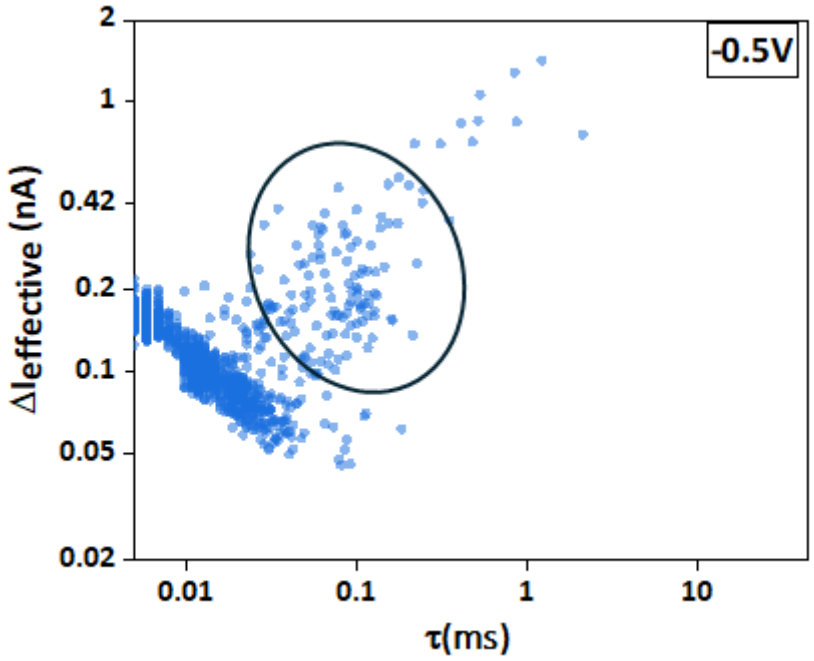

Figure 5.16 Scatter plot of ΔI_e vs τ for in-vitro transcribed UCA1 lncRNA translocation experiment ($c_{RNA,\ out}$ = 100 pM and G_{pore} = 35.7 nS in 4M LiCl 1xTE at V_{bias} = -0.5 V)

The scatter plot shows a dense cluster at τ ≈ 10µs with low ΔI_e values (< 100 pA) and another slightly sparse cluster (inside the eclipse) at τ ≈100µs with high and variable ΔI_e values (0.20 pA to 1000 pA). The shorter duration cluster corresponds to the background noise or short fragments translocation, as shown earlier with the carrier translocation experiments. However, the longer duration sparse cluster with variable current response is possibly related to the RNA translocation (shown inside the eclipse). Furthermore, τ ≈ 100µs indicates very short translocation time in comparison with the 7.2kb carrier DNA (1.2 ms), although RNA size is around 2314 bases (~1500 nm). This suggests two possible scenarios.

First, as supported by the AFM micrographs and computational simulation shown in section 5.3.3, RNA possesses a strong secondary structure which shortens the observed length compared to the calculated length. This indicates that the RNA does not fully unravel during translocation which consequently decrease the observed duration (τ). [28, 38] Moreover, the variable current response indicates the presence of RNA-pore interaction. A second scenario suggests incomplete translocations or colliding events followed by strong interaction pore.[60] As observed in the plot, a spread of data points with higher current magnitude (0.8-1nA) are observed at longer durations. This implies the presence of RNA-pore interaction, which is again explained by the strong secondary structure formation as observed in the carrier events in Figure 5.14.

In conclusion, the characteristic τ of the unbound UCA1 lncRNA is much shorter than the carrier DNA. Accordingly, this control experiment was conducted to confirm that the unbound UCA1 lncRNA does not overlap with the carrier DNA translocation characteristics.

5.3.5. Target Bound Experiment

The main aim of this chapter was to evaluate the capability of combined carrier and nanopipette sensor to detect lncRNA without fragmentation by binding to both ends of the RNA, thereby reducing the influence of the secondary structure. First, to assess the ability of the carrier to bind with RNA, gel electrophoresis was performed (cf. Chapter 2, section 2.1.1). Carrier binding with RNA experiment was conducted by incubating 25 µL of the carrier with 2 µL of the IVT- UCA1 transcript at 40^{O}C for 16 hrs in the thermocycler (cf. Chapter 2, section 2.1.3.4), the long duration of incubation was chosen to increase the chance of binding, since the capture probes are long (>50nts) and UCA1 has a complex secondary structure, which in

turn need longer time to hybridize with each other. Shorter incubation time (<4 hrs) revealed no new visible bands on the gel.[61, 62] Figure 5.17 presents the results of 1 % agarose TAE gel of the RNA-carrier bound experiment. Thus, the 1st lane shows 1kb ladder, the 2nd lane has the hybridised carrier (unbound) sample, and the 3rd lane has the RNA-carrier bound sample. Particularly, the 3rd lane shows the formation of 2 faint bands on the top between 10 kbp-6 kbp, and smearing region starts approximately at 2 kbp in accordance with the ladder. As indicated by the white arrows, the smearing at 2 kbp corresponds to the excess RNA, the upper faint band corresponds to the carrier, and on the top at approximately 10kbp, a new band formed is expected to be the RNA-carrier hybrid compared to the unbound carrier in the 2nd lane.

Prior to RNA binding experiment, excess fragments in the unbound carrier were removed using Zymo-Select-a-Size clean up kit as provided in section 5.3.3. Hence, this will reduce the competition between the remaining fragments and the unbound carrier for binding with RNA.

