UNIVERSITAT DE BARCELONA

Salt dependent DNA translocation dynamics across nanopores

Alejandro Colchero Truniger

PhD. Thesis

Salt dependent DNA translocation dynamics across nanopores

Author: Alejandro Colchero Truniger

Supervisors: Félix Ritort Farran & Isabel Pastor del Campo

UNIVERSITAT DE BARCELONA

Salt dependent DNA translocation dynamics across nanopores

Programa de Doctorado en Física

Autor/a: Alejandro Colchero Truniger

Director/a: Félix Ritort Farran & Isabel Pastor del Campo

Tutor/a: Giancarlo Franzese

Barcelona, Enero 2025

UNIVERSITAT DE BARCELONA

Retwisting the helix

Salt dependent DNA translocation dynamics across nanopores

Alejandro Colchero Truniger

January, 2025

Contents

I INTRODUCTION

0 General Introduction **3**
- 0.1 Biophysics . . . 3
- 0.2 Single-Molecule techniques . . . 4
- 0.3 Summary of the thesis . . . 8

1 Introduction to electrical measurements with nanopipettes **11**
- 1.1 Brief History of nanopore translocation . . . 11
- 1.2 Basics of nanopore measurements . . . 12
  - 1.2.1 Different Nanopore Types . . . 12
  - 1.2.2 The nanopipette translocation setup . . . 13
  - 1.2.3 Electric signal in nanopore-based sensing . . . 16
  - 1.2.4 On nanopipette measurements resolution: Time, amplitude and geometry . . . 17
- 1.3 Noise in nanopore experiments . . . 18
- 1.4 Nanofluidics phenomena relevant for nanopores . . . 21
  - 1.4.1 Electrostatic length scales and their influence . . . 21
  - 1.4.2 Relevant nanofluidics effects . . . 23
- 1.5 Conductance Model For Nanopipettes . . . 25
  - 1.5.1 Geometric model for nanopipettes . . . 25
  - 1.5.2 Modelling Surface Charge . . . 27
  - 1.5.3 Conductivity of Salt Solutions . . . 30
  - 1.5.4 Electric field at the tip of the nanopipette . . . 32
- 1.6 DNA translocation through nanopores . . . 33
  - 1.6.1 DNA current signal . . . 33
  - 1.6.2 Dwell times . . . 34
  - 1.6.3 DNA translocation rate . . . 35

2 The optical tweezers setup **37**
- 2.1 Optical trapping . . . 37

2.2 Mini-tweezers optical setup . . . 39
2.2.1 The Optical path . . . 39
2.2.2 The microfluidic chamber . . . 42
2.2.3 Experimental configuration . . . 43
2.2.4 Calibration . . . 44
2.2.5 Mini-tweezers pulling protocol . . . 46

3 Nucleic Acids **47**
3.1 Nucleic acids . . . 47
3.1.1 DNA structure . . . 49
3.1.2 Elastic models to describe DNA . . . 51
3.2 Helicases . . . 54

II ELECTRICAL MEASUREMENTS WITH NANOPIPETTES

4 Nanopipette conductance and noise **57**
4.1 Nanopipette geometry . . . 57
4.2 Conductance measurements of Nanopipettes . . . 59
4.3 Modelling Nanopipette Conductance . . . 61
4.3.1 Modelling surface charge . . . 62
4.3.2 Nanopipette Conductance at different concentrations . 63
4.3.3 Surface vs. bulk conductance . . . 64
4.3.4 Diameter estimation at different salt concentrations . 66
4.4 Flicker noise in nanopipettes . . . 68
4.5 Conclusions . . . 71

5 Translocation of λ-DNA in high monovalent salt concentrations **73**
5.1 λ-DNA translocation through nanopipettes . . . 74
5.2 λ-DNA analysis workflow . . . 77
5.3 Investigating charge blockade . . . 82
5.4 Cation size effect on current blockades . . . 86
5.5 Phenomenological scaling for dwell times . . . 89
5.6 Cation size effect on charge blockade . . . 94
5.7 Conclusions . . . 94

6 Study of λ-DNA configurations during translocation **97**
6.1 Introduction . . . 97
6.1.1 Dynamics of DNA translocation . . . 99
6.1.2 Chapter summary . . . 101
6.2 Configuration analysis . . . 102

6.2.1 Levels detection and level assignation 102
6.2.2 Code degeneracy 104
6.2.3 Knots detection 106
6.3 Effect of voltage on levels residence times and level transitions 106
6.3.1 Transition-level matrix 108
6.3.2 End-to-end time analysis 110
6.4 Effect of salt concentration on confined configurations 112
6.4.1 Conclusions of voltage and concentration effects 114
6.5 Dwell-time dispersion of different configurations 115
6.5.1 Modeling DNA translocation through nanopores . . . 115
6.5.2 Configuration dependent effective diffusion 119
6.5.3 Origin of configuration dependent velocity fluctuations 123
6.5.4 Simulating dwell-times distributions using FP 126
6.6 Conclusions 129

III DNA UNZIPPING WITH OPTICAL TWEEZERS

7 Unzipping experiments of a long DNA hairpin in monovalent salts 133
7.1 Experimental Protocols 134
7.2 Unzipping experiments in monovalent salts 137
7.2.1 Unzipping over a broad range of concentrations 138
7.3 Discussion with previous nanopore results 144
7.4 Conclusions 146

IV COLLABORATIONS

8 Single-molecule picometer resolution nanopore tweezers 151
8.1 Introduction to SPRNT 151
8.1.1 The MspA biological nanopore 154
8.2 Experiments with Hel308 155
8.3 Measuring gp41 with SPRNT 157
8.3.1 Conclusions 160

V FINAL CONCLUSIONS

9 Final Conclusions and future perspectives 165

VI APPENDIX

A A Single-Molecule Translocation Beginners Guide 171
A.1 Design and Fabrication of microfluidic chips for single-molecule translocation . . . 171
A.1.1 Material Selection for the microfluidic chip . . . 171
A.1.2 PDMS mold design and fabrication . . . 172
A.2 Fabrication of nanopipettes using a laser-based pipette puller 174
A.2.1 Controlling environmental parameters when fabricating nanopipettes . . . 176
A.2.2 Pipette puller maintenance . . . 176
A.3 Characterizing the nanopipettes with electron microscope . . 177
A.4 Assembling the microfluidic chamber . . . 178
A.5 Filling and using the chip for an experiment . . . 179

B Synthesis of DNA hairpin for Optical Tweezers experiments 181
B.1 Required DNA fragments and oligos for the synthesis . . . 181
B.2 Synthesis steps . . . 182
B.2.1 Digestion of λ-DNA with EcoRI . . . 182
B.2.2 Phosphorylation of λ-DNA/EcoRI to put 5'-Pho at cosL end . . . 183
B.2.3 Purification of λ-DNA/EcoRI with QIA quick purification Kit . . . 185
B.2.4 Oligonucleotides labeling . . . 185
B.2.5 Annealing reaction . . . 186
B.2.6 Ligation reaction . . . 187

C Appendix Chapter 5 189

D MATLAB codes developed for DNA translocation analysis 193
D.1 MATLAB app for the pretreatment of DNA translocation data 193
D.2 MATLAB app for the analysis of DNA translocation data . . 194

E Appendix Chapter 4 197
E.1 Simulations of λ-DNA using the Worm like chain model . . . 197
E.2 Effect of voltage on levels residence times and level transitions 199
E.3 Summary of analyzed data . . . 200

F A SPRNT (Single-molecule picometer resolution nanopore tweezers) beginners Guide 205
F.1 Performing SPRNT experiments . . . 205
F.1.1 Formation of the lipid membrane . . . 205
F.1.2 Flowing the MspA nanopore to get an insertion . . . 207
F.1.3 Running the experiment . . . 208

Resum de la tesi en català

Dinàmica de translocació de l'ADN dependent de la sal a través de nanopores

Aquesta tesi utilitza dues tècniques complementàries de molècula individual, la microscòpia de nanopipeta i les pinces òptiques, per investigar l'impacte de diverses sals monovalents a concentracions elevades en l'ADN. La tesi comença caracteritzant les propietats de la conductància i el soroll de les nanopipetes en un ampli rang de concentracions. Un cop caracteritzades les nanopipetes, s'utilitzen per dur a terme experiments de translocació de l'ADN per examinar com diferents cations influeixen en els seus paràmetres de translocació. Aquests experiments també ens permeten explorar els efectes de la del voltatge aplicat i la concentració de sal en les configuracions de plegament de l'ADN durant la translocació. A més, explorem per què les configuracions compactes de plegament de l'ADN presenten temps d'estada més baixos en comparació amb les configuracions de plegament més allargades. Finalment, s'utilitzen pinces òptiques per dur a terme experiments de d'estirament en una forquilla d'ADN, proporcionant informació sobre l'estabilitat de l'ADN en condicions d'alta força iònica relacionades amb els experiments de translocació de l'ADN.

La tesi es divideix en sis parts. La Part I proporciona una introducció a les tècniques experimentals utilitzades al llarg d'aquest treball, juntament amb els marcs teòrics i conceptes que seran necessaris per a les parts II, III i IV. Aquesta part es divideix en tres capítols. El capítol 1 és una introducció al camp de nanopor. El capítol descriu la configuració i els conceptes bàsics necessaris per realitzar mesures elèctriques amb nanopipetes i discuteix els límits de resolució de la tècnica. També es descriuen les principals fonts de soroll quan es duen a terme experiments amb nanopors. També, s'introdueixen alguns fenòmens nanofluídics rellevants quan es treballa a nanoescala, juntament amb alguns fonaments teòrics sobre la regulació de la càrrega superficial. Finalment, es mostren alguns resultats previs de translocació d'ADN a través de nanopors. El capítol 2 introdueix la trampa òptica i la configuració del *mini-tweezers* que es va utilitzar per realitzar els experiments en aquesta tesi. El capítol 3 introdueix els components

bàsics i l'estructura dels àcids nucleics, centrant-se en l'ADN, que serà la biomolècula estudiada al llarg de la tesi. El capítol conclou presentant els fonaments teòrics de dos models elàstics emprats per descriure les propietats elàstiques dels polinucleòtids.

La Part II conté resultats d'experiments amb nanopipetes. Comença amb el capítol 4, on s'estudia la conductància de les nanopipetes per a diferents concentracions de sal, comparant les contribucions de la conductància en volum i superfície i l'efecte del pH sobre la càrrega superficial. A més, es comparen dos models de conductivitat per per modelitzar la conductància de les nanopipetes amb la concentració. A més, s'analitza el soroll de parpelleig les nanopipetes, explorant com canvien els paràmetres que descriuen el soroll amb concentració i tensió. El capítol 5 presenta experiments de translocació λ-ADN en diferents sals monovalents. Aquest capítol investiga l'efecte de la concentració i el tipus de catió en els paràmetres de translocació com el temps d'estada, el bloqueig de la corrent elèctrica i el bloqueig de càrrega de les translocacions λ-DNA. El capítol se centra en com la interacció catió-ADN canvia amb la mida del catió. El capítol 6 estudia les diferents configuracions de plegament que es produeixen durant la translocació de l'ADN i com depenen de la tensió i la concentració de sal. Per a això, es fa una anàlisi dels diferents nivells que es produeixen durant la translocació λ-DNA i el temps de residència dels diferents nivells. L'anàlisi també ens permet extreure dades generals sobre la translocació de l'ADN a través de nanopipetes. Finalment, s'exploren les causes de la menor dispersió del temps d'estada de les configuracions de plegaments més compacte respecte a les configuracions de plegament més allargades.

La part III inclou experiments amb pinces òptiques. En el capítol 7, les pinces òptiques s'utilitzen per dur a terme experiments d'estirament de llargues forquilles d'ADN a altes concentracions iòniques de diverses sals monovalents. S'investiga la força mitjana d'obertura i l'efecte de la concentració en l'estabilitat de la forquilla. El capítol conclou amb una discussió conjunta dels resultats dels experiments de translocació i pinces òptiques en altes concentracions de sal.

La part IV conté resultats d'una estada internacional d'un mes. El capítol 8 inclou una breu introducció a la tècnica SPRNT (Single-molecule picometer resolution nanopore tweezers), que es va utilitzar per realitzar experiments amb l'helicasa Hel308. A més, es presenten alguns experiments preliminars de la amb l'helicasa gp41 helicasa fent servir SPRNT.

La part V conté les conclusions finals de la tesi i les perspectives futures del treball.

La Part VI consta de tots els Apèndixs. Els Appendices E i C complementen els resultats d'alguns dels capítols, mentre que els Appendices A, B, F i D descriuen en detall els protocols experimentals més importants i

els codis MATLAB utilitzats en el desenvolupament de la tesi.

Paraules clau: *Nanopore, translocació, ADN, sals monovalents, pinces òptiques.*

Part I

INTRODUCTION

Chapter 0

General Introduction

0.1 Biophysics

Biophysics is a bridge science between Biology and Physics. This discipline uses the tools and methodology developed in physics to study biological systems. Biophysicists make an effort to condense the complex and qualitative descriptions of biology into quantitative and more simplistic descriptions. By doing this, biophysicists uncover the underlying laws describing biological systems. One example is the use of quantum mechanics to understand the fundamental process of photosynthesis [1].

Nevertheless, very often, biophysicists contribute not only by describing the underlying laws but also by developing new experimental tools that are more sensitive than the previously available ones. The application of the new experimental tools to biological systems uncovers new mechanisms or processes in biological systems that were hidden from biologists due to a lack of precision of the previous techniques. An example of this was the X-ray diffraction image obtained by Rosalind Franklin (Fig. 1a), which led to the discovery of the double helix structure of DNA in 1953 by Watson and Francis Crick [2]. This groundbreaking image was achieved thanks to applying an experimental technique that physicists were using to study crystals, to study DNA, a biological molecule.

The interest of physicists in biology, particularly in the fundamental rules governing life, started to grow during the 20th century. This interest was probably inspired by the book *What is Life?* [3] by Erwin Schrödinger in 1944 (Fig1b). In the book, Schrödinger made the hypothesis that genetic information would be contained in the form of an aperiodic solid with a configuration of covalent chemical bonds. This book, together with the subsequent resolution of the structure of DNA, started a revolution in molecular biophysics that has not come to an end yet. Indeed, some people claim that if the 20th century was the century of physics, the 21st century will

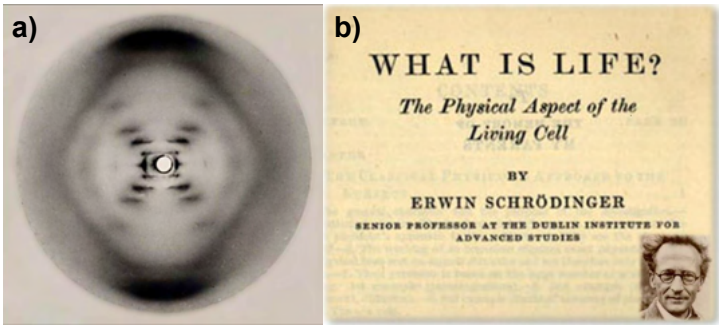


Figure 1: a) Photo 51, the First X-ray image showing the X-ray diffraction pattern of DNA. b) front page of Schrödinger's book *What is Life?*, published in 1944.

be the century of biology [4]. The recent advances in genetics, molecular biology, and synthetic biology, together with the advances in technology and the fusion with other disciplines like physics, computer sciences, and mathematics, are laying the grounds for a new revolution in biology.

There is still a lot of exciting physics to be uncovered by studying biological systems. There are also many biological questions to be answered thanks to the new experimental tools provided by biophysicists. All this knowledge will continue with the ongoing revolution in biology. In view of a quote by Richard P. Feynman, "What I cannot build. I do not understand". Nowadays, we can describe and understand countless biological processes that occur in organisms. In many cases, we have even adapted the tools given to us by biology and have used them, such as in the case of CRISPR cas9 for genome editing or nanopore sequencing, where we profit from biological nanopores and motor proteins of biological organisms to develop DNA sequencing technologies. However, we are still very far from being able to design and build biological nanopores or motor proteins from scratch.

0.2 Single-Molecule techniques

Single-molecule techniques have revolutionized the study of biomolecules by providing detailed insights into their behavior and interactions at a more fundamental level. Unlike bulk measurements, these methods allow researchers to observe the heterogeneity, dynamics, and rare events that are often obscured in bulk studies. By directly probing individual molecules, they uncover critical mechanisms underlying processes like protein folding, enzyme activity, DNA-protein interactions, and molecular motor function, offering a deeper understanding of the molecular basis of life. These advances are made possible through various techniques, such as fluorescence-

based methods, force spectroscopy techniques like optical tweezers, magnetic tweezers, atomic force microscopy, or, more recently, nanopore-based techniques. Each of them contributes with their unique capabilities for exploring biomolecular phenomena.

During the 20th century, most discoveries in molecular biology involved bulk experiments, where a large number of molecules were measured simultaneously. These experiments are very useful for studying the average behavior of molecules. However, they do not grasp the finer details of how these molecules behave and the fluctuations that occur at the nanoscale. The development of single-molecule techniques has allowed the study of biological molecules 'one at a time' with more precision. One example where single-molecule techniques have uncovered the details hidden in macroscopic/bulk experiments is the case of molecular motors moving along DNA. Even though from bulk assays, one may think the movement of helicases is smooth and continuous as they move over DNA. Single-molecule techniques, such as optical and magnetic tweezers, or more novel ones like SPRNT, presented in Chapter 8, have shown that helicase movement is far from continuous and that their movement consists of a series of quick steps alternated with pauses and even backsteps.

Below is a brief summary of the most relevant single-molecule techniques for studying biomolecules:

1. **Fluorescence Resonance Energy Transfer (FRET).** FRET is a highly sensitive experimental technique used to study single molecules by detecting energy transfer between two fluorescent dyes, a donor and an acceptor. This energy transfer occurs when the two fluorophores are in close proximity, typically within 1–10 nm, and depends on their distance and relative orientation, as shown in Fig. 2a. By monitoring changes in FRET efficiency, researchers can investigate molecular interactions, structural changes, and dynamic processes at the nanoscale. FRET is widely applied in biophysics and molecular biology to unravel conformational dynamics, protein folding, DNA-protein interactions, and complex assembly mechanisms in real-time [5].

2. **Magnetic tweezers (MT).** MT is an experimental technique used to study single molecules by applying controlled forces and torques using magnetic fields. A magnetic bead attached to a molecule of interest, as outlined in Fig. 2b, is manipulated to stretch, twist, or unfold the molecule, enabling precise measurements of mechanical properties, such as elasticity and torsional stiffness. This technique is widely used to explore DNA mechanics, protein folding, and molecular interactions under physiologically relevant conditions [6].