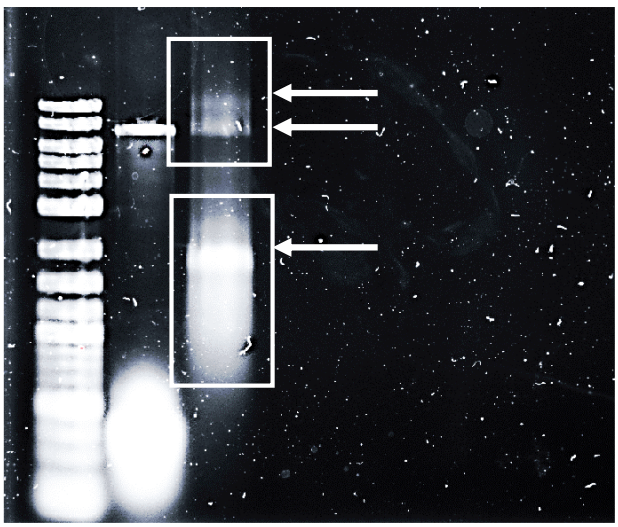

Figure 5.17 1% TAE agarose gel electrophoresis of RNA -bound M13 carrier with capturing probe using Zymo-a-select a size clean up kit, lane1 is NEB 1kbp ladder, lane 2 is M13 unpurified and unbound construct and lane 3 is the purified M13 construct with Zymo-a-select a size clean up kit, which was incubated with RNA after purification.

As stated earlier, the translocation experiment followed the same experimental setup (4M LiCl 1xTE buffer). Figure 5.18 presents effective currents (ΔI_e) vs τ plot for carrier-RNA bound at V_{bias} = -0.5 V ($c_{DNA,\ out}$ = 50 pM and G_{pore} = 24.4 nS in 4M LiCl 1xTE).

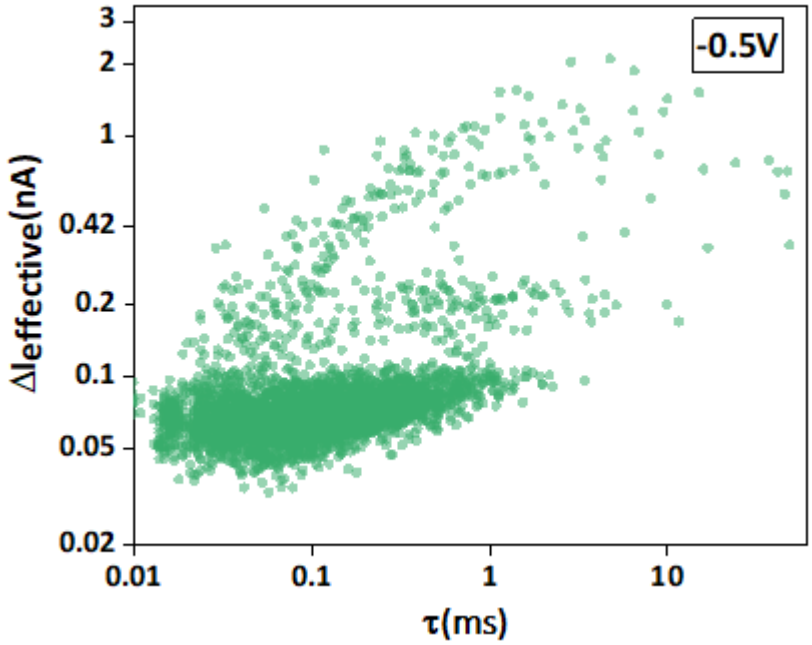

Figure 5.18 Scatter plot of ΔI_e vs τ for carrier-UCA1 bound translocation experiment ($c_{DNA,\ out}$ = 50 pM and G_{pore} = 24.4 nS in 4M LiCl 1xTE, V_{bias} = -0.5 V).

From the effective current scatter plot, three clusters are formed. First, a low intensity cluster with current magnitude ranging from 50pA to 80pA, this is the background noise as discussed in the free RNA experiment, with relatively sparse duration (10 µs-1 ms). The second cluster presents the RNA cluster at ~100 µs with ΔI_e values ranging from 0.20 pA to 1000 pA. This is in line with what observed with the free RNA experiment, which indicates the translocation of excess RNA. However, this cluster expands to longer durations(>10ms) compared to the free RNA experiment. Interestingly, a small thinly distributed cluster emerges at ~0.3-1ms with 100-200pA current magnitude. Events within this cluster correspond to the carrier translocation region. Nevertheless, the cluster has small number events (99 events) added to

the widespread location of datapoints in terms of duration, it was harder to separate these events to linear and folded, or subject them later to subevent analysis. It was worth noting that this experiment was run for longer time due to low concentration of the carrier but the RNA as it is in excess it blocked the pore after shorter acquisition time. This was an ongoing problem encountered through multiple experiments with intact RNA.

Based on the results above, successful binding of the RNA with the carrier was not fully approved. This methodology suffered from RNA-pore interactions which resulted in reduced events recorded. Further, this interaction has impacted the operational span of the pipette. These observations highlight the challenge of detecting intact lncRNA by the mentioned strategy. Accordingly, more work should be done to optimize the assay design or considering other capture strategies.