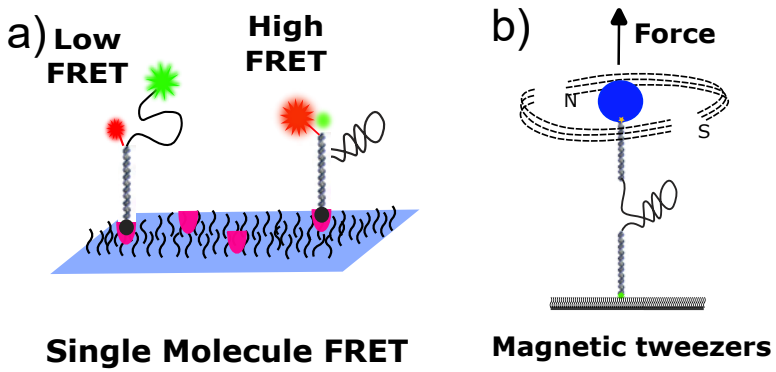


Figure 2: a) FRET principles. b) Magnetic tweezers principles.

3. **Atomic Force Microscopy (AFM).** AFM is a versatile experimental technique used to study single molecules by measuring forces or imaging surfaces with absorbed molecules on them, with nanometer resolution. A sharp probe on a flexible cantilever interacts with the sample, as sketched in Fig.3a, enabling the study of molecular structures, chemical properties, and dynamic processes. Moreover, the AFM can manipulate and apply forces to unfold biomolecules, providing detailed insights into protein mechanics and their elastic properties.

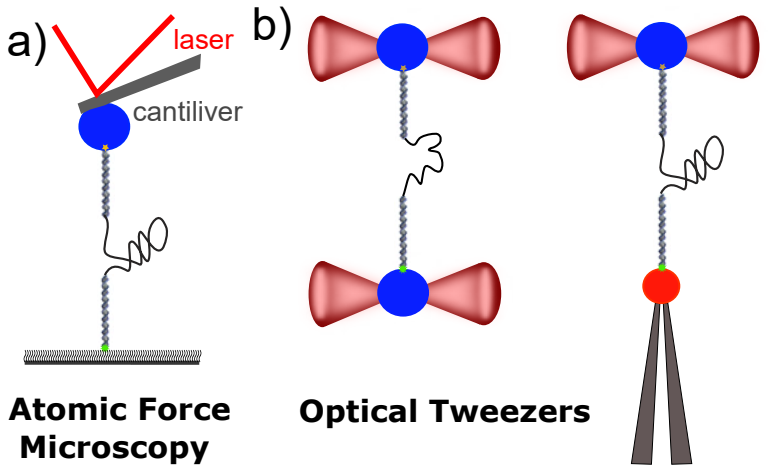


Figure 3: a) Sketch of AFM for studying single molecules. b) Optical tweezers.

4. **Optical tweezers (OT).** OT are a precise experimental technique that uses highly focused laser beams to trap and manipulate microscopic particles, such as beads attached to single molecules, as shown in Fig.3b. By applying controlled forces, OT enables the study of

molecular mechanics, conformational changes, and interactions between molecules in real time. Optical tweezers are widely used to investigate biomolecular processes, including DNA-protein interactions, molecular motor activity, and protein folding. Moreover, OT are also used to study nucleic acids and proteins at a more fundamental level by investigating their elastic properties, the base-pairing energy, or its structure under different conditions.

5. **Nanopores.** Nanopore is an experimental technique used to study single molecules by monitoring the ionic current as molecules pass through a nanometer-scale pore, as sketched in Fig.4a. Changes in the current provide detailed information about the molecule's size, shape, and charge. Nanopores are also widely used for sequencing DNA due to the astonishing sensitivities they can achieve, which allows distinguishing between the different DNA nucleotides. Moreover, nanopores can also be used to study nucleic acids or protein properties or interactions with other molecules. The extreme sensitivity of biological nanopores has provided a great platform to develop nanopore base techniques that combine biological nanopores with a motor enzyme that controls the translocation of nucleic acids, as shown in Fig.4b. Two recently developed nanopore-based techniques are DNA sequencing and Single-molecule picometer resolution nanopore tweezers (SPRNT), which are used to sequence DNA and study enzyme motion over DNA.

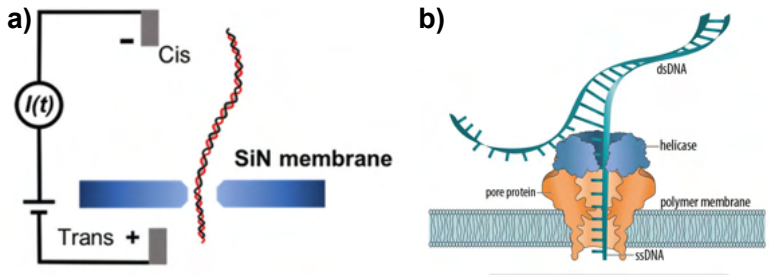


Figure 4: a) Solid state nanopore for studying DNA. b) Biological nanopore combined with a molecular motor, which can be used for DNA sequencing or SPRNT experiments.

Interestingly, except for nanopores, all the other techniques require photons in one way or another to function. In FRET, the light intensity correlates with the distance between the fluorophores. In OT and AFM, photons do not interact with the sample; these techniques use laser deflection to estimate the force. On the contrary, MT and some OT setups detect the

position of beads by optical imaging and then use these images to estimate the force. Nanopore-based techniques measure ion flow and do not require photons to measure the signal. This provides some advantages, such as the fact that light diffraction, which usually limits optical techniques, does not limit nanopores, and differences in molecule size of ~ 1 nm can be measured by the ion signal in nanopores. However, there are also some disadvantages, as the ions required to measure the signal affect the molecule environment and modify the nanopore's conductance behavior, limiting the conditions in which the molecules can be reliably measured.

Another important fact is that while nanopores are not a force spectroscopy technique, when molecules translocate through them, the electric field generated at the nanopore exerts a force on the molecules comparable to the force applied by other force spectroscopy techniques, such as OT, MT, or AFM. The force exerted on the molecules by the electric field can not be measured unless combined with the other force spectroscopy techniques. Hence, nanopores are not considered a force spectroscopy technique. However, the response of molecules to the applied force can be measured by detecting their folding configuration during translocation, which can provide some information about the molecule's elastic properties, which are usually studied with force spectroscopy techniques.

In this thesis, three different single-molecule techniques have been used. The primary technique used is nanopipettes, which were used to study the translocation mechanisms of DNA in high monovalent salts. Optical tweezers are used to perform force spectroscopy experiments of a DNA hairpin in high monovalent salt environments. Finally, the novel nanopore-base SPRNT technique is used to study the motion of helicases over DNA.

0.3 Summary of the thesis

This thesis employs two complementary single-molecule techniques, nanopipette microscopy, and optical tweezers, to investigate the impact of high concentrations of various monovalent salts on DNA. The thesis begins by characterizing the conductance and noise properties of nanopipettes across a wide range of concentrations. Once nanopipettes have been characterized, they are used to conduct DNA translocation experiments to examine how different cations influence DNA translocation parameters. These experiments also allow us to explore the effects of applied voltage and salt concentration on the folding configurations of DNA during translocation. Moreover, we explore why compact DNA folding configurations exhibit lower dwell times compared to more extended folding configurations. Finally, optical tweezers are used to perform unzipping experiments on a DNA hairpin, providing insights into the stability of DNA under high ionic strength

conditions relating to the DNA translocation experiments.

The thesis is divided into six parts. Part I provides an introduction to the experimental techniques used throughout this work, together with the theoretical frameworks and concepts that will be required for parts II, III, and IV. This part is divided into three chapters. Chapter 1 is an introduction to the nanopore field. The chapter describes the setup and the basic concepts required to perform electrical measurements with nanopipettes and discusses the resolution limits of the technique. The main sources of noise when performing nanopore experiments are also described. Some important nanofluidics phenomena relevant when working at the nanoscale are also introduced, along with some theoretical basics about surface charge regulation. Finally, some previous results of DNA translocation through nanopores are presented. Chapter 2 introduces optical trapping and the setup the mini-tweezer that was used to perform the experiments in this thesis. A brief description of how to perform experiments in the mini-tweezers and the calibration procedures are provided. Chapter 3 introduces nucleic acids' basic components and structure, focusing on DNA, which will be the biomolecule studied throughout the thesis. The chapter concludes by presenting the theoretical basics of two elastic models used to describe polynucleotide elastic properties.

Part II contains results from experiments with nanopipettes. It starts with Chapter 4, where the conductance of nanopipettes is studied over a broad range of salt concentrations, comparing the contributions of bulk and surface conductance and the effect of pH on surface charge. Moreover, two conductivity models for the conductivity of salt solutions with concentration are compared to determine which one better reproduces the nanopipette conductance. Additionally, the flicker noise of nanopipettes is analyzed, exploring how the parameters describing the noise change with concentration and voltage. Chapter 5 presents experiments of λ-DNA translocation in different monovalent salts. This chapter investigates the effect of concentration and cation type on translocation parameters such as the dwell time, the current blockade, and the charge blockade of λ-DNA translocations. The chapter focuses on how the cation-DNA interaction changes with cation size. Chapter 6 studies the different folding configurations that occur during DNA translocation and how they depend on voltage and salt concentration. For this, an analysis of the different levels occurring during λ-DNA translocation and the residence time of the different levels is done. The analysis also allows us to extract general facts about DNA translocation through nanopipettes. Finally, the causes of the smaller dwell time dispersion of more compact folding configurations with respect to more extended folding configurations are explored.

Part III includes experiments with optical tweezers. In chapter 7, optical

tweezers are used to perform unzipping experiments of long DNA hairpins in high ionic concentrations of various monovalent salts. We investigate the effect of high ionic strength on the mean unzipping force and the hairpin stability. The chapter concludes with a joint discussion of the results of translocation and optical tweezers experiments in high salt concentrations.

Part IV contains results from a 1-month research stay at the University of Seattle. Chapter 8 includes a short introduction to the SPRNT (Single-molecule picometer resolution nanopore tweezers) technique, which was used to perform experiments with Hel308. Moreover, some preliminary experiments of the gp41 helicase with SPRNT are presented.

Part V contains the final conclusions of the thesis and future perspectives of the work.

Part VI consists of all the Appendices. Appendices E and C complement the results of some of the chapters, while Appendices A, B, F and D describe in detail the more important experimental protocols and the MATLAB codes used in the development of the thesis.

Keywords: *Nanopore, translocation, DNA, monovalent salts, optical tweezers, DNA hairpin.*

Chapter 1

Introduction to electrical measurements with nanopipettes

Nanopipettes have proven to be an excellent tool to detect single molecules [7, 8] and to measure small currents, such as, the tiny currents in electrophysiology studies [9, 10]. Nanopipettes are characterized by nanoscale-size pores at their tips that confer high sensibility for detecting small currents. This chapter aims to introduce the basics of how to perform electrical measurements using nanopipettes. It presents the setup required for performing electrical measurements with nanopipettes. Then, it introduces electrostatic lengths and nanofluidics phenomena, which are relevant when performing electrical measurements at the nanoscale. These concepts are important for interpreting the electrical measurements acquired by nanopipettes. Moreover, a model for the conductance of nanopipettes is presented. This model relates different experimental parameters that contribute to the conductance of nanopipettes. Having presented the experimental setup and the basic concepts, we briefly present the different noise sources that affect electrical measurements with nanopipette. Finally, the use of a nanopipette to measure and study DNA is introduced.

1.1 Brief History of nanopore translocation

The first experiments with nanopores took place at a much larger scale (let us say that they were performed with 'macro-pores'). In the late 1940s, Wallace H.Counter used holes poked into a cellophane cigarette wrapper with a needle to count and size cells [11]. In this method, later patented in 1953, a pair of electrons were placed at the two sides of the orifice filled

with an electrolyte solution [12]. The electrostatic current was measured as a function of time. Direct resistive spikes were measured as a suspension of cells or particles was pushed through the hole. Yet, Coulter's patent is one of the few scientific patents that have revolutionized clinical practice to this day.

Later, in the 1970s, DeBlois and Bean were able to fabricate sub-micrometer track-etched pores and used them to detect nanoscale particles and viruses [13]. During the 1990s, due to developments in hole-making techniques and in sensing of low currents, a new revolution in hole-based detection happened. During the time gap of four decades, the hole size had changed from a few micrometers (10^{-5} m) to a few nanometers (10^{-9} m), a four-order-of-magnitude improvement. Similarly, the minimum analyte size that could be detected went from cells to single biomolecules (DNA, RNA, or proteins). As we moved into the nanoscale realm, the volume-to-surface ratio of the pores decreased, and surface-related effects became more prominent. So, for nanopores in the nanoscale realm, it was essential to understand the chemistry of the nanopore surface [14, 15] and how it could influence properties such as their conductivity.

Electrophysiologists made an essential push toward the birth of the nanopore field. Since the early 1970s, electrophysiologists have been able to monitor protein ion channels in synthetic planar lipid bilayer geometries and developed the tools for it. Indeed, even nowadays, one of the most common instruments to measure the small currents through nanopores, the Axopatch 200-B, is an instrument designed and extensively used for patch clamp electrophysiology experiments. Although nanopores is a much younger branch of biophysics than electrophysiology, it has grown much since its inception in the 1990s due to the investments in potential technology for studying different properties of biomolecules (charge [16], conformation [17]...), for sequencing DNA [18, 19] and studying motor proteins [20].

1.2 Basics of nanopore measurements

This section introduces the different types of nanopores that can be used for single-molecule sensing and presents the experimental setup required to perform nanopipette measurements. We then discuss the origins of nanopores' high sensitivity, which allows nanopores to detect single molecules [21] and even sequence DNA [18]. Finally, their resolution merits are discussed.

1.2.1 Different Nanopore Types

Three main types of nanopores can be used to perform experiments with biomolecules: biological nanopores, solid-state nanopores, and nanopipettes.

Although these nanopore techniques may differ in the pore preparation and the microfluidics needed to perform experiments, the instruments/setup required to measure and acquire the currents through them are the same. Therefore, although the focus will be placed on nanopipettes as they are the technique used in this work, the aspects discussed in this Section apply to all of them.

- Biological nanopores are transmembrane proteins found in living organisms, where they perform various functions. Isolating them onto a lipid membrane (lipid bilayer) can be used to study different molecules (Fig. 1.1a) . Biological nanopores normally have very small diameters that significantly increase their resolving power and are very reproducible because the proteins that conform to them always fold similarly. On the other hand, they can not be used in a wide range of experimental conditions, as lipid bilayers are delicate, and proteins may misfold in more extreme experimental conditions. This restricts their experimental conditions (pH, temperature, voltage...).

- Solid-state nanopores are much more versatile, as they can be used in a broad range of conditions. They can also be fabricated in larger sizes, from just a few nanometers to hundreds of nanometers. Solid-state nanopores can be fabricated by very different methods, such as using focus ion or electron beams for drilling a hole into thin membranes, chemically etching, or dielectric breakdown [22]. However, all the methods require thin membranes like silicon or graphene (Fig. 1.1b). These fabrication techniques are usually more precise than the method used to fabricate nanopipettes. However, they are more time-consuming and require more expensive equipment.

- Nanopipettes (Fig. 1.1c) can also be used in a wide range of conditions. Nanopipettes are a cheaper and faster alternative to fabricating nm-sized holes. They are generated using laser-based pipette pullers, where glass capillaries are pulled to pipettes with nm-sized tips. Although sometimes overlooked, nanopipettes do not require a bulky membrane. Therefore, they are ideal for non-invasive experiments with cells [7, 8] or combined with other techniques, such as optical tweezers [23, 24] or optical techniques like Raman spectroscopy [25]. Due to these last-mentioned features, nanopipettes were used to perform all the experiments in this thesis.

1.2.2 The nanopipette translocation setup

The setup required to make electrical measurements with nanopipettes is shown in Fig. 1.2. The setup consists of a costume-made microfluidic

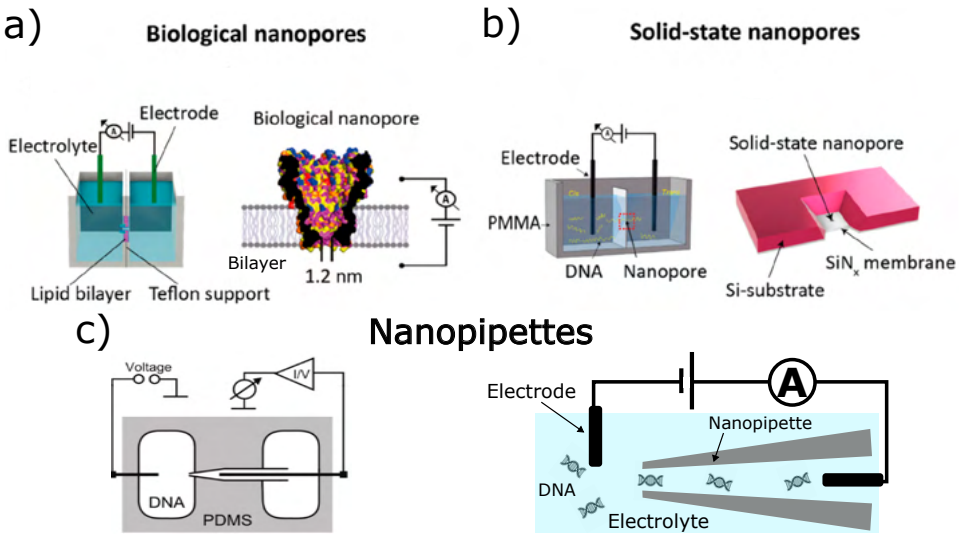


Figure 1.1: **Nanopore types** a) Biological nanopores. The two reservoirs are connected by an aperture. A lipid bilayer is formed over the aperture, and then a biological nanopore is inserted into the bilayer. b) Solid-state nanopores. The two reservoirs are connected by a small nanopore drilled into a thin membrane. c) Nanopipettes. The two reservoirs are connected by a nanopipette with a few nm aperture at the tip.

chamber that contains six nanopipettes (see Fig. 1.3a). The microfluidic chamber is made by polymerizing PDMS (Polydimethylsiloxane) in a mold shown in Fig. 1.2b. Then, nanopipettes are pulled using a P-2000 laser-based pipette puller, Fig. 1.2c, cut, and mounted into the PDMS chip from the central to a lateral chamber. More details on the microfluidic chamber preparation, mold design, and the protocols used in the pipette puller can be found in Appendix A.

As seen in Fig. 1.2, each nanopipette connects the central reservoirs, where the tip of the nanopipette is placed, with a lateral reservoir. The Ag/AgCl electrodes are placed in the central and lateral chamber at the ends of one of the nanopipettes, as shown in red in Fig. 1.2. The reservoirs are filled with the desired buffer for the experiment. The electrodes are used to apply a voltage difference across the nanopipette and then measure the ion current through the nanopipette. The electrode's working principle is as follows. An oxidative electrochemical reaction $\mathrm{Ag} + \mathrm{Cl}^- \rightarrow \mathrm{AgCl} + \mathrm{e}^-$ occurs at the anode, placed in the trans chamber ($-$) (Fig. 1.2 bottom left scheme). This results in the capture of a chloride ion (Cl^- from solution by the electrode, which causes an electron (e^-) to migrate from the electrode into the circuit, producing a current and generating a charge imbalance at the electrode, that results in a cation migration towards the nanopipette,

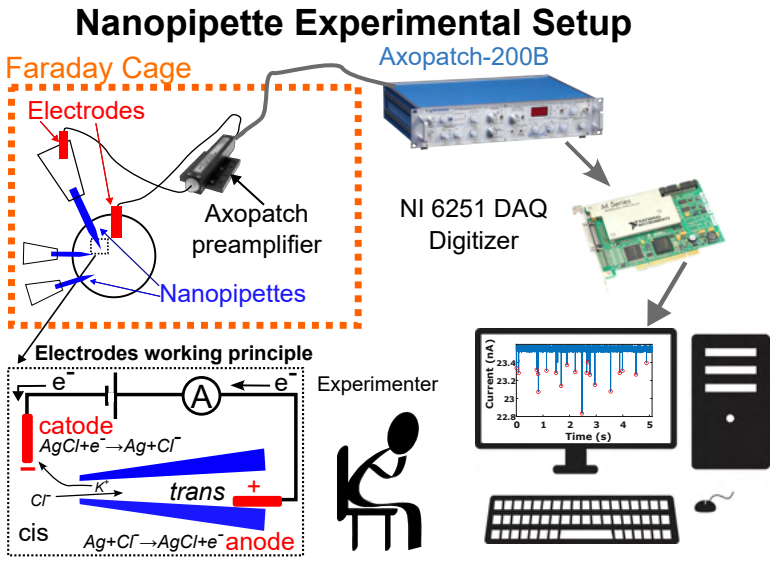


Figure 1.2: **Nanopipette Setup.** Schematic representation of the setup required to perform electrical measurements with nanopipettes. In the left bottom, the inset shows the electrochemistry of Ag/AgCl electrodes.

represented by a K^+ ion in Fig. 1.2, in the bottom left scheme. The reverse reaction occurs at the cathode (+), $AgCl + e^- \rightarrow Ag + Cl^-$, where an electron from the circuit releases a chloride ion from the electrode, that migrates towards the nanopipette. Ag/AgCl electrode can effectively be used to measure electric currents in solutions containing Cl^- ions, however other solutions not containing Cl^- require alternative electrodes.

The small current measured by the electrodes, which is in the nanoampere (nA), is then amplified by the Axopatch pre-amplifier. As the measured electric currents are very small, they need to be measured and amplified inside a Faraday cage to avoid any noise coming from external electromagnetic radiation, as this noise would also be amplified. The pre-amplifier amplifies the signal and converts the nA current signal into a voltage signal. This is done because analog to digital converters always work by digitalizing a voltage signal. The voltage signal is then sent via a coaxial cable to the Axopatch-200B (as shown in Fig. 1.2). The Axopatch-200B can then further amplify and low pass filtered by the signal. Finally, the voltage signal is sent from the Axopatch-200B to an NI PCI-6251 acquisition card, which converts the analog signal from the Axopatch into a digital, discrete signal (Fig. 1.2). The PCI-6251 digitalizes the signal with a 16-bit resolution and an acquisition speed of up to 1.25 MHz. For our experiments, we usually acquire at 250 kHz. A LabVIEW program is used as an interface to save the data on the computer.

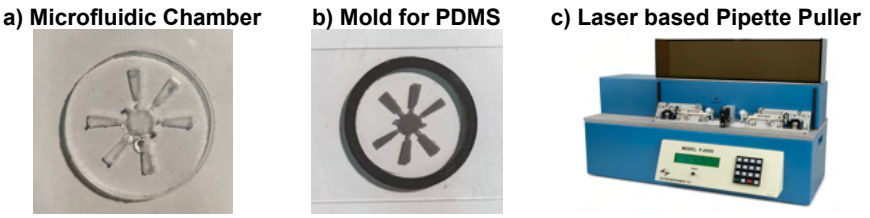


Figure 1.3: **Microfluidic chamber preparation.**a) Image of the microfluidic chambers coating six nanopipettes. b) Mold where PDMS is polymerized to make the microfluidic chamber. c) Laser-based pipette puller is used to fabricate the nanopipettes.

In a typical translocation experiment with biomolecules, the desired molecule is flowed into the central chamber. The electrodes are placed one in the central chamber and one at the end of the nanopipette through which molecules translocate. Applying a voltage bias will make electrically charged molecules translocate through the nanopipette to the lateral chamber. For the case of negatively charged molecules, placing the positive electrode in the lateral chamber will make the molecules translocate from the central to the lateral chamber (Fig. 1.2 bottom left inset). The chamber to which the molecule translocates is called the trans chamber, while the chamber from which they translocate is the cis chamber. When the molecules pass through the narrow section at the tip of the pipette, they will produce a resistive or conductive current spike, depending on the molecule and the experimental conditions of the experiment. These spikes are monitored in real time on the computer screen while they are recorded on the computer using a LabVIEW program (fig. 1.2.2).

1.2.3 Electric signal in nanopore-based sensing

The ions used for pore-based sensing are faster, smaller, and more concentrated than the molecule that wants to be detected. This yields three consequences: (1) For each analyte molecule that crosses the pore, many ions go through the pore; (2) The pore must permit the flow of ions and be wide enough to fit the analyte; (3) As the hydrodynamic radius of the ions is about 0.1 nm, to a first approximation, the flow of ions should be able to report size differences of the order of this hydrodynamic radius.

To give some rough numbers, for a typical nanopipette with a current of ~ 10 nA, approximately $6 \cdot 10^7$ ions pass through the nanopipette during the $\sim$ 1 ms translocation time of a 48502 bp DNA molecule [26]. This is not unique to nanopore-based sensing: in single-molecule microscopy, a single photon is a poor indicator of a process, and often, many photons are required to ascertain the identity and position of a fluorescent molecule in

a sample. Moreover, in analogy with light-based microscopy where lower wavelength light increases resolution (Rayleigh criterion [27]), smaller ions better resolve our sample when doing electrical measurements.

1.2.4 On nanopipette measurements resolution: Time, amplitude and geometry

Nanopore experiments yield three different types of merits regarding their resolution. The first is their temporal resolution. The fastest time resolution of any measurement is usually limited by the sampling rate and limited to the maximum bandwidth of the measurement. For example, nanopore current signals are typically collected using a low-pass filter of $10-100$ kHz. In this case, the time resolution of the experiment is $\delta t = 10 - 100\ \mu s$ (Fig. 1.4a, δt). The low-pass filtered data are then typically acquired at a higher frequency, about 250 kHz, to oversample the fastest frequencies in the signal so that they can be better resolved.

The second type of resolution is the electric current resolution, which is the smallest deviation of the electric current we can detect. A typical 16-bit digitizer can digitalize voltage signals from $\pm 10\ V$, with a voltage noise of 280 μV_{rms}. If data are collected with a total gain of $0.5\ V/nA$, the current noise caused by the digitizer acquisition is $\frac{280\mu V}{0.5\ V/nA} \sim 1\ pA$. The current measurement range of the digitizer will be 40 nA, from $\pm\ 20\ nA = \frac{\pm 10\ V}{0.5\ V/nA}$, which will be digitalized into 2^{16} values. Hence, the current resolution is $\delta I \sim 1\ pA \sim 40\ nA/\ 2^{16}$ (Fig. 1.4a, δI). Therefore, the digitizer acquisition noise is similar to the δI, as it does not make sense to digitalize with a resolution δI smaller than the acquisition noise.

However, this is only the "theoretical" maximum current resolution that can be achieved. As shown in the next section (Sec. 1.3), other noise sources influence the measured current signal, and the magnitude of these noise sources is bandwidth-dependent. Therefore, time and current resolution are actually related. Measurements with higher bandwidths (higher temporal resolution) will come with the trade of a higher current noise; hence, the ability to distinguish small current deviations will be smaller.

The third type of resolution is a geometric resolution. The smallest constriction of the pore describes geometric resolution. The constriction is the part of the pore that contributes more to the total resistance of the pore, as the resistance of a conductor is given by $R = \rho \cdot L/A$. Where ρ is the conductor's resistivity, L its length, and A its cross-section. When the molecule passes through the constriction, it produces a drop in the current, as it partially blocks the ion flow through the pore; the narrower and shorter this constriction, the bigger the current blockade and the higher the geometric resolution. The amplitude of the spike is, in a first approximation,

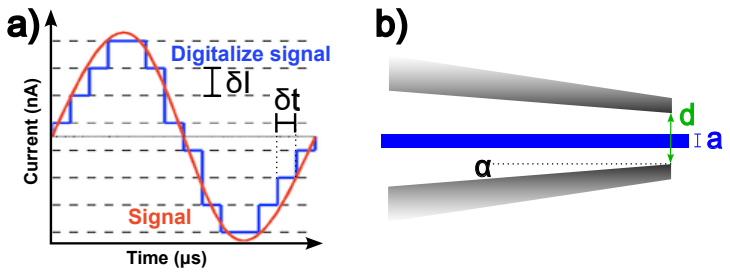


Figure 1.4: **Temporal, amplitude, and geometric resolution.** a) Temporal and amplitude resolution that is determined by the digitizer. b) The geometric resolution is determined by the nanopipette geometry.

proportional to the ratio of the cross-section between the analyte ($\propto a^2$) and the cross-section of the pore ($\propto d^2$). So that it is given by $\Delta I \propto a^2/d^2$ (Fig. 1.4b). However, it should be pointed out that this is only valid when the molecules are longer than the pore ($L_{\text{mol}} \gg L_{\text{pore}}$) so that it fills the whole pore length. As shown by Meller et al. [28], by translocating oligonucleotides of varying lengths ranging from $5 - 50$ nucleotides through an α-hemolysin pore, the current blockade (ΔI) was constant for sizes over 12 nucleotides. This experiment showed that when the oligos were long enough to occupy the total length of the α-hemolysin pore ΔI would become independent of the molecule length. This is especially important when working with nanopipettes, as they usually have a conical geometry with a small angle α. In this case, the sensing region length of the nanopipettes can be a few hundred nanometers, in contrast to the case of solid-state nanopores (or α-hemolysin), where the sensing length is given by the thickens of the membranes, which is usually a few nanometers [29].

1.3 Noise in nanopore experiments

In nanopore experiments, the ionic current driven through the pore by a constant applied bias voltage is measured. Any departure from the baseline current for the open-pore current measurement might be considered noise in the absence of analyte molecules. Noise is described by its power spectral density (PSD), which gives the contribution of noise to the variance of the signal per frequency (nA^2/Hz), as some frequencies are more prone to introduce noise into the signal. Therefore, the variance σ^2 of the signal is given by the integral of the PSD over the full bandwidth. The PSD is calculated by Fourier transforming the acquired current signal over time. Noise is generally undesirable because it can obscure or distort the true signal. The typical PSD of a nanopore experiment can be observed in Fig. 1.5a.

The noise sources in nanopore experiments can generally be divided

into (1) low-frequency $1/f$ noise, (2) frequency-independent noise sources like shot noise and thermal current noise, (3) high-frequency noise, dielectric and capacitative noise (see Fig. 1.5b). However, as will be seen in Chapter 4, these ranges are estimates, as they depend on the nanopore type and the experimental conditions.

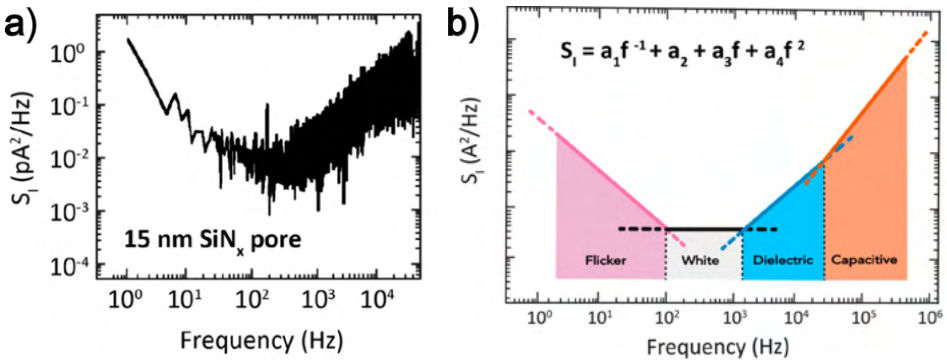


Figure 1.5: a) Current PSD for a 15.6 nm solid-state nanopore in a Silicon Nitride membrane (SiNx). b) Schematic of the current PSD for a typical nanopore. Common types of noise are highlighted in the various frequency ranges. Images from [30]

The different types of noise in nanopores are:

1. **Flicker noise or 1/f noise**. $1/f$ noise, commonly known as flicker or pink noise, is usually the predominant source of noise. Its power drops with frequency, following a $1/f^{\beta}$ scaling. Although this kind of noise is present in many physical and biological systems, it is still not well understood. $1/f$ noise in nanopores has been associated with slow fluctuations in the number and mobility of the charge carriers [31, 32], nanometer-sized bubbles in the pore channel [33], noise arising from the electrodes [34], mechanical fluctuations [35], etc. It was found that Hooge's phenomenological equation [31] could effectively describe the 1/f noise in solid-state [34] nanopores. The electric current power spectral density (S_I) of nanopores in the low-frequency range is described by [36, 31]:

$$S_{I,1/f} = \frac{\alpha_H I^2}{f^{\beta} N_C} \tag{1.1}$$

where Hooge's parameter, α_H, is an empirical parameter that quantifies the magnitude of 1/f noise fluctuations, I is the electric current, and N_c the number of charge carriers in the pore volume [34]. For nanopores $\beta \sim 1$ [30].

2. **Thermal Noise.** In the mid-frequency range, a frequency-independent white noise, also known as Johnson-Nyquist noise, is observed. This thermal current noise is fundamental to any dissipative element in a circuit [37, 38], and adds to the current noise as:

$$S_{I,\text{thermal}} = \frac{4k_B T}{R} \tag{1.2}$$

where k_B is the Boltzmann constant, T is the temperature, and R is the resistance of the nanopore.

3. **Shot noise.** Shot noise is another frequency-independent noise due to the quantization of charge and is generated when charge carriers flow across a potential barrier [39]. Its current-dependent contribution to the noise can be expressed as:

$$S_{I,\text{shot}} = 2Iq \tag{1.3}$$

where q is the charge of a single carrier. Shot noise and thermal noise are comparable in magnitude for the conditions that are typically used for nanopore experiments.

4. **Dielectric Noise.** Dielectric noise originates from the loss of conductance of the membrane and chip support. This noise can be described by:

$$S_{I,\text{dielectric}} = 8kT\pi C_{chip} D f \tag{1.4}$$

where C_{chip} is the parasitic capacitance, and D is the dissipation factor of the dielectric materials constituting the membrane and support chip, which is dimensionless. For the best solid dielectric (such as, for example, quartz), the dissipation factor is of the order of $10^{-4} - 10^{-5}$, while for poorer dielectrics it can be as high as 0.01-0.1 [40].

5. **Capacitance Noise.** For higher frequencies ($f \gtrsim 10$ kHz), the current noise is determined by the input voltage noise thermal voltage noise v_n across the total capacitance C_{tot} at the amplifier input:

$$S_{I,\text{cap}} = 4\pi^2 {C_{tot}}^2 {v_n}^2 \cdot f^2 \tag{1.5}$$

where v_n is the input voltage noise ($3\ nV/Hz$) for the commonly used amplifier Axopatch 200B [30], Molecular Devices, San Jose, USA). C_{tot} is the total capacitance, including the membrane and support chip capacitance C_{chip}, the capacitance C_{amp} at the input of the amplifier, and the capacitance C_w of the wiring between the electronics and the pore.

As we will see in Chapter 4, $1/f$ noise and frequency-independent sources are the main sources of noise in our nanopipette experiments. Capacitive and dielectric noise are not the dominant noise sources in our experiments. Our experiments are performed using a 35 kHz low pass filter. Nanopipettes are made from quartz capillaries, and quartz has an extremely low dissipation factor (D). Moreover, for nanopipettes, the chip capacitance (C_{chip}) is also low, depending on the size and thickness of the chip where the nanopore is inserted. Nanopipettes have tiny tips with thick walls, so their capacitance is also low. Therefore, dielectric noise is also not important below 35 kHz in our setup.

1.4 Nanofluidics phenomena relevant for nanopores

In this section, several phenomena related to the field of nanofluidics will be explained. These phenomena are caused by the influence of surface charge when performing electrical measurements at the nanoscale. At the nanoscale, surface effects become particularly relevant as they scale quadratically with the system size, while bulk effects scale with the cubic power. Therefore, these effects become relevant only at the nanoscale and overcome bulk effects. We will start with a brief introduction of several length scales that are crucial for determining the importance of the surface charge when studying ion transport, which is a problem on the nanoscale. Next, some relevant phenomena that arise from the relation between these length scales will be presented.

1.4.1 Electrostatic length scales and their influence

The length scales presented in this section depend on the ionic strength of the solution in which the surface is placed and on its electric charge. We will briefly mention in which conditions these effects are relevant for nanopipette experiments.

Debye length scale

Fixed charges on a surface attract oppositely charged ions in the solution, creating an electrical double layer (EDL) to maintain electrical neutrality. The EDL is the region where the surface charge is balanced/screened by the cloud of counterions (ions with an opposite charge to the surface). In this region, the ion concentration profile is different from the bulk values because of the interaction of the ions with the surface charge. The concentration of counterions is higher than that in bulk, whereas the concentration of co-ions

(ions with the same charge as the surface) is lower than that in bulk in this region (Fig. 1.6a). Therefore, the EDL is a region in the electrolyte solution where local charge neutrality is not held, as a local charge is present in this region due to the counterions/co-ion imbalance.

The Debye length emerges naturally from the Poisson–Boltzmann theory and characterizes the electrostatic screening of surfaces in the electrolyte solution [41]. It is defined in terms of the salt concentration ρ_s according to the expression:

$$\lambda_D = (8\pi\ell_B\rho_s)^{-1/2} \tag{1.6}$$

The value of Debye length (λ_D) is inversely proportional to the square root of the bulk salt concentration (ρ_s) and typically ranges from tens of nanometers (30 nm for $\rho_s = 10^{-4}mM$) and angstroms (0.3 nm for $\rho_s = 1M$). It is important to note that the Debye length is independent of the surface charge density σ. The Bjerrum length (ℓ_B) is defined as the distance at which the electrostatic interaction between two charged species becomes of the order of the thermal energy for two ions, with valence Z, embedded in a dielectric medium with dielectric constant ε (e is the elementary charge). The Bjerrum length is given by:

$$\ell_B = \frac{Z^2e^2}{4\pi\varepsilon k_B T} \tag{1.7}$$

with k_B the Boltzmann constant and T the temperature.

The Debye length is especially relevant at low concentrations. Inside nanopipettes, when the λ_D of the nanopipette's inner surface is large enough, the Debye layer overlap occurs. This enhances the nanofluidics phenomena we explain in the next section.

The Dukhin length

The Dukhin length is defined based on the comparison between the bulk and the surface electric conductance when the electric current through a nanochannel is taken as a consequence of an applied electric field. The Dukhin length characterizes the channel scale, below which surface conduction dominates over the bulk one. It is given by:

$$\ell_{Du} = \frac{|\Sigma|}{\rho_s} \tag{1.8}$$

where $|\Sigma|$ and ρ_s are the surface and bulk charge densities. To put it in numbers, for a surface charge density $\sigma = 20\ mC/m^2$ ($\approx 0.12\ e/nm^2$), the Dukhin length is $\ell_{Du} \sim 0.2$ nm, for $\rho_s = 1$ M, while $\ell_{Du} = 2\ \mu m$ for $\rho_s = 10^{-4}$ M.