5.4. Conclusion

In conclusion, this chapter explored an alternative application for carrier-based nanopore as a biosensing strategy for detecting UCA-1 lncRNA, a potential biomarker for bladder cancer. Nanopipettes were able to characterise both the unbound carrier and RNA transcript, solely. However, the RNA-carrier bound experiments showed statistically a low event count. As a result, subevent analysis was not feasible to determine the position of the binding. Moreover, RNA-carrier bound translocation signals exhibited variations and complex patterns that were difficult to interpret. Even though the results are somewhat inconclusive, possibly due to pore RNA interaction and pore clogging, new directions could be investigated further in the future. For example, to reduce interaction with the pore, surface modification could be explored.

Another possibility could be due to the strong secondary structure of RNA-carrier complex that is much larger than pore dimension, hindering complete translocation. Accordingly, to help interpret the behaviour of the carrier-RNA bound structure, we would probably think about the determination of target-bound secondary structure in liquid. This can improve our understanding about the dimension and the shape of the 3D structure of the complex. Hence, it can be concluded that it is hard to detect lncRNA using our above-mentioned carrier-based approach. Thus, an alternative method of capture should be devised.

5.5. References

1. P. C. Champe, R. A. Harvey and D. R. Ferrier, *Biochemistry*, Lippincott Williams & Wilkins, 2005.
2. C. J. David and J. L. Manley, *Genes Dev*, 2010, **24**, 2343-2364.
3. G. J. Goodall and V. O. Wickramasinghe, *Nat Rev Cancer*, 2021, **21**, 22-36.
4. M. N. Islam, M. K. Masud, M. H. Haque, M. S. Al Hossain, Y. Yamauchi, N. T. Nguyen and M. J. A. Shiddiky, *Small Methods*, 2017, **1**, 1700131.
5. C. C. Pritchard, H. H. Cheng and M. Tewari, *Nat Rev Genet*, 2012, **13**, 358-369.
6. A. Kappel and A. Keller, *Clin Chem Lab Med*, 2017, **55**, 636-647.
7. Y. Qian, L. Shi and Z. Luo, *Front Med (Lausanne)*, 2020, **7**, 612393.
8. C. Badowski, B. He and L. X. Garmire, *NPJ Precis Oncol*, 2022, **6**, 40.
9. Z. Ding, W. Ying, Y. He, X. Chen, Y. Jiao, J. Wang and X. Zhou, *Medicine (Baltimore)*, 2021, **100**, e24805.
10. F. Wang, X. Li, X. Xie, L. Zhao and W. Chen, *FEBS Lett.*, 2008, **582**, 1919-1927.
11. H. Wang, Z. Guan, K. He, J. Qian, J. Cao and L. Teng, *Oncotarget*, 2017, **8**, 64638-64650.
12. M. Wang, Z. Zhang, D. Pan, Z. Xin, F. Bu, Y. Zhang, Q. Tian and X. Feng, *Biosci. Rep.*, 2021, **41**.
13. A. Y. Y. Loh, C. H. Burgess, D. A. Tanase, G. Ferrari, M. A. McLachlan, A. E. G. Cass and T. Albrecht, *Anal. Chem.*, 2018, **90**, 14063-14071.
14. N. A. Bell and U. F. Keyser, *Nat Nanotechnol*, 2016, **11**, 645-651.
15. R. Stark, M. Grzelak and J. Hadfield, *Nat Rev Genet*, 2019, **20**, 631-656.
16. R. Yelagandula, A. Bykov, A. Vogt, R. Heinen, E. Ozkan, M. M. Strobl, J. C. Baar, K. Uzunova, B. Hajdusits, D. Kordic, E. Suljic, A. Kurtovic-Kozaric, S. Izetbegovic, J. Schaeffer, P. Hufnagl, A. Zoufaly, T. Seitz, Vcdi, M. Fodinger, F. Allerberger, A. Stark, L. Cochella and U. Elling, *Nat Commun*, 2021, **12**, 3132.
17. M. Soler, C. S. Huertas and L. M. Lechuga, *Expert Rev Mol Diagn*, 2019, **19**, 71-81.
18. L. G. Carrascosa, C. S. Huertas and L. M. Lechuga, *Trends Analyt Chem*, 2016, **80**, 177-189.
19. S. Fang, H. J. Lee, A. W. Wark and R. M. Corn, *J. Am. Chem. Soc.*, 2006, **128**, 14044-14046.
20. H. Xie, Y. H. Yu, F. Xie, Y. Z. Lao and Z. Gao, *Anal. Chem.*, 2004, **76**, 4023-4029.

21. A. Shaver and N. Arroyo-Curras, *Curr Opin Electrochem*, 2022, **32**, 100902.

22. M. Santhanam, I. Algov and L. Alfonta, *Sensors (Basel)*, 2020, **20**, 4648.

23. M. Rahman, M. J. N. Sampad, A. Hawkins and H. Schmidt, *Lab Chip*, 2021, **21**, 3030-3052.

24. Y. Xing, A. Rottensteiner, J. Ciccone and S. Howorka, *Angew. Chem. Int. Ed. Engl.*, 2023, **62**, e202303103.

25. W. Yi, C. Zhang, Q. Zhang, C. Zhang, Y. Lu, L. Yi and X. Wang, *Top Curr Chem (Cham)*, 2023, **381**, 13.

26. R. Ren, S. Cai, X. Fang, X. Wang, Z. Zhang, M. Damiani, C. Hudlerova, A. Rosa, J. Hope, N. J. Cook, P. Gorelkin, A. Erofeev, P. Novak, A. Badhan, M. Crone, P. Freemont, G. P. Taylor, L. Tang, C. Edwards, A. Shevchuk, P. Cherepanov, Z. Luo, W. Tan, Y. Korchev, A. P. Ivanov and J. B. Edel, *Nat Commun*, 2023, **14**, 7362.