1.4.2 Relevant nanofluidics effects

Electro-osmosis

Electro-osmosis (EO) is the phenomenon by which an electric field induces the flow of a liquid. As previously mentioned, inside the EDL there is a non-vanishing charge density, $\rho_e = e(\rho_+ - \rho_-)$. When an electric field (E_e) is applied, a net force $F_e = \rho_e E_e$ drives the motion of the fluid in the EDL. This motion can drag the fluid outside the EDL, farther away from the surface and start moving it. The final velocity of the fluid results from the balance between the driving electric force and the friction force of the surface. The Stokes and Poisson equations can be used to calculate the velocity profile for fluids containing ions. However, the velocity of the fluid far from the surface is given by the simpler Smoluchowski equation [42]:

$$v_\infty = -\frac{\varepsilon\zeta}{\eta}E \tag{1.9}$$

where E is the electric field parallel to the surface, ε is the dielectric constant, η the viscosity and ζ the so-called zeta (electrostatic) potential. ζ is the electrostatic potential at the ''shear plane'', i.e., the position close to the wall where the hydrodynamic velocity vanishes. The velocity profile of electro-osmotic flow as a function of the distance to the surface, $v(z)$, is shown in Fig.1.6b.

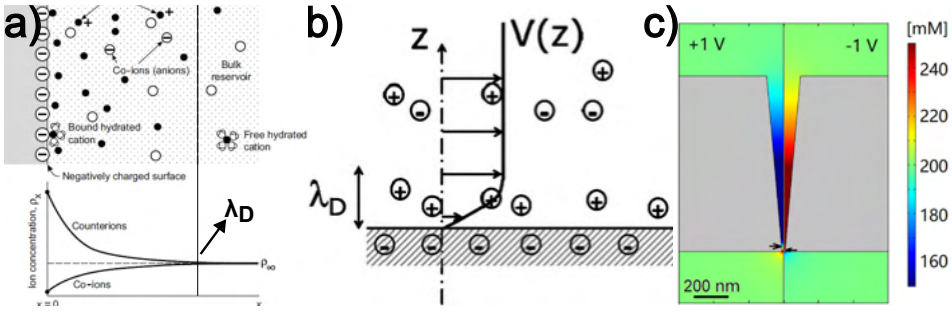


Figure 1.6: a) Debye layer (λ_D) of a surface in a solution containing ions. The concentration distributions of counter ions and co-ions are shown in the bottom plot. b) Velocity ($v(z)$) profile caused by electro-osmosis flow due to the presence of a charged surface. c) Ion concentration profile for two opposite voltage polarities $\pm 1V$ for a nanopipette in a bulk concentration of 100 mM. The scale on the right is the total concentration of cations plus anions. Image from [43].

Donnan equilibrium

In nanopipettes or nanochannels, when the Debye layer of the surfaces is of the order of the channels or nanopipette dimensions, the Debye layers

overlap. Therefore, a concentration gradient is generated because of the charges brought into the nanochannel to screen off the nanochannel surface charge. This gradient generates a potential drop between the nanochannel interior and the external reservoir. The potential is called the Donnan potential, and the process by which it arises is called Donnan equilibrium. The Donnan potential occurs to maintain a uniform chemical potential of all the ions. It is important to note that the Donnan potential happens even when no voltage is applied to the nanopipette or nanochannel.

Ion current rectification

Due to the just mentioned Donnan equilibrium, asymmetric pore geometries or symmetric nanopores with nonhomogeneous surface charges lead to nonlinear diode-like voltage-current curves in symmetric electrolyte conditions. This effect is called current rectification. Such ionic rectification has been specifically reported in quartz nanopipette [44, 45]. The magnitude of current rectification depends on the ionic strength of the solution, and the diode-like behavior becomes more pronounced with dilution when surface charge effects are more relevant.

Rectification occurs due to the following mechanism: The asymmetric charge distribution inside the nanopipette produces an asymmetric ion concentration and, therefore, a concentration gradient, as some regions need more counterions to screen the nanopipette charge [46]. This occurs without a voltage due to Donnan equilibrium. When a voltage is applied, it modulates the ion concentration inside the nanopipette. Voltages of different polarity modulate the ion concentration differently, such that for one polarity, the ion concentration at the nanopipette tip is higher than for the opposite polarity. This can be observed in Fig. 1.6c, which shows the case of an asymmetric nanopore in a 100 mM salt concentration. In the figure, it can be observed that positive polarity produces a lower ion concentration inside the nanopore than the negative polarity. This is known as a charge polarization effect. Therefore, the nanopipette conductance is higher for the negative polarity when the concentration inside the nanopipette is higher, and a higher current is measured for this polarity $|I(V_-)| > |I(V_+)|$. The rectification ratio is the ratio of currents recorded for voltages of similar amplitude but opposite polarities $|I(V_-)|/|I(V_+)|$.

Ion-specific effects

As explained before, the Debye layer is described using the mean-field Poisson–Boltzmann (PB) theory [41]. In this theory, ions are defined as point charges, and only their valency enters the description. Moreover,

correlations between charges are neglected. The theory considers the thermodynamic equilibrium balancing the ion's entropy to their electrostatic interaction with the surface (attraction for the counter-ions and repulsion for the co-ions), which leads to the concept of the Debye length.

However, the PB description misses an important class of effects associated with the specific nature of ions that affect the fine structure of the EDL [47, 48]. Such ion-specificity effects are, therefore, not predicted by the traditional PB framework. In this work, we will investigate the effect of different monovalent cations on DNA's charge.

1.5 Conductance Model For Nanopipettes

Continuing with the microscopy analogy, in super-resolution microscopy, it is important to have a good model for the point spread function of our microscope to calculate the position of fluorescent molecules with nanometer precision. In nanopore experiments, it is crucial to have a good model for the conductance (or resistance) of our nanopore. A good conductance model will relate the different parameters contributing to the nanopore conductance, such as size, geometry, charge distribution, or voltage. This model would allow us to determine some of these parameters when the other parameters are known. However, we lack a complete model describing the exotic nanofluidics phenomena we presented in the previous section, such as current rectification.

In particular, for experiments with nanopipettes, it is essential to estimate the tip diameter of our nanopipette from conductance measurements. Many translocation parameters, such as the translocation time or the translocation rate, depend on the nanopipette diameter. Nanopipettes are generated by pulling them from a quartz capillary so that no information about their size is available a priori. Measuring nanopipettes under an electron microscope involves cutting them to very small sizes and coating them so they cannot be used afterward. Therefore, a good option is to estimate their diameter using a model for their conductance.

1.5.1 Geometric model for nanopipettes

Nanopipettes have a truncated cone geometry that can be described by three parameters: the diameter at the tip of the cone (d), the diameter at the bottom (D), and its length (L). Using Ohm's law, $dR = \rho \cdot dl/A(l)$ and integrating it for a truncated cone geometry, the resistance of the cone R_c filled with a solution of resistivity ρ, can be calculated using:

$$R_c = \frac{4L\rho}{\pi dD} \simeq \frac{2\rho}{\pi d \cdot tan(\alpha)} \tag{1.10}$$

where α is the conical angle of the nanopipette (Fig. 1.7), and it is possible to approximate $tan(\alpha) = (D - d)/2L \simeq D/2L$ that is equivalent to assuming an infinitely long truncated pore, which is very reasonable for our nanopipettes as will be seen in 4.1.

Apart from the resistance due to the nanopipette geometry, the access resistance to the ends of the truncated cone [49] needs to be considered. This resistance considers the convergence of the field lines to the small nanopore. The access resistance (R_c) is given by:

$$R_a = \frac{\rho}{2d} + \frac{\rho}{2D} \simeq \frac{\rho}{2d} \tag{1.11}$$

It can be seen that only the access resistance to the small pore aperture is relevant. R_a is connected in series to the cone resistance (R_c) so that the total resistance is $R_{tot} = R_a + R_c$. The conductance of the nanopipette ($G = 1/R_{tot}$) can be written as [50]:

$$G = \frac{\pi g d D}{4L + \frac{\pi D}{2}} \simeq \frac{2\pi g d \tan(\alpha)}{4 + \pi \tan(\alpha)} \tag{1.12}$$

where g is the conductivity of the solution filling the nanopipette ($g = 1/\rho$). However, as it was mentioned in Sec. 1.4.1, this model does not include a major contribution to the conductance due to surface charge, which is especially relevant at low salt concentrations where $\ell_{Du} > d$. To include this effect in the model, we add a surface charge (σ) dependent term to the conductance [50, 29]. The new conductance considers an extra term that corresponds to an extra resistance in parallel to the nanopipette's total resistance. This resistance comes from considering that a small portion at the tip of the conical nanopipette has a conductivity equal to $4|\sigma|\mu/d$, where μ is the mobility of the counterions close to the surface. Note that the term $4|\sigma|\mu/d$ has conductivity units (S/m). The extra term considers the movement of the ions inside the Debye layer, which are screening the surface negative charge. These counterions also contribute to the total residence of the nanopipette, especially at low concentrations where the conductivity (g) of the solution is small. The dimensions of the nanopipette region where surface conductance is relevant are given by d, D^* and L^*, and are shown in Fig. 1.7. This portion is the sensing region of the nanopipette, which is the region that mainly contributes to the nanopipette resistance. The conductance is then given by:

$$G = \frac{1}{R_{total}} = \frac{2\pi g d \tan(\alpha)}{4 + \pi \tan(\alpha)} + \frac{\pi d D^*}{4L^*} \frac{4|\sigma|\mu_+}{d} \tag{1.13}$$

This extra conductance comes from the extra counterions that accumulate inside the nanopipette to screen the nanopipette's inner surface charge (σ).

In Eq. 1.13, it has been considered that the surface has a negative charge, so that μ_+ is the mobility of the counterions (cations) in the Debye layer. Although complete, this model still misses voltage-dependent effects on the nanopipette conductance, such as the rectification effects mentioned in Sec. 1.4.2. This model can be used to predict the diameter at the tip of a nanopipette if the solution's conductivity and the nanopipette material's surface charge are known. The latter are described in the next section.

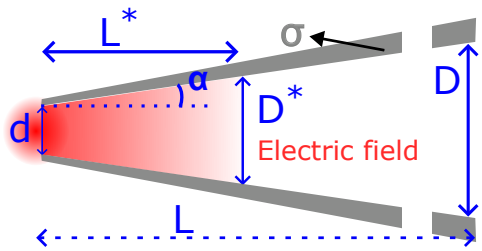


Figure 1.7: **Nanopipette geometric model.** Nanopipette with a surface charge (σ), tip diameter (d), length (L) and cone angle (α). D is the diameter at the bottom of the nanopipette. L^* and D^* are the length and bottom diameter of the sensing region of the nanopipette.

1.5.2 Modelling Surface Charge

The present conductance model contains the surface charge, which depends on the solution conditions, like pH or ion concentration. Charged surfaces in solution acquire their charge due to ionizable sites. These sites, which are in contact with the solution, can become dissociated by losing an ion into the solution. However, these binding sites are rarely fully dissociated, and their dissociation fraction depends on the solution conditions. Therefore, the surface charge (σ)and electric potential at the surface (ψ_0) depend on the solution's conditions.

Surface Chemistry

In quartz, each silicon atom is covalently connected to four oxygen atoms SiO_4. Let's consider a quartz surface in a solution containing monovalent and divalent ions. When the quartz is in contact with water molecules, silanol groups ($Si-OH$) form at the surface due to the chemical reaction $Si-O-Si \ + \ H_2O \rightleftarrows 2Si-OH$. Therefore, the only exchangeable ions with the surface are the very reactive protons in the solution (H^+); the other ions in the solution are inert ions that cannot bind to the surface to modify their charge. In this case, the surface charge is controlled by occupancy of the silanol groups ($Si-OH$), which is regulated by the chemical reaction:

$$\mathrm{SiOH} \overset{K_d}{\rightleftarrows} \mathrm{SiO^-} + \mathrm{H^+} \tag{1.14}$$

where K_d is the surface dissociation constant and is given by:

$$K_d = \frac{[\mathrm{SiO^-}]_0[\mathrm{H^+}]_0}{[\mathrm{SiOH}]_0} = \frac{\alpha}{1-\alpha}[\mathrm{H^+}]_\infty e^{-\frac{e\cdot\psi_0}{k_B\cdot T}} \tag{1.15}$$

where subindex 0 indicates the values at the surface, and the subindex ∞ the values in bulk. α is the surface association fraction ($\alpha = \sigma/\sigma_0$), ψ_0 is the electric potential at the surface and e the elementary charge. The proton concentration in bulk is given by $[\mathrm{H^+}]_\infty = 10^{-pH}$. If the surface is in chemical equilibrium, and K_d remains constant as pH changes, then the surface charge is given by [51]:

$$\sigma = \alpha\sigma_0 = \frac{K_d\sigma_0}{K_d + [H]_\infty e^{-\frac{e\psi_0}{k_BT}}} = \frac{\sigma_0 \cdot 10^{-pK}}{10^{-pK} + 10^{-pH}e^{\left(-\frac{e\cdot\psi_0}{k_B\cdot T}\right)}} \tag{1.16}$$

the pK is the pH value at which half of the surface sites are dissociated (so that $\alpha = 0.5$). It is related to the surface dissociation constant by $K_d = 10^{-pK}$. Eq. 1.16 can be rewritten to express the surface potential in the function of the surface charge:

$$\psi_0 = \frac{k_BT}{e}\ln\left(\frac{\sigma}{\sigma_0 - \sigma}\right) + \frac{k_BT\ln 10}{e}(pK - pH) \tag{1.17}$$

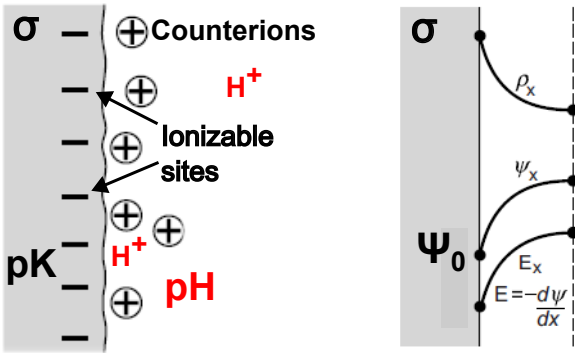


Figure 1.8: **Gouy-Chapman model.** a) Regulation of the surface charge (σ) of a surface with a given pK via protonation/deprotonation of the ionizable sites of the surface, controlled by the solution pH. b) The distribution of counter ions ρ_x, electric potential (ψ_x) and electric field (E_x), away from the surface

Grahame Equation

The Grahame equation relates the charge of a surface in a saline solution with the potential at the surface. For a negatively charged surface in a solution that contains monovalent and divalent ions, the Grahame equation is [51]:

$$\sigma = \sqrt{8\varepsilon_0\varepsilon k_B T}\sinh\left(e\psi_0/2k_BT\right)\left\{\left[H^+\right]_\infty + \left[M^+\right]_\infty + \left[D^{2+}\right]_\infty\left(2 + e^{-e\psi_0/k_BT}\right)\right\}^{1/2} \tag{1.18}$$

where $\varepsilon_0\varepsilon$ is the permittivity of the solution, and ψ_0 is the electric potential directly at the surface. $[H^+]_\infty$ is the concentration of protons in the solution that is given by $[H^+]_\infty = 10^{-pH}$. $[M^+]_\infty$ is the monovalent salt concentration, and $\left[D^{2+}\right]_\infty$ is the divalent salt concentration. Examples of salts could be NaCl and $MgCl_2$.

The model presented in this section is the Gouy-Chapman model, which works well for surfaces with low surface charges. In this model, the concentration of ions ($c_\pm(x)$) away from the surface flows an exponential function, as seen in Fig. 1.8 given by the equation:

$$c_\pm(x) = c_0 e^{\mp\frac{\psi e x}{k_B T}} \tag{1.19}$$

where c_0 is the concentration of ions in the bulk, and e the elementary charge. As the potential (ψ) is negative for a negatively charged surface, the concentration of cations (+) at the surface is larger than the concentration of anions (-).

Calculating the surface charge

The surface potential (ψ_0) is then calculated by solving the equation 1.20, and can be then used to calculate the surface charge (σ) using Grahame equation 1.18.

$$f(\psi_0) = \frac{\sigma_0 \cdot 10^{-pK}}{10^{-pK} + 10^{-pH}\cdot\exp\left(-\frac{\psi_0}{25.7}\right)} - 0.117\sinh\left(\psi_0/51.4\right)\sqrt{[M^+]_\infty + [D^{++}]_\infty\cdot\left(2+\exp\left(-\frac{\psi_0}{25.7}\right)\right) + 10^{-pH}} = 0 \tag{1.20}$$

In this equation, ψ_0 is in mV units, the concentrations are in molar concentration, and the surface charge is in C/m^2.

To get an idea of how the charge of the surface changes with the salt concentration of monovalent or divalent ions in solution, let's solve the equations for the specific case of a surface that has a charge of $-20\ mC/m^2$ when

fully dissociated (σ_0), and $pK = 7.5$. These are common values for surface charge and pK. The equations are solved for pH values of 7, 8, and 9, for a salt concentration range 10^{-4} to 10 M of monovalent or divalent salts.

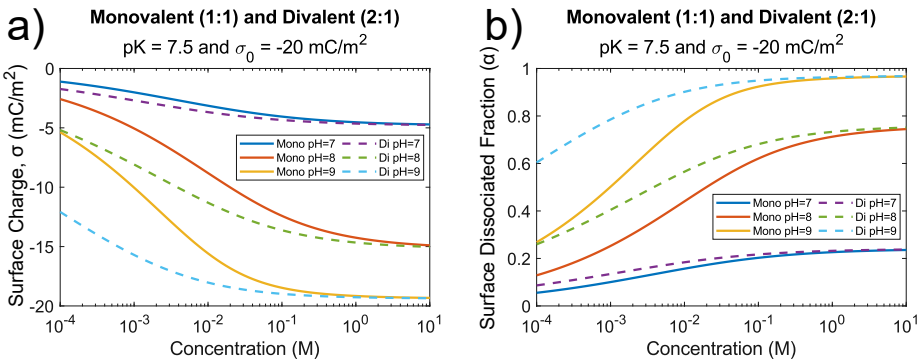


Figure 1.9: a) Surface charge (σ) of a quartz surface with a $pK = 7.5$ and $\sigma_0 = -20\ mC/m^2$, at pH's 7, 8, and 9, for monovalent (1:1) and divalent (2:1) salts. b) Surface association fraction (α) for the same conditions as a).

From Fig.1.9, it can be seen the effect that pH, salt concentration, and ion type have on the surface charge. The higher the pH (lower concentration of protons H^+ in the solution), the fewer protons are available to bind to the sites to neutralize the charge partially. So, the negative surface charge is higher (σ), and reaction 1.14 is shifted to the right, increasing the dissociation fraction (α). When the concentration of inert ions (monovalent or divalent ions that do not bind to the surface) increases, they displace some protons from near the surface. The electric potential attracts all the positive ions in the solution to the surface, such that the Debye layer is mainly conformed by these cations. The protons compete with the inert ions to be near the surface. When the inert ion concentration increases, this increases the portion of inert ions in the Debye layer, decreasing the concentration of protons near the surface. The displacement of protons from near the surface (decrease of $[H^+]_0$) shifts the chemical equilibrium 1.14 to the right so that the surface charge also increases. Finally, because divalent ions are more effective at displacing protons because of their higher charge, the surface charge is higher for divalent ions than monovalent ions for the same conditions.

1.5.3 Conductivity of Salt Solutions

The study of the conductivity of ionic solutions is crucial for science and technology. The ability of ions in solution to carry and transfer electricity influences many processes in various fields of physics and chemistry. Because of this, various theoretical treatments, such as Debye-Hückel [52],

have focused on explaining why and how conductivity decreases with the addition of more ions to a solvent.

The model for the nanopipette conductance presented in Eq. 1.13 contains the conductivity of the salt solution filling the nanopipette. In the nanopore field, the conductivity of the ionic solution filling the nanopore is generally assumed to be linear with the salt concentration [50, 26]. Therefore, it is given by $g_L = (\mu_+ + \mu_-)n(c) \cdot e$, where $n(e)$ is the number of ions per volume that depends on the salt concentration (c), e is the charge of the ions, and μ_+ and μ_- are their electrical mobilities. The values of some ion mobilities relevant to this work are presented in Table 1.1. However, as will be shown, this simple approximation fails, especially when working at high salt concentrations in the $1 - 4$ M range. The study of theoretical models to account for the decrease in conductivity at higher salt concentrations is a complex task and is out of the scope of this work. Here, we used empirical data to model the nonlinear conductivity ($g_N L$) of LiCl, NaCl, KCl, and $MgCl_2$ solutions at different concentrations. Data are taken at 298 K ($25°C$), and are taken from [53, 54]. The data are then interpolated with a spline function to extract the conductivity at each desired concentration.

Table 1.1: **Ion mobilities at infinite dilution at** 25°. Mobility of ions, μ_+ and μ_-, present in salt solutions of LiCl, NaCl, KCl, and $MgCl_2$. Mobilities are in $10^8 \times (m^2/s \cdot V)$ [55].

Cl^-	Li^+	Na^+	K^+	Mg^{2+}
7.91	4.01	5.19	7.62	11.0

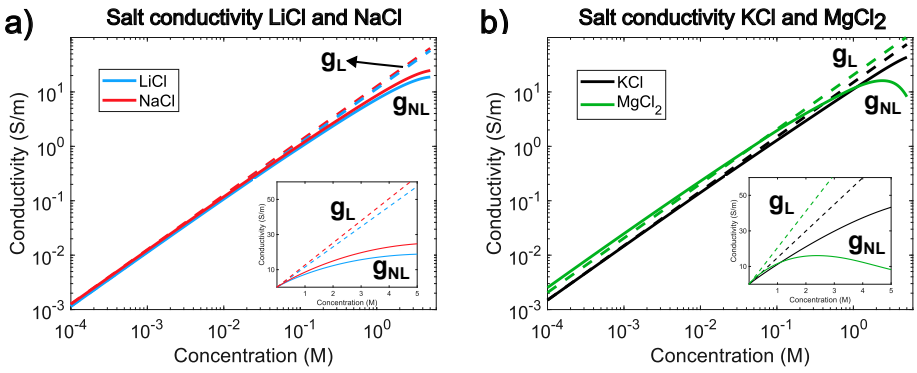


Figure 1.10: **Salt solutions conductivity models.** a) Linear (g_L) and non-linear (g_{NL}) conductivity models for LiCl and NaCl. b) Same for KCl and $MgCl_2$. Conductivities are at 298 K ($25°C$). The linear models are obtained by using the Eq. $g = (\mu_+ + \mu_-)n(c) \cdot e$, and the ion mobilities in Table 1.1. The non-linear models are obtained by interpolating the conductivity data in [53, 54].

Figure 1.10 shows a comparison of the linear conductivity model (g_L, dashed lines), using the ion mobilities values at infinite dilution (Table 1.1), and the non-linear conductivity model (g_{NL}, straight lines in Fig. 1.10) for four different salt types. It can be observed that there is a big difference between the two models at high concentrations. Table 1.2 shows the concentrations for which the linear conductivity (g_L) deviates by more than 20% from the non-linear conductivity (g_{NL}).

Table 1.2: **Comparison linear and nonlinear conductivity models.** Concentration for which there is a 20% difference between salt conductivity models.

	LiCl	NaCl	KCl	$MgCl_2$
$\lvert g_L - g_{NL}\rvert / g_{NL} > 0.2$	$105mM$	131 mM	191 mM	226 mM

In Chapter 4, we will present results supporting the non-linear conductivity models in nanopore experiments over the linear model. The non-linear model will be then used throughout this work to predict the conductivities of these four salt solutions.

1.5.4 Electric field at the tip of the nanopipette

Apart from modeling the nanopipette conductance, another important quantity that needs to be considered is the electric field at the tip of the nanopipette. The strength of the electric field will determine the strength of the electrophoretic force pulling on DNA during DNA translocation. For a conical geometry, using the continuity equation $j = g \cdot E = I/A$, that relates the ion flux (j) with the electric field (E), the solution conductivity (g) and the electric current (I) through a nanopipette section (A). The electric field at a distance (x) away from the nanopipette tip is given by:

$$E(x) = \frac{I}{g \cdot A(x)} = \frac{VdD\pi}{4L \cdot A(x)^2} \tag{1.21}$$

Here, we have used $I = V/R_c$ and Eq. 1.10 for the resistance (R_c) of the conical nanopipette. Moreover, d, D, and L are the dimensions of the conical nanopipette, shown in Fig. 1.7. The electric field at the tip ($x = 0$) is:

$$E_{\text{tip}} = \frac{2Vtan(\alpha)}{d} \tag{1.22}$$

where α is the cone angle, and the approximation $tan(\alpha) \approx D/2L$ has been used. The electric field does not depend on the salt concentration and is inversely proportional to the nanopipette tip diameter. In consequence, smaller nanopipettes have a higher electric field at the tip.

1.6 DNA translocation through nanopores

Nanopores have emerged as a groundbreaking molecular biology and biophysics tool, offering unparalleled versatility and sensitivity for studying single molecules. These nanoscale channels allow us to detect and study single molecules in real time without chemical labeling. Common applications are the study of DNA, where nanopores have proven invaluable for sequencing [18, 56], studying DNA configuration [57, 17], or studying DNA properties like charge [50, 29], or complexation with ligands [58].

A voltage is applied across the nanopore during DNA translocation while measuring the electric ion current through the nanopore. When DNA molecules get near the tip of the nanopore, the strong electric field generated at the pore captures and threads the DNA through the nanopore. During the translocation of the DNA molecule, the electric current is affected by the presence of the DNA inside the nanopore, and a current blockade is observed (Fig. 1.11a). The time it takes for a DNA molecule to translocate depends on the experimental conditions. From now on, we will refer to this as the dwell time (t_D) of DNA inside the nanopore (Fig. 1.11a). Here, we will briefly present some previous results studying the current blockade (or enhancement) during DNA translocation, the dwell time, and the DNA translocation rate.

1.6.1 DNA current signal

DNA has been shown to reduce or increase the electric current when translocating through nanopores, depending on the experimental conditions [50]. When translocating in solutions at a low salt concentration (below 150 mM), the electric current, and thus the electric conductance ($\Delta G = \Delta I/V$), increases, while it decreases above 150 mM. DNA is a highly negatively charged molecule, with one electron charge per phosphate group; therefore, at low salt, the presence of DNA inside the nanopore increases the number of ions. Hence, the electric current increases ($\Delta I > 0$) when DNA translocates. On the other hand, in solutions with high ionic strength, the DNA molecule excludes part of the volume and, consequently, reduces the amount of charge carriers inside the nanopore, reducing the electric current ($\Delta I < 0$). The following equation models this behavior for the case of a cylindrical nanopore [50, 59]:

$$\Delta G = \frac{1}{L}\left(-\frac{\pi}{4}d_{DNA}^2 g + \mu^* q_{DNA}^*\right) \tag{1.23}$$

where g is the conductivity of the salt solution given by $g = (\mu_+ + \mu_-)n(c)\cdot e$, L is the length of the cylindrical nanopore, and d_{DNA} is the DNA diameter (2.2 nm) (Fig. 1.11b). μ^* is the effective mobility of the cations screening

the DNA charge, and λ^* is the effective charge of DNA per unit length. If we now consider n strands of DNA inside the nanopore for high ionic strengths, where $d_{DNA}^2 g \gg \mu^* \lambda^*$, the relative current blockade when n strands simultaneously translocate, can be written as [60]:

$$\frac{\Delta I_n}{I_0} = \frac{\Delta G_n}{G_0} = -n \left(\frac{d_{DNA}}{d} \right)^2 \tag{1.24}$$

where we have used the result that the conductance of a cylinder is given by $G_0 = \pi g d^2 / 4L$. Therefore, G_0 is the conductance of the open pore and $I_= = 1/G_0$ is the current through it. This indicates that the conductance drop is independent of the concentration and only depends on the relative sections of the DNA and the cylindrical nanopore.

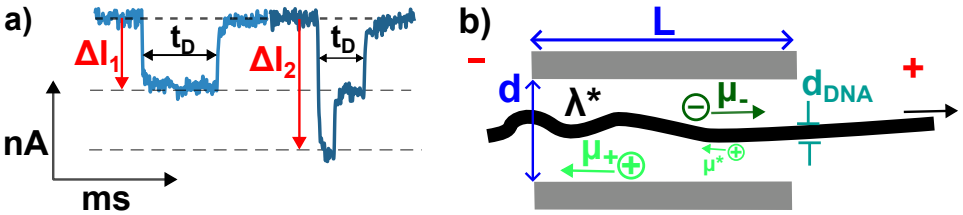


Figure 1.11: a) DNA translocation events, the discretized levels of the current blockade ΔI_n are shown for the events. b) Scheme of DNA, with diameter (d_{DNA}), translocating through a nanopore with diameter d and length L. The movement direction of DNA, cations (μ_+), and anions (μ_-) is indicated. The lower mobility of cations shielding DNA is represented by the smaller mobility (μ^*).

1.6.2 Dwell times

DNA dwell times have been extensively studied. Dwell times have proven to be dependent on multiple experimental parameters like salt concentration [61], salt type [61], voltage [62], pore diameter [63], and solvent viscosity [62]. The dependence on pore diameter is not straightforward. For nanopores slightly larger than DNA's width ($\sim$ 2 nm), dwell times increase for smaller diameters due to the stronger DNA-pore interaction [64]. In this range, dwell types are dependent on the polynucleotide sequence due to different interactions of each nucleotide (A, T, C, or G) with the nanopore [65]. However, when pores are larger than DNA's width, pore-DNA interactions are weaker and less frequent. In this case, smaller diameters produce faster translocation times, as the higher electric field in the smaller pores produces a higher driving force. Fig. 1.12 shows the dwell-time distribution for λ-DNA in a 10 nm nanopore, where DNA-pore interactions do not occur. It can be observed that the distribution has a peak at 2 ms. The presence of fragments can be observed from a large number of counts at very short dwell

times. Moreover, concatenates of multipleλ-DNA can be observed from the long tail of events with longer dwell times.

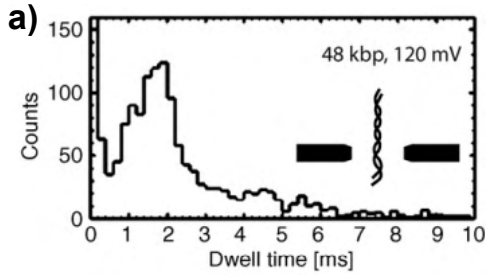


Figure 1.12: a) Dwell time distributions for linear λ-DNA dsDNA through a 10 nm nanopore [66].

It has been reported that average dwell times linearly increase with the inverse of voltage ($t_D \propto 1/V$), such that higher voltage produces shorter dwell times [67]. Moreover, viscosity also affects DNA dwell times, as viscosity increases the drag on the DNA molecules, increasing the dwell time [68], and shows a linear behavior $t_D \propto \eta$ [61]. Salt concentration also affects dwell times due to the stronger screening of DNA's electric charge. In [61], a relation of 1:1.5:2 was observed for dwell times in 1:2:4 M of LiCl.

1.6.3 DNA translocation rate

The process of DNA capture into nanopores has received considerable attention [69, 70]. For practical reasons, it is crucial to understand how long polymers find their way into nanopores. For low DNA concentrations, where molecules do not interact with each other, the process of capture is a Poisson process, for which the normalized probability of arrival follows an exponential distribution:

$$P(t) = e^{-Rt} \tag{1.25}$$

where t is the time delay between successive translocations, and R is the mean capture rate.

Furthermore, the dependence of the capture rate on the applied voltage is relevant. Two different behaviors, linear and exponential, have been observed for the mean translocation rates as a function of voltage. Each corresponds to a different process dominating the capture rates.

First, the linear dependence of capture rate with voltage has been attributed to the diffusion-limited capture process (Fig. 1.13a). In this case, rates are determined by the diffusion of DNA molecules into the capture radius of the nanopore [69]. The voltage profile away from the nanopore decays with the distance to the nanopore (x), following the equation [71]:

$$V(x) = \frac{d^2}{8hx}\Delta V \tag{1.26}$$

where d is the pore diameter, h is the pore height, and ΔV is the applied voltage.

When DNA comes close to the nanopore, its diffusion becomes biased due to the weak electric field near the nanopore, and the DNA migrates toward the nanopore. Since the bias diffusion depends on the strength of the electric field, the capture radius is linearly dependent on the voltage. Therefore, the capture rate linearly depends on the voltage (Fig. 1.13a).

Second, the exponential dependence of the capture rate has been interpreted as an activated process where DNA must overcome an entropic barrier to start the translocation process (Fig. 1.13b). Increasing the applied voltage reduces the entropic barrier for DNA entry to the pore [70]. Therefore, the capture rate increases exponentially with voltage.

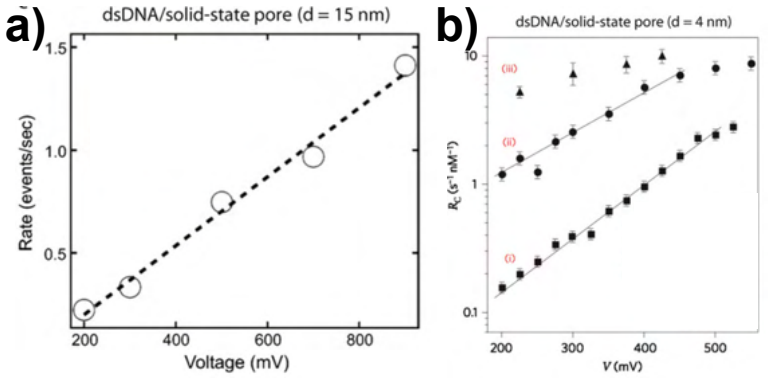


Figure 1.13: a) Capture rate as a function of voltage for λ-DNA for a 15 nm diameter solid-state nanopore [69]. b) Semi-log plot of the capture rate of (i) 400 bp dsDNA, (ii) 3500 bp DNA, (iii) 48.5 kbp λ-DNA into a 4 nm diameter nanopore [70].

For the specific case of λ-DNA, the voltage dependence of the capture rate has been studied in solid-state pores of 15 nm diameter [69]. A linear dependence was observed in the voltage range of 200–900 mV, suggesting that capture is diffusion-limited (Fig. 1.13a). The fact that no barrier was observed can be explained by the large nanopore diameter in the scale of the persistence length of DNA ($\sim$ 50 nm), which facilitates bending and DNA entrance into the nanopore [57, 72]. In contrast, an entropic barrier was observed for dsDNA in solid-state nanopores of 4 nm diameter (Fig. 1.13b) [70].

Chapter 2

The optical tweezers setup

This chapter provides an introduction to the foundational principles of optical trapping and force spectroscopy experiments. Optical tweezers (OT) have high force sensitivity, which makes them ideal for investigating biomolecules, including nucleic acids and proteins. The chapter is organized into different sections. First, the fundamental principles underlying optical trapping are outlined. This is followed by a detailed description of the experimental apparatus, called "mini-tweezers," utilized for the OT experiments in this work. Subsequently, the experimental protocols necessary for conducting OT experiments are presented. Finally, the chapter concludes with an explanation of the calibration procedures required for the mini-tweezers.

2.1 Optical trapping

Photons carry linear momentum. Therefore, a ray of light has a linear momentum ($\overrightarrow{p}$) given by:

$$\overrightarrow{p} = \mathrm{N}_\lambda \frac{\hbar}{\lambda} \hat{e_p} \tag{2.1}$$

where λ is the wavelength of light, $\hbar$ is Planck constant, N_λ is the number of photons of the light ray, and $\hat{e_p}$ is the unit vector that indicates the direction of propagation of the ray.

Light's momentum has important consequences when light interacts with matter, as it can exert forces on it. Here, the ray-optics approximation will be used to interpret the light-matter interaction. First, we consider the simple case of a particle in a collimated laser beam. Figure 2.1a shows a collimated beam of light interacting with a spherical particle or bead, sitting off-center of the beam axis, which has an index of refraction higher than the surrounding medium ($n_p > n$). The beam has a Gaussian intensity profile. Let's assume light does not reflect from the bead, so all light rays

are transmitted. The incident rays of light (black lines in Fig. 2.1) change direction when crossing the particle's surface. This phenomenon is known as refraction and is described by Snell's law $sin(\theta_m)n = sin(\theta_p)n_p$. Light rays are deflected upon entering and exiting the particle. The change of direction of light is translated into a change of light momentum, from $\overrightarrow{p_i}$ to $\overrightarrow{p_t}$, being larger for the more intense ray (numbers 2 Fig. 2.1), as it carries more photons Eq. 2.1. Due to momentum conservation, the change of light momentum equals that of the bead ($\overrightarrow{p_{b,1}}$ and $\overrightarrow{p_{b,2}}$), producing a total momentum $\overrightarrow{p_b}$. According to Newton's second law, the change in momentum generates a force on the bead ($\overrightarrow{F_b}$), given by $\overrightarrow{F_b} = \frac{d\overrightarrow{p_b}}{dt}$. The particle is pushed towards the center of the beam, where the intensity is maximum, along the beam's axis direction. However, this scenario is not optimal, as the particle is only trapped in the optical plane but is pushed along the optical axis.