27. T. Albrecht, *Annu Rev Anal Chem (Palo Alto Calif)*, 2019, **12**, 371-387.

28. F. Boskovic and U. F. Keyser, *Nat Chem*, 2022, **14**, 1258-1264.

29. J. Zhu, R. Tivony, F. Boskovic, J. Pereira-Dias, S. E. Sandler, S. Baker and U. F. Keyser, *J. Am. Chem. Soc.*, 2023, **145**, 12115-12123.

30. F. Boskovic, C. Maffeo, G. Patino-Guillen, R. Tivony, A. Aksimentiev and U. F. Keyser, *ACS Nano*, 2024, **18**, 15013-15024.

31. C. Chau, F. Marcuccio, D. Soulias, M. A. Edwards, A. Tuplin, S. E. Radford, E. Hewitt and P. Actis, *ACS Nano*, 2022, **16**, 20075-20085.

32. C. Plesa, N. Van Loo, P. Ketterer, H. Dietz and C. Dekker, *Nano Lett.*, 2015, **15**, 732-737.

33. O. I. Wilner and I. Willner, *Chem Rev*, 2012, **112**, 2528-2556.

34. S. Dey, C. H. Fan, K. Gothelf, J. Li, C. X. Lin, L. F. Liu, N. Liu, M. A. D. Nijenhuis, B. Saccà, F. C. Simmel, H. Yan and P. F. Zhan, *Nat Rev Methods Primers*, 2021, **1**, 13.

35. A. Ptasinska, C. Whalley, A. Bosworth, C. Poxon, C. Bryer, N. Machin, S. Grippon, E. L. Wise, B. Armson, E. L. A. Howson, A. Goring, G. Snell, J. Forster, C. Mattocks, S. Frampton, R. Anderson, D. Cleary, J. Parker, K. Boukas, N. Graham, D. Cellura, E. Garratt, R. Skilton, H. Sheldon, A. Collins, N. Ahmad, S. Friar, D. Burns, T. Williams, K. M. Godfrey, Z. Deans, A. Douglas, S. Hill, M. Kidd, D. Porter, S. P. Kidd, N. J. Cortes, V. Fowler, T. Williams, A. Richter and A. D. Beggs, *Clin Microbiol Infec*, 2021, **27**, 1348 e1341-1348 e1347.

36. H. Zhang, M. Hiratani, K. Nagaoka and R. Kawano, *Nanoscale*, 2017, **9**, 16124-16127.

37. N. A. Bell and U. F. Keyser, *J. Am. Chem. Soc.*, 2015, **137**, 2035-2041.

38. M. Mihovilovic, N. Hagerty and D. Stein, *Phys. Rev. Lett.*, 2013, **110**, 028102.

39. *National Center for Biotechnology Information*, 2024, https://www.ncbi.nlm.nih.gov/nuccore/NR_015379.

40. H.-M. Eun, in *Enzymology Primer for Recombinant DNA Technology*, ed. H.-M. Eun, Academic Press, San Diego, 1996, pp. 1-108.

41. G. Weber, in *Adv. Protein Chem.*, eds. C. B. Anfinsen, J. T. Edsall and F. M. Richards, Academic Press, 1975, vol. 29, pp. 1-83.

42. W. A. Kibbe, *Nucleic Acids Res.*, 2007, **35**, W43-46.

43. E. C. Hulme and M. A. Trevethick, *Br. J. Pharmacol.*, 2010, **161**, 1219-1237.

44. D. C. Rio, M. Ares, Jr., G. J. Hannon and T. W. Nilsen, *Cold Spring Harb Protoc*, 2010, **2010**, pdb prot5445.

45. J. L. Viovy, *Rev. Mod. Phys*, 2000, **72**, 813-872.

46. K. Halvorsen, M. E. Kizer, X. Wang, A. R. Chandrasekaran and M. Basanta-Sanchez, *Anal. Chem.*, 2017, **89**, 5673-5677.

47. Y. L. Lyubchenko, L. S. Shlyakhtenko and T. Ando, *Methods*, 2011, **54**, 274-283.

48. Z. Liu, Z. Li, H. Zhou, G. Wei, Y. Song and L. Wang, *Microsc. Res. Tech.*, 2005, **66**, 179-185.

49. D. Nečas and P. Klapetek, *Open Phys*, 2012, **10**, 181-188.

50. K. H. S. Main, J. I. Provan, P. J. Haynes, G. Wells, J. A. Hartley and A. L. B. Pyne, *APL Bioeng*, 2021, **5**, 031504.

51. A. Cavallo, U. Masullo, A. Quarta, A. Sannino, A. Barca, T. Verri, M. Madaghiele and L. Blasi, *Appl. Sci. (Basel)*, 2020, **10**, 6056.

52. Q. J. Chi, G. X. Wang and J. H. Jiang, *Physica A*, 2013, **392**, 1072-1079.

53. I. V. Novikova, S. P. Hennelly and K. Y. Sanbonmatsu, *Int J Mol Sci*, 2013, **14**, 23672-23684.

54. D. Japrung, A. Bahrami, A. Nadzeyka, L. Peto, S. Bauerdick, J. B. Edel and T. Albrecht, *J. Phys. Chem. B*, 2014, **118**, 11605-11612.

55. R. Spokoini-Stern, D. Stamov, H. Jessel, L. Aharoni, H. Haschke, J. Giron, R. Unger, E. Segal, A. Abu-Horowitz and I. Bachelet, *RNA*, 2020, **26**, 629-636.

56. I. L. Hofacker, W. Fontana, P. F. Stadler, L. S. Bonhoeffer, M. Tacker and P. Schuster, *Monatsh. Chem.*, 1994, **125**, 167-188.

57. R. Lorenz, S. H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler and I. L. Hofacker, *Algorithms Mol Biol*, 2011, **6**, 26.

58. R. Kumar Sharma, I. Agrawal, L. Dai, P. S. Doyle and S. Garaj, *Nat Commun*, 2019, **10**, 4473.

59. C. Plesa, D. Verschueren, S. Pud, J. van der Torre, J. W. Ruitenberg, M. J. Witteveen, M. P. Jonsson, A. Y. Grosberg, Y. Rabin and C. Dekker, *Nat Nanotechnol*, 2016, **11**, 1093-1097.

60. M. Wanunu, J. Sutin, B. McNally, A. Chow and A. Meller, *Biophys. J.*, 2008, **95**, 4716-4725.

61. A. R. Chandrasekaran, L. Zhou and K. Halvorsen, *Trends Biochem. Sci*, 2022, **47**, 367-368.

62. L. Zhou, A. Hayden, A. R. Chandrasekaran, J. Vilcapoma, C. Cavaliere, P. Dey, S. Mao, J. Sheng, B. K. Dey, P. Rangan and K. Halvorsen, *Cell Rep Methods*, 2021, **1**.

CHAPTER 6
CONCLUSIONS AND OUTLOOKS

CONTENTS

Synopsis: This chapter presents a summary of the research conducted in this thesis. The purpose and motivation of this research were mentioned. Then, key findings, novel results and challenges identified from this research are discussed. Finally, future directions and recommendations are suggested.

6.1. Conclusion and Outlooks

Single molecule sensing with quartz nanopipettes has progressed in recent years, enabling high sensitivity detection of biomolecules including DNA, RNA and proteins.[1] The specific properties of nanopipettes have promoted their wider applicability in nanopore labs. Further, these properties included the material it made from which exhibits low-noise, and good surface wettability. Other properties included the ease and the low-cost of their fabrication procedure.[1-5] Interestingly, nanopipettes possess long tapered geometry that converges to the narrowest end of the tip, which constitutes the sensing region. This convergence creates asymmetry between the cis and the trans compartments.[6] Previous work has explored controlled delivery and localisation of ssDNA and dsDNA.[7, 8] Recently, this asymmetry was found to increase τ of linear and functionalised DNA molecules moves out from the trans compartment.[9-12]

In this thesis, nanopipettes were explored as a versatile tool for two distinct purposes:

The first objective aimed at investigating local DNA trapping induced by geometric constraints. Hence, in **Chapter 3** we have demonstrated through a detailed analysis and multiple experiments how asymmetric geometry of the nanopipette can influence the translocation dynamics of kilobase DNA into and out of the nanopipette tip. It appears that the slow transport from the tip region into the bulk of the nanopipette can occur due to combinations of spatial constraints, friction and weakening electric field are important in this area.[13] In this regard, we have initiated a detailed analysis on the translocation dynamics of kilobase linear DNA from both compartments (cis and trans). Further, we exploited variable lengths of DNA standards starting from 4 kbp to 48.5 kbp DNA. As mentioned in the Chapter

2 and Chapter 3, the experimental procedure incorporated adding DNA molecules to the cis compartment (outside the pipette), then, translocate the DNA into the pipette for a relatively prolonged duration as opposed to previous studies, which incorporated short duration of translocation runs. Interestingly, at shorter DNA lengths (4 kbp DNA), once the translocation direction was reversed (from the inside to the outside), we observed a comparable DNA capture rate and increased fraction of linear DNA events. Further, τ for DNA obtained from in-out translocation was longer than that obtained through out-in translocation aligning with previous studies. Moreover, with longer DNA molecules (≥ 7 kbp), we identified the occurrence of new phases during the translocation process, depending on the number of DNA molecules remaining inside the pipette. Thus, we observed a crowded state characterised by abrupt shift in τ after the translocation of certain number of DNA molecules into the pipette. This state was approached faster as DNA length increases. For 10 kbp and 48.5 kbp DNA an additional phase transition was observed marked by a sudden cessation of DNA translocation and blockage of the nanopipette. However, this phase was completely reversable as soon as reverse translocation was initiated. Accordingly, from those results we concluded that DNA is locally trapped near the tip inside the pipette.

This finding has further motivated us to explore a new bioanalytical paradigm in **Chapter 4**. Hence, we conducted two separate experiments where we incubated two different sizes of streptavidin-functionalised spherical gold nanoparticles (5 and 40 nm in diameter), inside a nanopipette. Then, a biotin functionalised 10 kbp DNA was loaded into the pipette and allowed to bind with the particles .Afterwards, DNA-particle capture was assessed by reverse translocation. Toward this end, we have demonstrated the capability of this approach by demonstrating successful particle-DNA binding through detailed subevent analysis. We

further incorporated specific statistical analysis to discriminate between both nanoparticle sizes from the two experiments conducted.