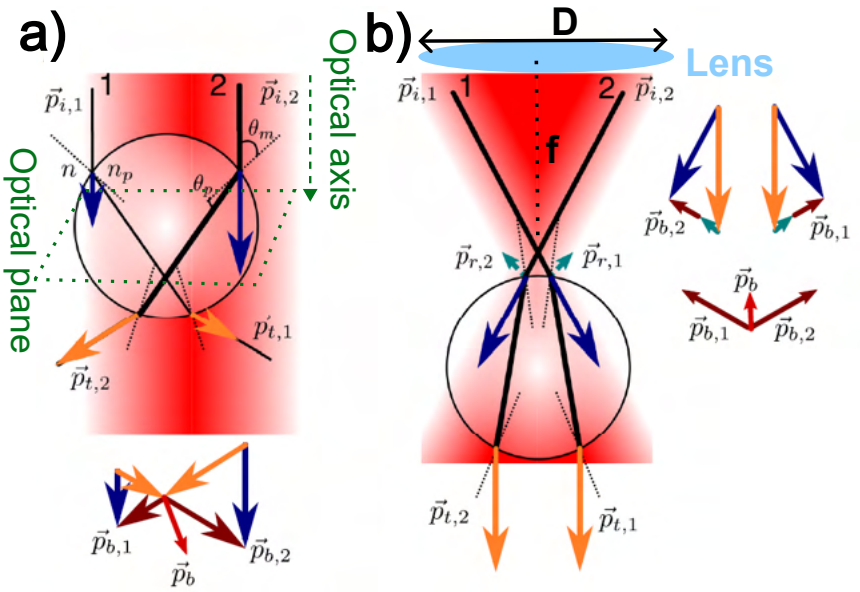


Figure 2.1: (a) Ray optics diagram of a particle in a collimated beam, with a Gaussian intensity profile. Rays are represented as black lines with a thickness proportional to their intensity. The arrows indicate the initial ($\overrightarrow{p_i}$, blue), transmitted ($\overrightarrow{p_t}$, orange) light momentum and the change in momentum of the particle, $\overrightarrow{p_b}$. (b) Particle in a focused Gaussian beam. Turquoise arrows ($\overrightarrow{p_r}$) are the weak reflected light's momentum change. Figure adapted from [73].

To create a stable trap that can trap objects in the three axes, we need to create an intensity gradient in the direction of the laser's propagation. This is achieved by focusing the laser beam using a microscope objective (Fig. 2.1b) of large numerical aperture $NA = D/f$, where D is the lens diameter and f its focal length. In this case, the scattering force is overcome by the

gradient force along the optical axis. In this case, the particle experiences a net force towards the focus point of the beam (red $\overrightarrow{p_b}$ in Fig. 2.1). Therefore, the particle will be trapped in a 3-D harmonic well, slightly beyond the focal point where the gradient and scattering force compensate each other. It is possible to trap objects by just using one objective and one laser beam. However, a more stable trap is achieved when two counter-propagating lasers are focused on the same spot.

Although, as previously mentioned, the ray-optics approach is only an approximation, it provides an intuitive explanation of optical trapping. To study objects smaller than the focus size given by the Rayleigh criterion $\sim \lambda/2$, the generalized Lorenz-Mie theory needs to be considered [74]. This theory uses the wave formulation of light derived from Maxwell's equations [75].

2.2 Mini-tweezers optical setup

This section explains the basics of the mini-tweezers setup, designed by Steve B. Smith and Carlos Bustamante in 2003 and adapted to our lab in 2005 [76, 77, 78]. The mini-tweezers have proven very useful for studying the properties of biomolecules [79, 80], such as DNA, RNA, and proteins, or bigger samples, such as red blood cells [81, 82]. Thanks to their high sensitivity, the mini-tweezers have been successfully used for the study of the folding and unfolding of nucleic acids and proteins under equilibrium and non-equilibrium conditions[83, 84] and to study the behavior of molecular motors[85].

As their name implies, the mini-tweezers is a very compact optical tweezers setup. The setup has two contra-propagating laser beams that form a more stable and stiffer optical trap than single laser setups. The compactness of the setup provides advantages regarding vibrations and versatility. Thanks to their small size, the mini-tweezers can be easily put inside a refrigerated box to make temperature-controlled experiments [86]. However, mini-tweezers also have drawbacks, as the compact space of the instrument renders modifications to the optical path and alignment more difficult than in setups built on optical tables. The main parts of the setups and how experiments are performed are explained in this section.

2.2.1 The Optical path

The setup has two infrared lasers ($\lambda = 845$ nm). It is important to note that the complete beam path described in this section appears twice in the setup, once for each laser (green and yellow paths in Fig. 2.2). The laser light is transmitted through an optical fiber to the wiggler, as seen in the scheme

of the optical path in Fig. 2.2). The wiggler is a cylindrical brass tube that contains optical fiber. The brass tube moves around a pivot point using two piezoelectric actuators that allow for precise movements of the laser beams in the plane perpendicular to the optical axis. The piezoelectrics move the optical trap inside the microfluidic chamber, with a movement range of about $\sim 8 - 12\ \mu m$. Let's pick the green color path in Fig. 2.2 for our explanation. The light from the optical fiber is first split using a thin pellicle; about $\sim 8\%$ of the laser light coming out of the fiber is reflected by and sent to a Position Sensitive Detector (PSD). The PSD are quadrant photodiode detectors that register the exact position of the laser spot on their sensor. These first detectors measure the laser position and are called position-PSDs. The remaining laser light's $\sim 92\%$ is collimated and directed to a polarizing beam splitter (PBS) that redirects the vertically polarized light to the objective (green light behind left objective Fig. 2.2). By changing the polarization of the light coming from the laser; we can regulate the amount of light reaching the microscope objective. The polarization is changed by bending the optical fiber using duct tape. The reflected light from the PBS is circularly polarized by a quarter wave plate ($\lambda/4$) before entering the objective.

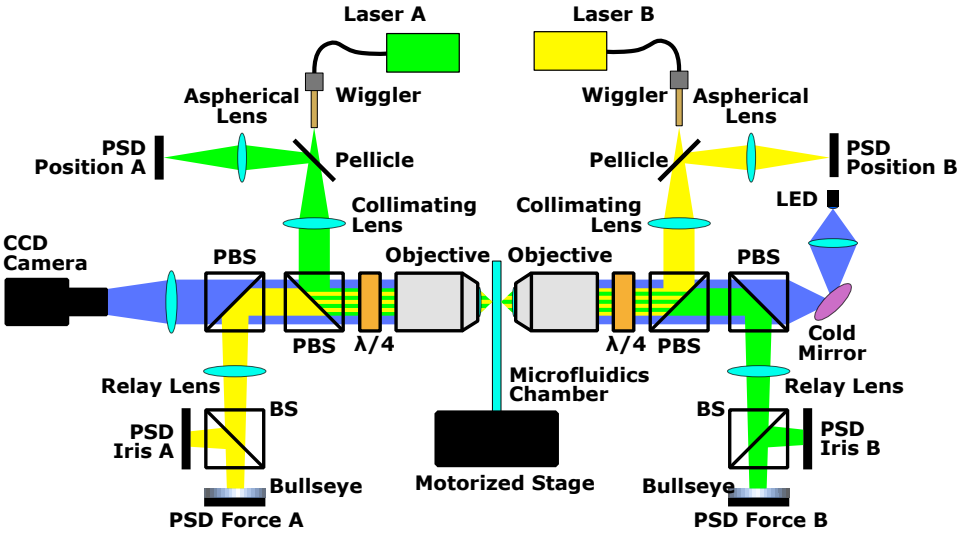


Figure 2.2: Schematics of the Mini-tweezers optical path. A detailed description can be found in the text.

Two high NA objectives (Olympus UplanApo 60x, with $NA = 1.20$) focus the laser's light inside the microfluidic chamber to form the optical trap. Light is then collected and collimated again by the opposed objective. A second $\lambda/4$ changes light polarization again, such that light now has a perpendicular polarization to the one it had after reflecting the first PBS.

The change in polarization is such that light gets now transmitted through the first PBS and reflected by a second PBS, which is placed in an opposite orientation. The light then reaches the relay lens, which focuses it on the PSD force detectors. A final beam splitter (BS) divides the light between both detectors, the PSD Iris and the Force PSD. This detector measures the force components. While the Iris measures the z component of the force along the optical axis, the Force PSD measures the x and y components of the force in the optical plane. However, the use of the Force PSDs relies on some tricks to perform the force measurements. The Iris PSD has a bullseye filter in front of it. When a bead in the optical trap gets displaced in the z-direction, the intensity profiles at the bullseye get broader or narrower. Hence, the iris receives more or less light intensity. The light intensity is used to determine the z-component of the force. The Force PSD detectors are positioned such that the laser spot on the detectors does not move when the lasers are moved. Only when an optically trapped object deflects light does the spot on the detector move. This is achieved by placing the Force PSD in a conjugated plane of the pivot point of the wiggler. Therefore, as the laser is moved by pivoting on this point, the point on the PSD does not move, and the light spot only moves on the detector when the bead is displaced in the optical trap. The principle is similar to that of the atomic force microscope (AFM), which measures force by the movement of the laser spot on a quadrant photodiode due to the deflection by the cantilever. Note that the setup includes all the mentioned optics twice, one for each laser. Additionally, an LED and a CCD camera are used to capture a live image of the microfluidic chamber.

In addition to the optics, the microfluidic chamber is the other crucial part of the setup because it is where the experiments are performed. As the mini-tweezers only generate one optical trap, the chamber contains a micropipette, in which a bead can be immobilized by air suction, as will be explained in detail in the next section. The microfluidic chamber is moved by a motorized stage that moves the chamber in all three dimensions. These motors can move the chamber a wider range, about ~ 0.5 cm in x and y, and a shorter distance in z $\sim 300\mu m$. Hence, when performing experiments, there are two independent ways to move the optical trap with respect to the micropipette and the microfluidic chamber. One is by moving the chamber with the motors, while the optical trap remains formed on the same spot in 3D space. The other way is to move the lasers using the piezoelectric, which moves the optical trap inside the microfluidic chamber. The latest has a much shorter range and is used to perform precise movements of the optical trap during the experiments. The motor movements are used to move longer distances and collect the beads used for the experiment.

2.2.2 The microfluidic chamber

The microfluidic chamber where the experiments are performed is placed in the narrow space between the microscope objectives (Fig. 2.2). As the objectives have a short focal distance, the chamber is very thin, about $\sim$ 300 μm. A sketch of the chamber is shown in Fig.2.3. The chamber has three channels. The central channel contains the micropipette and is where the experiments are performed. The two lateral channels are used to flow the two types of beads used for the experiment: streptavidin-coated beads (SA) and anti-digoxigenin-coated beads (AD). The beads are flown into the central channel through small dispenser tubes that connect the lateral channels with the central one. By using dispenser tubes, only a few beads reach the central channel. This is important because if the concentration of beads in the central channel is too high, they continuously fall into the optical trap, and performing experiments is impossible. Therefore, one needs to control the concentration of beads in the central channel during the experiments.

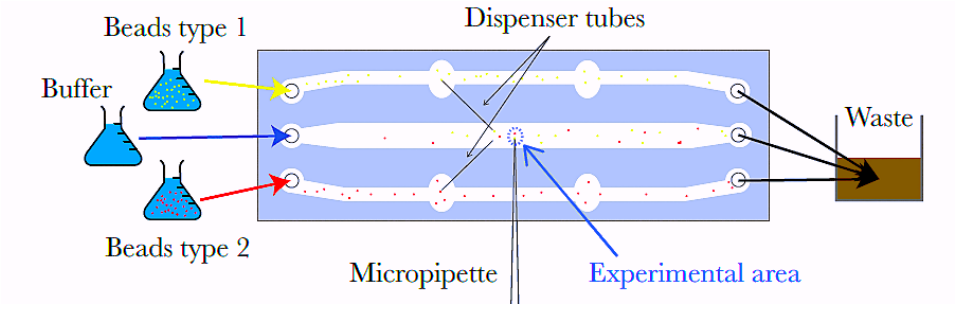


Figure 2.3: Schematics of the microfluidic chamber. Image from [87]

The microfluidic chamber is handmade from two glass coverslips and two pieces of parafilm. A laser cutter is used to cut the three-channel pattern into the parafilm pieces. Additionally, the laser cutter is used to make small holes into one of the coverslips that will be used to flow the buffers. The parafilm is then glued to the glass coverslips using UV glue (NOA61 from Thorlabs). This is done to achieve a good sealing between the glass and the parafilm. Then, the dispenser tubes and micropipette are placed as shown in Fig.2.3. For the micropipette and the dispenser tubes, thin glass capillaries with a 100 μm outer diameter are used. The pipette is pulled on a homemade coil-based pipette puller and should have a diameter of $\sim$ 1 μm at the tip. Once the pipette and tube have been placed, they are sandwiched between the coverslips. The chamber is then sealed by putting it on a hotplate at 120°, such that the two parafilms melt and fuse to each other. Finally, the glass capillary of the micropipette is threaded into a

plastic tube to seal the micropipette so that a micropipette can be used to trap beads by air suction onto the micropipette. UV blue is used to seal the capillary and the plastic tube.

2.2.3 Experimental configuration

The molecule to be studied must be flanked by dsDNA handles that connect the molecule ends to beads. The combination of the molecule with the handles is called molecular construction. It must be tethered between two beads in order to perform force spectroscopy experiments in the mini-tweezers. In the case of a DNA hairpin, the molecular construction consists of a hairpin between two dsDNA handles, as shown in image Fig.2.4. Each of the handles is labeled with biotin or digoxigenin at their ends. For the experiments, two types of beads were used. One type is the streptavidin-coated beads (SA beads), which are smaller in diameter,$\sim$ 2 μm. The others are anti-digoxigenin-coated beads (AD beads), about $\sim$ 3 μm in diameter, so they can be visually differentiated thanks to the CCD camera image. The labeled ends of the molecular construct bind to the corresponding bead. The biotin forms a biotin-streptavidin bond with the SA bead. The digoxigenin forms an antidig-digoxigenin bond with the AD bead Fig.2.4.

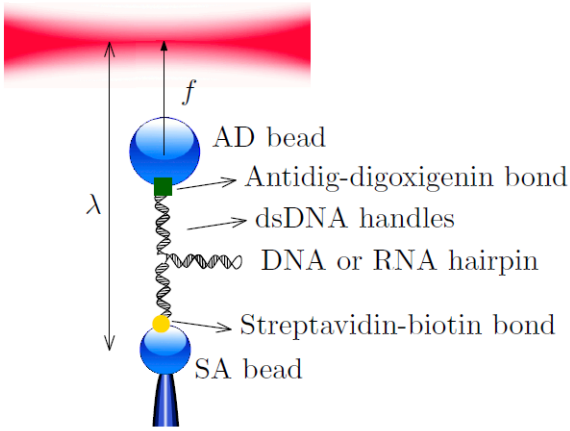


Figure 2.4: Setup to study the mechanical unfolding and folding of a nucleic acid hairpin (DNA or RNA).

To achieve the final arrangement of having the molecule tethered between two beads, with the SA bead immobilized on the micropipette and the AD in the optical trap, we proceed in the following manner. Before the experiment, the DNA molecule is incubated with the AD beads so the construct attaches to the AD beads. Then, during the experiment, the SA

beads are first flown into the microfluidic chamber. As the setup only has one optical trap, the SA beads are trapped and then brought to the micropipette and fixed on it by air suction (Fig.2.4 bottom). Afterward, the AD beads with the attached molecular construct are flown into the chamber and trapped with the optical trap. By bringing both beads together, the biotin-streptavidin bond forms between the construct and the SA bead on the micropipette. All these procedures are performed by watching the live image of the CCD camera and moving the microfluidic chamber with respect to the optical trap using the motors. Once the molecule is treated between the beads, the piezoelectrics move the optical trap to perform the desired experiment.

2.2.4 Calibration

One advantage of the 'mini-tweezers' is that they directly measure the force. Other OT setups measure the force by monitoring the position of the bead with respect to the optical trap using less precise imaging methods. In the mini-tweezers, the force components are measured by the Force PSDs (x and y components) and by the Iris PSD (z component). The measured force corresponds to the mechanical force components (x, y, and z) acting on the trapped bead. For the Force PSD it is given by:

$$f_{x|y} = C_{x|y} \cdot PSD_{x|y} \tag{2.2}$$

where $C_{x|y}$ is the calibration factor. These calibration factors are independent of the experiments and, therefore, independent of the bead diameter, solvent, and index of refraction of the bead or the medium. We will now briefly explain the two methods we used for calibrating the force [88]:

1. Stokes law. Moving the chamber with a bead in the optical trap generates a relative velocity between the bead and the surrounding fluid. Knowing the velocity at which the motors are being moved, the Stokes force ($f_{x|y}$) that is acting on the bead can be computed:

   $$f_{x|y} = \gamma \cdot v_{x|y} = 6\pi\eta R v_{x|y} \tag{2.3}$$

   where γ is the drag coefficient, $\eta = 0.89\ mPa \cdot s$ is the water viscosity at 298 K and $R = 1.5\ \mu m$ is the bead radius. By doing a linear fit between the force acting on the bead and the signal measured from the PSDs, the calibration constants ($C_{x|y}$) can be calculated as Fig.2.5a.

2. DNA overstretch. When dsDNA is overstretched by pulling on its ends, the molecule experiences a transition at high forces $65 - 70$ pN, which is observed as a plateau at a constant force (Fig.2.5b). Although

this transition depends on the experimental conditions [89, 90], for a sequence with $\sim$ 50% GC content, at 298 K and 1M NaCl, the transition happens at $\sim$ 67 pN Fig.2.5b. Therefore, the calibration factor can be computed by overstretching a dsDNA molecule; in our case, we use a 24508 bp fragment from λ-DNA and compare the force of the transition. Moreover, as the transition occurs at almost constant force, the measured distance corresponds to a change of molecular extension, that is, $\sim$ 0.25 nm/bp. Hence, the overstretch can also be used to check the distance calibration factor, as the plateau for a 24508 bp molecule should be $\sim 5.8\ \mu$ long.

Details for the calibration of the z component of the force can be found [91].

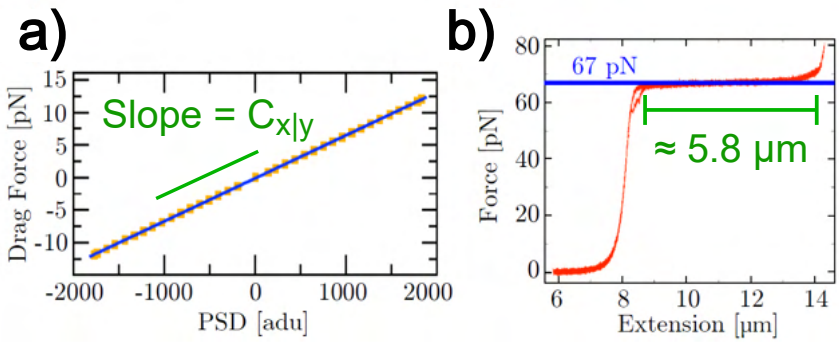


Figure 2.5: Force calibration methods. a) Stokes' law calibration. Drag force (obtained using Eq. 2.3) as a function of PSD values. Orange squares are experimental measurements, and the blue solid line is a linear fit, yielding the calibration factor. b) Overstretching transition of half of a λ-DNA molecule (24508 bp). The plateau occurs at $\sim$ 67 pN and has an extension of $\sim 5.8\ \mu m$. Figures are adapted from [91].