Based on results presented in **Chapters 3** and **4**, future directions can be tailored to either fundamental consideration or further advanced applications. Hence, a fundamental future study can be incorporated to study and resolve the structure of single molecule in a confined space. Nanopipettes can be employed as a device to trap single molecule (ex. DNA) and then the confirmational and structural orientation in the confinement can be further imaged by an analytical method considering atomic details like x-ray crystallography. This would increase our understanding in the future about the effect of confinement on single molecules. On the other hand, advancing solid state nanopore technology through controlling factors affecting DNA translocation is an important aspect for advancing the technology. As we have provided that the volume of the nanochannel is critical for the emergence of crowding effect. We suggest that this aspect could be further explored in depth. Since we used a comparable size of nanopipette (except for the 2nd experiment of 4 kbp DNA), The volume of the nanochannel can be further investigated experimentally to study and conclude the optimal conditions required to achieve local trapping.

Additionally, we can further explore a new design of chip based nanopores with longer channel and slightly converged end. Therefore, from the idea of reverse translocation we can enhance the temporal resolution. Further, we can reduce the critical concentration of DNA molecules and enhancing their capture rate.

Regarding analyte incubation, as we suggested in Chapter 4, we can extend this to method to a real word biomarker such as proteins and RNA. However, an interesting fundamental study

can be further extended on DNA-nanoparticles bound complexes to study their translocation behaviour. For example, the design of DNA particle bound construct with particles remaining inside can for example trap a single DNA molecule inside once this bound construct is reverse translocated. Since the DNA molecule is bound to a large particle from the inside, particle bound DNA molecule remains trapped inside for seconds of time as determined by the user. Accordingly, exploring this aspect can open the door toward high resolution-controlled DNA characterisation.

Unrelated to nanopipette geometry, **the second purpose** of this thesis was to explore the capability of nanopipettes in combination with a carrier DNA strategy to capture long noncoding RNA (lncRNA) without amplification or complex staple design, **Chapter 5**. The work incorporated the use of previously successful assay designed by previous members of the group to detect shorter fragments.[6] Although gel electrophoresis results have shown a potentially RNA-carrier bound complex. However, the associated bound translocation experiments results remained inconclusive with statistically low events recorded hindering successful implementation. Further, from the AFM imaging results and computational prediction of RNA structure we identified a strong secondary structure, increasing the chance pore interaction and clogging. Future work in this aspect could aim to apply methods to reduce pore-RNA interactions. Including surface coating[14], lowering salt concentrations, or either increasing the size of the pore to facilitate the translocation of RNA-carrier bound complex through the pore.

6.2. References

1. R. M. Kuanaeva, A. N. Vaneev, P. V. Gorelkin and A. S. Erofeev, *Biosensors (Basel)*, 2024, **14**.
2. J. Stanley and N. Pourmand, *APL Materials*, 2020, **8**.
3. D. Perry, D. Momotenko, R. A. Lazenby, M. Kang and P. R. Unwin, *Anal. Chem.*, 2016, **88**, 5523-5530.
4. R. L. Fraccari, M. Carminati, G. Piantanida, T. Leontidou, G. Ferrari and T. Albrecht, *Faraday Discuss.*, 2016, **193**, 459-470.
5. R. L. Fraccari, P. Ciccarella, A. Bahrami, M. Carminati, G. Ferrari and T. Albrecht, *Nanoscale*, 2016, **8**, 7604-7611.
6. A. Y. Y. Loh, C. H. Burgess, D. A. Tanase, G. Ferrari, M. A. McLachlan, A. E. G. Cass and T. Albrecht, *Anal. Chem.*, 2018, **90**, 14063-14071.
7. L. M. Ying, A. Bruckbauer, A. M. Rothery, Y. E. Korchev and D. Klenerman, *Anal. Chem.*, 2002, **74**, 1380-1385.
8. A. P. Ivanov, P. Actis, P. Jonsson, D. Klenerman, Y. Korchev and J. B. Edel, *ACS Nano*, 2015, **9**, 3587-3595.
9. F. Zheng, M. Alawami, J. Zhu, C. M. Platnich, J. Sha, U. F. Keyser and K. Chen, *Nano Lett.*, 2023, **23**, 11145-11151.
10. K. Chen, N. A. W. Bell, J. Kong, Y. Tian and U. F. Keyser, *Biophys. J.*, 2017, **112**, 674-682.
11. N. A. W. Bell, K. Chen, S. Ghosal, M. Ricci and U. F. Keyser, *Nat Commun*, 2017, **8**, 380.
12. F. Zheng and Q. Han, *Analyst*, 2024, **149**, 5131-5138.
13. R. A. Al-Waqfi, C. Khan, O. J. Irving, L. Matthews and T. Albrecht, *arXiv preprint arXiv:2501.14347*, 2025.
14. O. M. Eggenberger, C. Ying and M. Mayer, *Nanoscale*, 2019, **11**, 19636-19657.

APPENDIX

CONTENTS

Appendix 1 Overhang Sequences and UCA1 Plasmid Map

Table A1 lists the sequences of both long and short overhangs oligonucleotides used for designing capture probes targeting the UCA-1 lncRNA.. The table includes the sequence, provider and stock concentration of each oligo.

Table A2 sequences and concentration of Oligos used for designing capture probe for UCA-1

Oligo name	Sequence	Provider	Stock concentration*
RE site	5'-TCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGTAATC-3'	Merck	100 µM
Long overhang I	5'—CTCCGGCTTAGGTTGGGTTATATAACTATATGTCACTACGGCTCAGACCGCCGCGGTACGCTGCCAGCGATTACCCGAAGCTCGGCATCTGTGATGGTCCAAGGGGCTTCTGAGGCGATCGGGCAGCATCAGTCTTCAGCCACTAAG CCGAGGAGATC –3'	IDT	100 µM
Long overhang II	5'-CTTGAGCCATTTGGGAATTAGAGCCAGCAAAATCACCATGGCGCCGAAAGCAGCAGCGCTCGGTCGGAGTCTGACTCTTTTAGGAAGATTTCTTTTCTGTCACCTTTTCACATATAAGGAAAGGAAAGGAAAACCACATGAAACGCCTCTGTTTACAA -3'	IDT	100 µM
Short overhang I	5'-TCGCTGGCAGCGTACCGCGGCGGTCTGAGCCGTAGTGGC AAATCCAATCGCAAGACAAAGAACGCGAGAA -3'	IDT	100 µM
Short overhang II	5'-ACTCCGACCGAGCGCTGCTGCTTTCGGCGCCAGTAGCAC CATTACCATTAGCAAGGCCGGAAACGTCACC -3'	IDT	100 µM
190 oligo Pool mix	As adapted by Bell [1,2](38 nucleotides for each oligo)	IDT	30 µM each

**Oligos were dissolved in nuclease free water to get the desired stock solution concentrations*

Figure A1.1 provides an illustration of The UCA1 designed plasmid map. Accordingly, key features are shown including the position of the SfiI restriction enzyme (RE) site and the SP6 promoter sequence. The SfiI site is placed by the manufacturer for restriction digestion and cutting. SP6 promoter site was added at the 5' position to facilitate transcription by SP6 RNA polymerase.

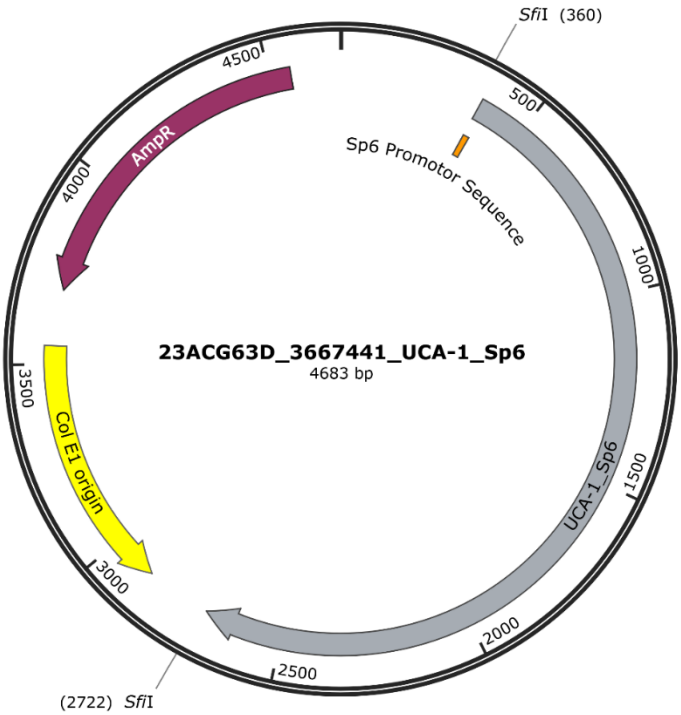

Figure A.*1*: UCA1-plasmid map illustration with position of SfiI RE site and sp6 promoter sequence

Appendix 2 Supporting Data for Chapter 3

This section includes additional supporting data for chapter 3 that was published as supporting information for the main manuscript.[3, 4]

1. Other DNA lengths, ΔI_e vs τ scatter plots

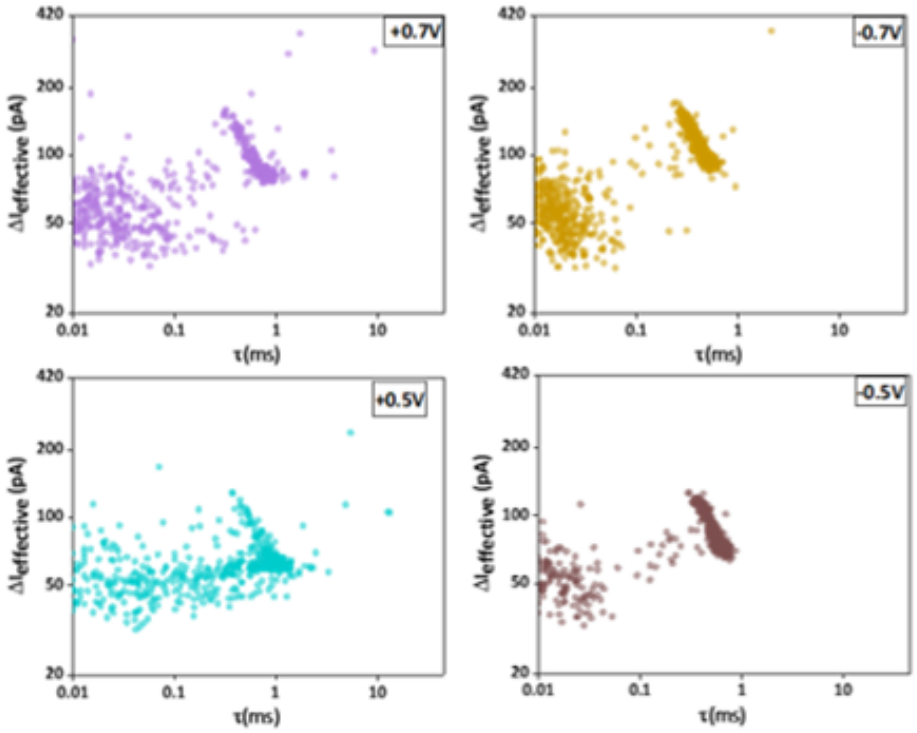

Figure A.2: Effective current scatter plots for 4 kbp DNA ($c_{DNA,out}$ = 300 pM), bias applied in sequence from left to right..