Finally, to calibrate the PSD-position detectors, we must determine the calibration factor between the voltage applied to the wiggler and the optical trap displacement. For this, a bead is sucked onto a micropipette, and the laser trap is centered on it such that the measured force is zero. The software is then used to keep the bead at constant zero fore, guaranteeing that the trap position will always be centered on the bead. By moving the chamber up and down (or right to left) with the motors, the piezos in the wiggler will then move the lasers to compensate for the motor's movement. The signal on the PSD-position detectors can be correlated with the movements of the distance moved by the motors.

Usually, the sequence for the calibration is the following: First, the Stokes law is used to calibrate the force (Fig. 2.5a). Second, the distance is calibrated. Finally, we perform an overstretch of several DNA molecules

to check the force and distance calibration. The plateau of the overstretch should be at $\sim 67\ pN$, and the extension of the plateau should be about $\sim 5.8\ \mu m$ (Fig. 2.5b). If the values are slightly off, we can correct the calibration factors so that the force and distance match these values.

2.2.5 Mini-tweezers pulling protocol

In the 'mini-tweezers,' the relative distance between the micropipette and the center of the optical trap can be controlled by moving the lasers. For most experiments, the control parameters are trap-pipette distance (λ) or the force (f). Controlling these parameters, different protocols can be used to perform experiments. For example, experiments with a constant distance between the micropipette and the trap, at a constant force, or where the distance or force is changed in a specific manner during the experiments.

The experiments carried out in this work are what we call pulling experiments, where λ is used as the control parameter. In this experiment, the optical trap is moved back and forth at a constant speed (v); therefore, $\lambda(t) = \lambda_0 + vt$. When λ increases (decreases), the force on the molecule increases (decreases) at a loading rate r. The force is therefore $f(t) = f_0 + rt$. r and v are related through the effective stiffness (k_{eff}) of the system $r = v \cdot k_{eff}$

This protocol will be used to carry out unzipping experiments of long DNA hairpins in Chapter 7.

Chapter 3

Nucleic Acids

Molecular biology encompasses a remarkable diversity of molecules that include among others, lipids, proteins, carbohydrates, nucleic acids (DNA and RNA), and adenosine triphosphate (ATP), the primary energy source of cells. The complexity of these molecules calls for the use of different experimental techniques to study them. Optical tweezers and nanopores are two of these techniques, and throughout this work, they will be used to study DNA. This chapter will introduce the basics of DNA, as it will be our primary sample throughout this work.

3.1 Nucleic acids

All organisms contain an "instruction manual," written and encoded in DNA. These instructions are genetic information stored inside the nucleus of cells in eucharistic organisms. The central dogma of molecular biology states that DNA is transcribed into RNA, which is then translated into proteins by the ribosomes. The proteins are then in charge of developing vital processes for the organisms. Furthermore, DNA is continuously replicated/copied when cells divide by mitosis.

Deoxyribonucleic acid (DNA) is a polymer composed of two polynucleotide chains that hybridize and coil around each other, forming a helix, as shown in Fig.3.1a. Although RNA is also a polynucleotide, very similar to DNA, it is usually found as a single strand. Nucleotides are the monomers forming nucleic acids. The nucleotides of the polynucleotide chains are composed of three different building blocks attached by covalent bonds. These building blocks are shown in Fig.3.1b and are:

1. **Sugar.** The sugar composing DNA is deoxyribose. Deoxyribose is a pentose sugar that is derived from ribose by the loss of an OH

group to an H group at the C2 position. While DNA is composed of deoxyribose, RNA is composed of ribose.

2. **Phosphate.** Together with deoxyribose, the phosphate groups make up the DNA backbone. This phosphate group is negatively charged in DNA physiological conditions. This makes DNA a highly negatively charged polymer. RNA has the same phosphate groups.

3. **Nitrogenous base.** Nucleobases or nitrogenous bases are nitrogen-containing compounds with a ring structure. DNA contains four different bases: Adenine (A), Cytosine (C), thymine(T), and Guanine (G). In the case of RNA, the thymine base is replaced by Uracil(U) Fig.3.1c.

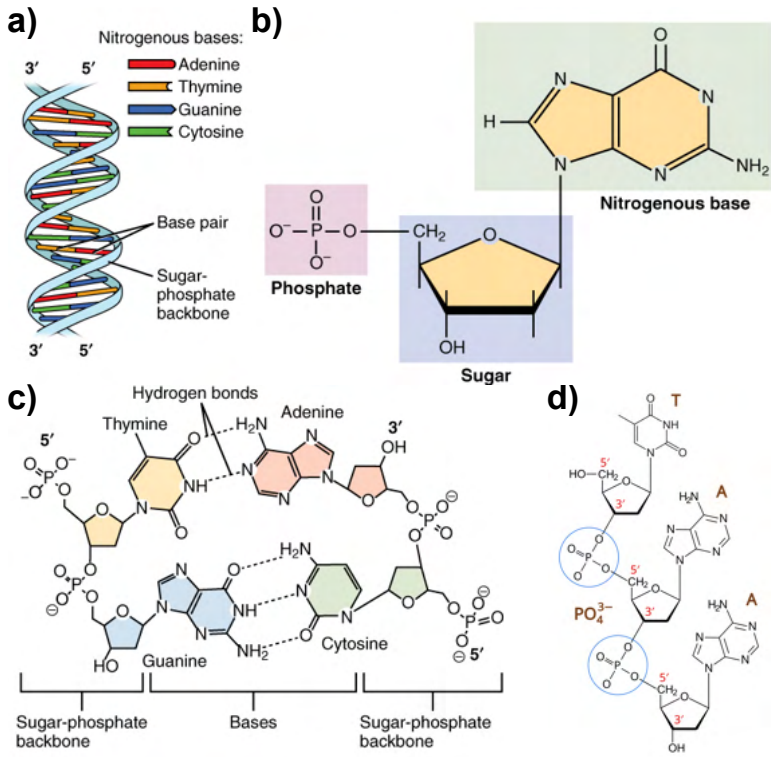


Figure 3.1: **DNA composition.** a) DNA helix structure shows the bases in different colors. b) Base-pairing between the different bases of DNA, showing the number of hydrogen bonds between them. A-T has two hydrogen bonds. G-C has three hydrogen bonds. c) Structure of a nucleotide composed of a phosphate group, a nucleobase, and a deoxyribose sugar.d) Schematic of three bond nucleotides shown the 5' and 3' ends.

DNA bases are divided into two different categories, Purines and Pyrimidines, depending on their size. The bases that contain one carbon-nitrogen ring and two nitrogen(N) atoms in the ring are called Pyrimidines, as seen for Thymine and Cytosine in Fig. 3.1b. Meanwhile, the bases that contain two carbon-nitrogen rings and four nitrogen atoms are called Purines, as seen for Adenine and Guanine in Fig. 3.1b.

The nucleotides are connected to each other via phosphodiester bonds that form between the phosphate group at the 5' end of one nucleotide to the 3' sugar carbon of the next nucleotide (Fig.3.1d).

The two polynucleotide strands run antiparallel to each other, such that while one goes from 5' to 3', the other has the opposite directionality. The strands interact with each other via hydrogen bonds that form between the different nucleobases. As seen in Fig.3.1c, the T of one of the strands forms two hydrogen bonds with the A from the other, while the C forms three hydrogen bonds with the G. These are the two types of base pares that are found in DNA, A-T, and G-C, and are called Watson-Crick base-pairing. The higher number of hydrogen bonds of the G-C base pairing grants higher stability. In RNA, A forms two hydrogen bonds with U (A-U).

3.1.1 DNA structure

Fig.3.2 shows the different DNA structures. DNA is usually found in the B-DNA structure, where the two DNA strands form a right-handed helix, held together by the hydrogen bonds formed between complementary nucleotides (Watson-Crick base-pairing). A representation of B-DNA is shown in Fig.3.2 (middle). Several factors contribute to B-DNA being the most stable structure for a wide range of conditions, including physiological conditions. Being the most relevant factors:

- The repulsion between the negatively charged phosphate groups of the DNA backbone. This repulsion is mediated by the interaction of the phosphates with ions or water molecules in the DNA environment, which can partially screen for DNA's negative charge.

- The hydrophobic interactions between adjacent bases lead to the so-called base staking interaction.

The more common B-DNA structure is a right-handed helix that requires 10.5 base pairs to complete a turn, has a rise of 3.3 Å per base pair, and a diameter of 2.0 Å. All the parameters of the different structures are shown in see Table 3.1 [92].

A-DNA is usually found in low-hydrated conditions, such as solutions containing ethanol. A-DNA has a right-handed helix structure, with 11 base

pairs per complete turn (Fig.3.2 left). Furthermore, the rise per basis pair is 2.3 Å, and its diameter is 2.3 Å (Table 3.1).

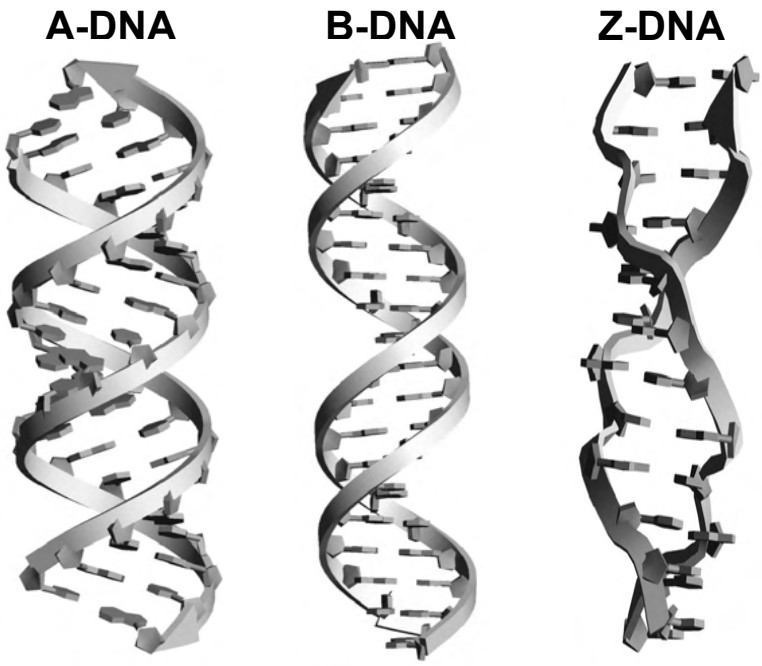


Figure 3.2: **3D-geometric structure of A-DNA, B-DNA, and Z-DNA.** Left: A-DNA structure. Center: B-DNA structure. Right: Z-DNA structure. Images from [93].

Z-DNA, surprisingly, is a left-handed helix in which the phosphate groups of the back bond form a zigzag pattern, hence the name of Z-DNA (see Fig.3.2 right). Z-DNA has 12 bases per turn, with a rise per base of 0.38 Å and has the smallest diameter of the three structures, of 1.8 Å(Table 3.1). Z-DNA is a higher-order structure known to form in alternating purine-pyrimidine or pyrimidine-purine dinucleotide repeat sequences, with the propensity of Z-DNA to form in the order $GC > CA > TA$ [94, 95]. Z-DNA has been found to form at very high NaCl concentration $\sim$ 4 M [96]. The high ionic concentrations reduce the interactions between interstrand phosphate groups, allowing them to come closer to each other. These phosphates are much closer in Z-DNA, about 8 Å, compared to B-DNA, where they are about 12 Å across the minor groove [93]. Moreover, Z-DNA can coexist with B-DNA in the same DNA molecule due to the formation of a B-Z junction, where DNA changes from the B to the Z structure. This junction requires a strong twist in DNA's backbone, as DNA needs to change from left to right-handed helix sense. Strikingly, this is achieved by not forming one base pair at the junction, and the bases are flipped out into the exterior of

the helix [97, 98]. Even low $\sim 1-2\%$ levels of Z-forming DNA sequence can significantly change the supercoiling properties of DNA.

Table 3.1: **Structural features of the three major forms of DNA.**. 3D-structures of A-DNA, B-DNA and Z-DNA [92].

Geometric attribute	A-DNA	B-DNA	Z-DNA
Helix sense	right-handed	right-handed	left-handed
bp/turn	11	10.5	12
Rise/bp along axis (Å)	2.3	3.3	3.8
Diameter (Å)	2.3	2.0	1.8

3.1.2 Elastic models to describe DNA

Different models have been developed to describe the behavior of polymer chains such as DNA. The simpler descriptions assume an ideal chain where the different monomers do not interact with each other. More complex descriptions included interactions between the monomers and excluded volume effects. These models are called real chains. Here, we present two ideal chain models to describe the elastic properties of polymers, known as the Freely-Jointed Chain (FJC) and the Worm-Like Chain (WLC). More information about further models can be found in [99].

Freely-Jointed chain

The simplest model to describe polymers is the FJC. In the model, the polymer consists of N rigid segments of length c that can orient freely in space, and that connect the polymer monomers. In some cases, the segments can have a physical interpretation, such as the inter-phosphate distance for DNA. The model is equivalent to a random walk with a step size c, neglecting any interactions between the monomers. The position of the monomer i, is $\overrightarrow{r}_i$ (see Fig.3.3a). The vector connecting two consecutive monomers is $\overrightarrow{t}_i$, where $\overrightarrow{t}_i = \overrightarrow{r}_{i+1} - \overrightarrow{r}_i$. The end-to-end vector of the polymer can be written as the sum of $\overrightarrow{t}_i$:

$$\overrightarrow{R} = \sum_{i=1}^{N} \overrightarrow{t}_i = 0 \tag{3.1}$$

The average end-to-end vector is zero in analogy with the mean displacement of the random walk. The vectors connecting two monomers are uncorrelated $\left\langle \overrightarrow{t}_i \cdot \overrightarrow{t}_j \right\rangle = \delta_{ij} c^2$ and therefore, like for the random walk, the

mean square displacement is $\left\langle \overrightarrow{R}^2 \right\rangle = Nc^2$. Here, $\langle ... \rangle$ represents the average over different conformations. However, if the polymer is semi-rigid or rigid, it will not be flexible over distances larger than the length of bonds between monomers (c) due to interactions between them. In this case, if the chain length is longer than the interactions between the monomers, any chain can be rescaled to the FJC. The recalling is done by defining a new segment length, named the Kuhn length (b), such that neighboring segments are uncorrelated. In this case, the average mean square distance becomes $\left\langle \overrightarrow{R}^2 \right\rangle = Nb^2$ (see Fig.3.3b).

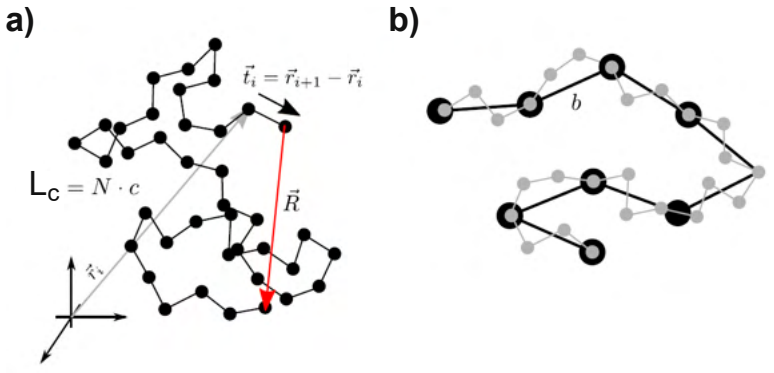


Figure 3.3: **Schematic depiction of the Freely-Jointed Chain model.** a) Representation of a Freely Jointed chain with monomer length c. The end-to-end vector, $\overrightarrow{R}$, is also shown. b) Schematic of the rescaling of the polymer, dividing it into longer segments, each of Kuhn length, b.

Worm-Like Chain

The WLC model considers the polymer as a thin, continuous, and flexible rod that is inextensible, in contrast to other models where the polymer is divided into monomers. It was proposed in 1949 to describe the elasticity of semi-flexible polymers [100]. Moreover, this model has been able to reproduce the force-distance curves that are obtained by pulling dsDNA using optical or magnetic tweezers [101]. In the WLC, the polymer is described by two parameters: its full stretched length, the contour length (L_c), and the persistence length (P), which models its flexibility. In Fig.3.4a, a representation of a WLC is shown. In the model, the polymer path is parametrized by $s \in (0, L_c)$, such that $\overrightarrow{r}(s)$ is the position vector along the chain, and $\hat{t}(s)$ is the unit tangent vector at position s. The tangent vector is then $\hat{t}(s) = \frac{\partial \overrightarrow{r}(s)}{\partial s}$, as shown in Fig.3.4b, and the end-to-end distance of the polymer is given by $\overrightarrow{R} = \int_0^{L_c} \hat{t}(s) ds$, red vector in Fig.3.4a.

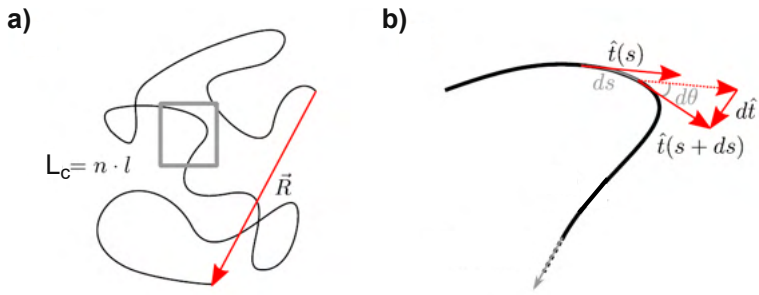


Figure 3.4: **Schematic depiction of the Worm-Like Chain model.** a) Representation of a Worm-Like Chain (WLC) rod with a contour length $L_c = n \cdot l$. The area inside the gray square is zoomed in on the right. (b) The tangent vector $\hat{t}(s)$ at position s.

Whereas the FJC model is purely entropic, the WLC adds an enthalpic contribution that takes into account a bending penalty of the polymer. The energy associated with bending is given by:

$$E = \frac{1}{2} k_B T \int_0^{L_c} P \cdot \left(\frac{\partial \overrightarrow{r}(s)}{\partial s} \right)^2 ds \tag{3.2}$$

where P is the polymer's characteristic persistence length, k_B is the Boltzmann constant, and T is the absolute temperature. The energy cost associated with bending has as a consequence that the correlation function between tangent vectors is not a delta function like for the FJC but has an exponential decay, with a decay contact of 1/P:

$$\left\langle \hat{t}(s) \cdot \hat{t}(0) \right\rangle = cos(\theta(s)) = e^{-s/P} \tag{3.3}$$

where θ is the bending angle at the position s in the rod, as shown schematically in Fig.3.4b. The force required to maintain the polymer extended at an extension (x), against entropic forces is given by [102]:

$$f_{WLC}(x) = \frac{k_B T}{P} \left(\frac{1}{4} \left(1 - \frac{x}{L_c} \right)^{-2} - \frac{1}{4} + \frac{x}{L_c} \right) \tag{3.4}$$

where L_c is the contour length of the molecule. For the case of DNA, $L_c = nb$, with n the number of bases of the DNA sequence and b the base-to-base distance.