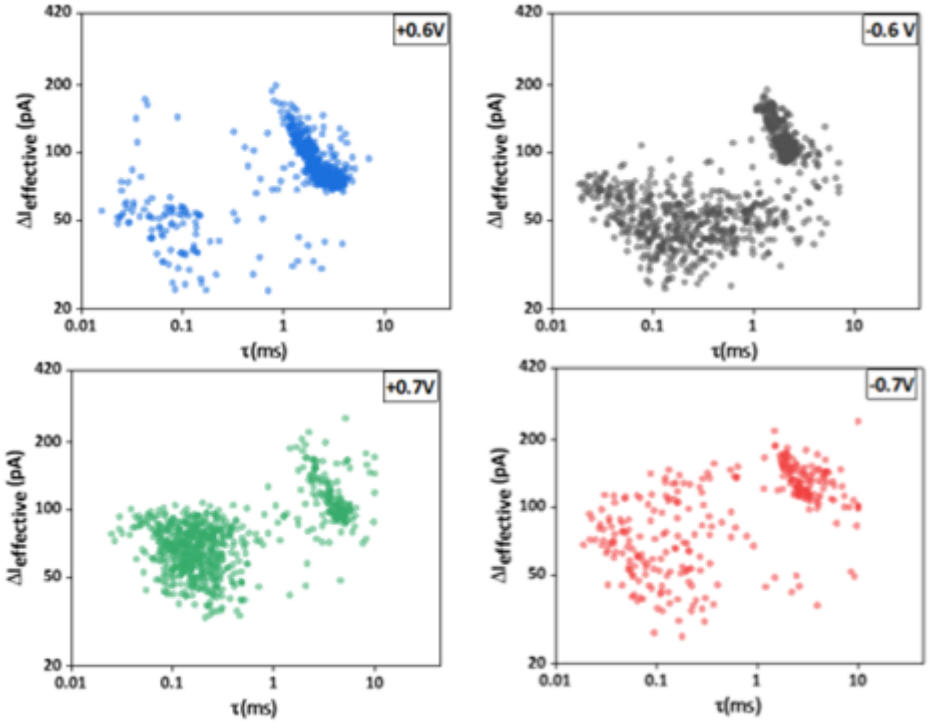

Figure A.3: Effective current scatter plots for 7 kbp ($c_{DNA,out}$ = 600 pM). The applied bias sequence was from left to right, top to bottom.

2. Additional translocation statistics for the DNA samples studied in chapter 3

DNA length (kbp)	V_{bias} (V)	ratio, linear vs. folded	τ_m (linear events) (s)	G_{pore} (nS)	no. of events
4	-0.6	0.596	0.000538	48.4	1028
4	-0.5	0.604	0.000596	48.4	1450
4	-0.7	0.569	0.000480	48.4	1778
4	-0.8	0.562	0.000442	48.4	1077
4	-0.4	0.509	0.000571	48.4	432
4	0.6	0.6348	0.000816	48.4	419
4	0.5	0.549	0.000686	48.4	271
4	0.7	0.6729	0.000837	48.4	691
4	-0.5	0.629	0.000615	26.76	383
4	-0.8	0.55	0.000551	26.76	673
4	-0.6	0.59055	0.000568	26.76	464
4	-0.7	0.59666	0.000486	26.76	534
4	-0.4	0.643	0.000723	26.76	253
4	0.6	0.7703	0.001060	26.76	283
4	0.7	0.834	0.000791	26.76	362
4	0.5	0.8	0.001070	26.76	280
7	-0.5	0.576	0.001600	27.5	520
7	-0.8	0.538	0.001500	27.5	1069
7	-0.6	0.556	0.002300	27.5	254
7	-0.7	0.42	0.004000	27.5	193
7	0.5	0.50993	0.006500	27.5	183
7	0.8	0.6319	0.002000	27.5	398
7	0.6	0.61966	0.003000	27.5	573
7	0.7	0.4918	0.004600	27.5	151
10	-0.8	0.4972	0.004040	36.9	1215
10	-0.6	0.572	0.001640	36.9	131
10	-0.7	0.462	0.002970	36.9	1051
10	0.7	0.552	0.003960	36.9	1035
48.5	-0.5	0.483	0.008590	31.6	172
48.5	0.8	0.5703	0.021520	31.6	154

Appendix References

1. N. A. Bell, K. Chen, S. Ghosal, M. Ricci and U. F. Keyser, *Nat Commun*, 2017, **8**, 380.
2. N. A. Bell and U. F. Keyser, *Nat Nanotechnol*, 2016, **11**, 645-651.
3. R. A. Al-Waqfi, C. Khan, O. J. Irving, L. Matthews and T. Albrecht, *arXiv preprint arXiv:2501.14347*, 2025.
4. R. A. Al-Waqfi, C. J. Khan, O. J. Irving, L. Matthews and T. Albrecht, *ACS Nano*, 2025, **19**, 16803-16812.