Finally, a useful quantity is the mean square end-to-end distance of the polymer. That for the WLC is given by:

$$\left\langle \overrightarrow{R}^2 \right\rangle = 2PL_c \left[1 - \frac{P}{L_c} \left(1 - e^{-L_c/P} \right) \right] \tag{3.5}$$

where it can be observed that for the limit $L_c \gg P$, then $\langle R^2 \rangle = 2PL_c$, showing that when the Kuhn length of the FJC is twice the persistence length, the FJC resembles the WLC model.

3.2 Helicases

Helicases are essential enzymes that unwind double-stranded nucleic acids, such as DNA or RNA. These enzymes are critical for many fundamental cellular processes, including DNA replication, repair, recombination, transcription, and RNA processing. Helicases achieve strand separation by translocating along the nucleic acid, disrupting the hydrogen bonds between complementary bases. This separation involves breaking the hydrogen bonds between the complementary nucleotides. Therefore, helicases require energy for this process, which they get from the hydrolysis of ATP. This translocation can occur in either the 5' to 3' or 3' to 5' direction, depending on the helicase type [103, 104].

Similarly, translocases are molecular motors that move along nucleic acids or other polymers, often without directly separating strands but facilitating processes like protein displacement or substrate transport. Both helicases and translocases are critical for maintaining genome stability and regulating gene expression [104].

There are many different types of helicase, depending on the organism or the process in which DNA is separated (DNA replication, transcriptions,...). In Chapter 8, we will show how SPRNT (Single-molecule picometer resolution nanopore tweezers), a recently developed technique to study enzymes with unprecedented precision, can be used to measure the movement along the DNA of two different helicases/translocases. The first is Hel308, an ATP-dependent translocase that translocates on ssDNA in the 3' to 5' direction. Hel308 helicases are conserved among archaea and eukaryotes, including humans. In contrast to other enzymes that move along ssDNA, which perform steps on DNA of 1 bp at a time, Hel308 was found to have substeps of less than 1 bp when moving along DNA [20, 105]. The second is gp41, which is a bacteriophage T4 gp41 replicative helicase. gp41 is a hexameric helicase, conformed by six equal monomers [106] that unwinds DNA with 5' to 3' polarity. In Chapter 8, SPRINT experiments with both of these helicases are presented.

Part II

ELECTRICAL MEASUREMENTS WITH NANOPIPETTES

Chapter 4

Nanopipette conductance and noise

Understanding the different parameters that contribute to the conductance of nanopipettes is essential when performing electrical measurements. In this Chapter, conductance measurements of nanopipettes over a wide range of concentrations explore the rectification effect at low concentrations. We compare the linear and non-linear conductivity models for salt conductivity. Finally, we explore the dominant, flicker noise in nanopipettes and how they depend on the applied voltage and concentration.

4.1 Nanopipette geometry

The electrical resistance of nanopipettes is mainly determined by their geometry and dimensions at the nanopipette tip region, as this region has the biggest contribution to the nanopipette's resistance. An electron microscope must be used to visualize the nanopipette tip, as its dimensions are far below the diffraction limit. In Fig.4.1 a) and b), images obtained using a scanning electron microscope (SEM) can be observed. From these images, one can observe the two drawbacks of using an SEM microscope. Nanopipettes with tips in the 20-60 nm range are very close to the resolution limit of this type of microscope. Furthermore, in these images, the internal geometry of the nanopipettes cannot be observed, so it is impossible to determine the exact diameter at the tip, as the thickness of the nanopipette walls is not known. The cone angle can also not be determined reliably from these images. To overcome this problem, it is possible to use a transmission (TEM) electron microscope. TEM microscopes have a higher resolution than SEM, and because they detect transmitted electrons, they can reveal the internal features of thin samples. In Fig.4.1 c)-f) images taken with a

TEM microscope can be seen. From these images, the tip diameter (d) is determined (blue in Fig.4.1), and the cone angle (2α), where α is the cone angle (red in Fig.4.1). These estimations using the TEM are much more precise than for the SEM, as the internal section of the nanopipette can be observed in the images.

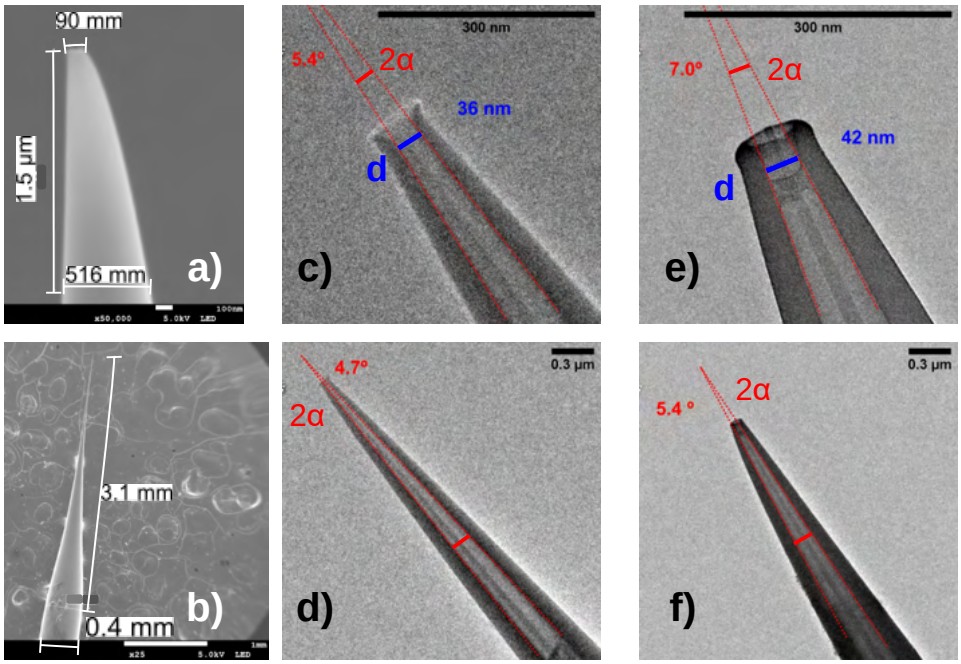


Figure 4.1: a) and b) SEM images of two nanopipettes at different scales. c) and d) TEM images of the same nanopipette at different scales. e) and f) TEM images of the same nanopipette at different scales.

The images in Fig.4.1 c) and d) (and e) and f)) are from the same nanopipette but at a different scale. It can be seen that the cone angle of the interior section of the nanopipettes remains similar even for distances $1 - 2\ \mu m$, away from the tip. The electrical resistance of a conical nanopipette with an angle of about 6^o is mainly concentrated in the first hundreds of nanometers of the nanopipette. Therefore, a conical geometry can be assumed for our nanopipettes, and there is no need to assume more complex geometries like a double cone geometry [23].

We can determine their mean diameter and cone angle by measuring several nanopipettes with the TEM microscope. The mean diameter is $d = 34 \pm 16$ nm, and the mean angle is $2\alpha = 5.8° \pm 0.5°$, and therefore $\alpha = 2.9° \pm 0.3°$ is the cone angle. There is some variability in tip diameters for a given protocol of the pipette puller, and the diameters are not so reproducible. More information about the pipette pulled protocols and how

to reduce the variability of the tip diameter can be found in App.A). The cone angles do not depend on the nanopipette tip diameter.

4.2 Conductance measurements of Nanopipettes

A good conductance model for nanopipettes is important to determine their diameter from conductance measurements in a broad range of concentrations. We performed experiments varying the salt concentration. Current measurements are performed at several positive and negative voltages for each concentration.

Experimental Protocol

Experiments are performed with LiCl, NaCl, KCl and $MgCl_2$ salts. Apart from the salt, the buffer contains 10 mM Tris. Moreover, the LiCl and KCl buffers are at pH = 7.5, while the NaCl and $MgCl_2$ buffers are at $pH = 9.0$. The salt concentration range from the experiments is from 0 mM (just the 10 mM Tris buffer) to $3.5 - 6$ M, depending on the salt type, as some of the salts saturate at concentrations lower than 6 M. The experimental procedure is as follows: a microfluidic chip containing six nanopipettes is filled with the buffer at the lowest salt concentration (more details about the preparation of the microfluidic chip containing the nanopipettes are explained in more detail in Sec.5.1 and App.A.) For each concentration, the electric current flowing through the nanopipette is measured at different voltages, ranging from $0 - 1000mV$. After the voltage is changed, we wait for the electric current to stabilize before making the current measurements, as sometimes the current drifts for a few seconds before reaching a stable value, and then the current through the nanopipette is recorded for about ~ 10 seconds. After measuring all nanopipettes, the buffer with the next concentrations is flown into the microfluidic chip. At least five times the volume of each microfluidic chip chamber is flowed to ensure a good buffer exchange. The chip is left overnight in the fridge, covered with a plastic foil to avoid evaporation. This is done to allow for good diffusion of the new buffer into the nanopipettes. The next day, a fresh buffer is flowed into the microfluidic chip again before performing the measurements to ensure the desired concentration if there was some evaporation overnight. This procedure is repeated for several days until all concentrations have been measured.

Current vs. Voltage I-V plots

The I-V plots (current vs. voltage) for a nanopipette filled with NaCl are shown in Fig.4.2, with b) being a zoom-in of the y-axes of a). In Fig.4.2a,

it can be observed that for the lower concentrations, the currents are larger (in absolute value) for negative than for positive voltages. Therefore, the nanopipette conductances (the slope of the I-V plots) are higher for negative voltages at low salt concentrations. This is due to the rectification effect explained in Section 1.4.2. Due to the asymmetric geometry of the nanopipettes and their negatively charged surface, charge polarization inside the nanopipette causes the salt concentration inside the pipettes to be higher for negative voltages, making the pipettes more conductive. This effect decreases when increasing the salt concentration, as shown in Fig.4.2 b). As concentration increases, the Debye layer decreases, reducing the rectification effect associated with the negative charge from the nanopipette surface.

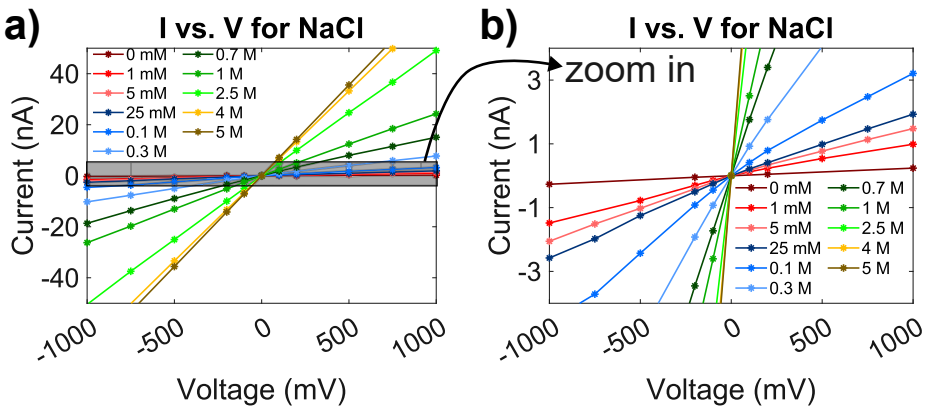


Figure 4.2: **I-V plot.** I-V plot for different concentrations of a nanopipette with a diameter of 22 nm in NaCl. a) Plot from -40 nA to 40 nA. b) Zoom-in of the y-axis in a) and shows the current range from -4 nA to 4 nA

Rectification in Nanopipettes

From the I-V measurements in the previous section, we can calculate the rectification ratio ($|I_N(V_-)|/|I_P(V_+)|$), which is given by the ratio of the electric currents at the same value of negative and positive voltages. The rectification ratio at different salt concentrations is shown for a monovalent salt (NaCl) and a divalent salt ($MgCl_2$), in Fig.4.3a and b, respectively. Nanopipettes of different tip diameters are shown in each case, and the mean value of all of them is shown in black.

It can be observed that the rectification ratio decreases with the salt concentration. As salt concentration increases, the Debye length (ℓ_B) and the Dunking length (ℓ_{Du}) decrease, and therefore, effects associated with the negatively charged surface of the nanopipette decrease. Comparing Fig.4.3

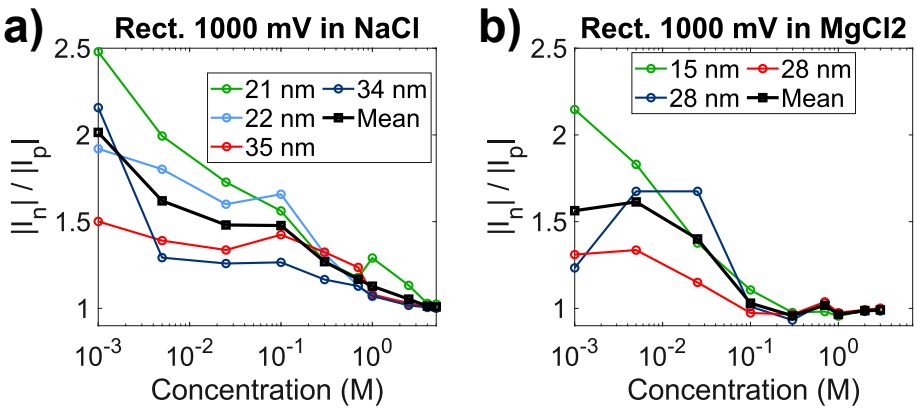


Figure 4.3: **Rectification in nanopipettes.** The rectification ratio ($|I_N(V_-)|/|I_P(V_+)|$) as a function of concentration is shown for different nanopipettes in NaCl a) and $MgCl_2$ b). The mean value of the ratio over the different nanopipettes is shown in black.

a) and b), it can be observed that the rectification ratio is reduced much faster for the divalent salt condition. At 100 mM the rectification ratio is almost 1 for $MgCl_2$, while it is still ~ 1.5 for NaCl. For KCl and LiCl, a similar behavior to NaCl is observed, where the rectification ratio at 100 mM is still over ~ 1.5. This agrees with the stronger effect that divalent ions have on the screening of surface charge, reducing related effects such as rectification.

4.3 Modelling Nanopipette Conductance

The nanopipette conductance ($G = 1/R = I/V$) at each salt concentration can be calculated by performing a linear fit to corresponding points in the I-V plot. The nanopipette's conductance is measured independently for positive and negative voltages for each salt concentration. The conductance of a nanopipette is modeled by Eq.1.13, presented in Chapter 1. The model requires the conductivity of the salt solutions $g(c)$ and the surface charge as a function of the concentration $\sigma(c)$.

To model the conductivity of the electrolyte solutions, the linear and non-linear models presented in Fig.1.10 in Chapter 1 are used. These two models are used for each of the four different salts to test which one of them works better in each case.

4.3.1 Modelling surface charge

The charge of the inner surface of the quartz nanopipettes, which contributes to the surface conductance, is modeled as explained in Sec. 1.5.2. However, the Gouy–Chapman model, presented in Sec. 1.5.2, works well for surfaces with small surface charge or a low density of dissociable sites at the surface. For quartz surfaces, it has been proposed that the surface density of dissociable sites, the number of dissociable sites per surface area, is $\Sigma \approx 8\ nm^{-2}$ [107]. For a completely dissociated surface, this would give a surface charge of $\sigma_0 = 1.28\ C/m^2$, which is a much larger surface charge than for the case presented in Sect. 1.5.2.

For high surface charges, the model is modified to include the existence of the Stern layer. The model is called the Stern model and accounts for the fact that when a charged surface is placed into a fluid containing ions, a layer of counterions builds up at the surface; that is, the Stern layer. This layer consists of specifically absorbed ions, especially counterions (cations for the case of the negatively charged quartz surface), that are absorbed to the surface in response to chemical interactions (see Fig.4.4). Further away from the Stern layer, a second layer of ions interacting via Coulomb forces with the surface electrically screens the first layer. The second layer is "loosely" associated with the surface so that the ions in this layer are free to move under the influence of electrical attraction or thermal motion. It is thus called the "diffuse layer" or Debye layer (see Fig.4.4). The Gouy–Chapman model presented in Chapter 1 only considers the diffusive layer.

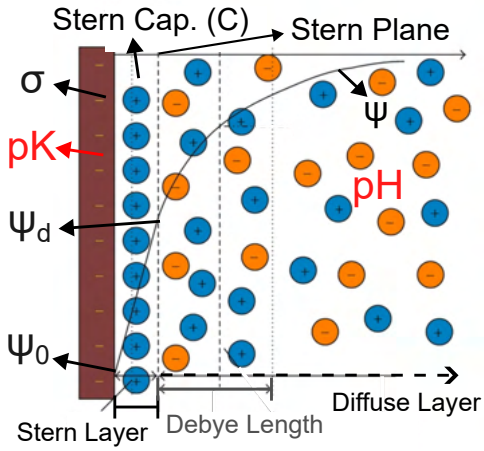


Figure 4.4: **Stern Model.** Schematic representation of the cation and anion distributions near a negatively charged surface in the Stern model. The potential (Ψ) as a function of the distance to the surface is also shown.

The basic Stern layer model considers the Stern layer as a region where the potential drops linearly between ψ_0 (the surface potential) and ψ_d (the diffuse layer potential) as shown in Fig.4.4. Further away from the surface, in the diffuse layer, the potential decays exponentially, as predicted by the Gouy–Chapman model. The potential drop in the Stern layer is characterized by the Stern layer's phenomenological capacity (C), defined by the equation:

$$C = \frac{\sigma}{\psi_0 - \psi_d} \tag{4.1}$$

where σ is the surface charge. Note that the phenomenological capacity C has units of farad per surface area. As the ions in the Stern layer are very strongly bound to the surface and will not move when an electric field is applied, the potential and the surface charge at the diffuse layer need to be considered. Using this definition for the Stern layer and Eq.1.17 for the potential at the surface, the potential at the diffuse layer can be expressed by the equation:

$$\psi_d = \frac{k_B T}{e} \ln\left(\frac{\sigma}{\sigma_0 - \sigma}\right) + \frac{k_B T \ln 10}{e} (pK - pH) - \frac{\sigma}{C} \tag{4.2}$$

with σ_0 the surface charge when it is fully dissociated, given $\sigma_0 = -e\Sigma$, where Σ is the surface density of chargeable sites and e the elementary charge. We have used the following reported values for solutions containing KCl, $\Sigma = 8\ nm^{-2}$, pK = 7.9 and $C = 0.3\ F/m^2$ [108, 107] to estimate the surface charge σ. It is important to note that in this chapter, we will assume the same phenomenological capacitance for all the salts as it was the only value we could find in the literature. However, it is known that the structure of the Stern layer is dependent on the cations that conform to it. Therefore, this value could be different for the other cations, especially for divalent cations.

Using Eq.4.2 and the Grahame equation (Eq.1.18) that relates the potential and the surface charge, the surface charge of the inner quartz surface of the nanopipettes can be calculated at any salt concentration for a given pH value. This surface charge will be used in the conductance model for nanopipettes presented in the next section.

4.3.2 Nanopipette Conductance at different concentrations

The conductance for different concentrations for four different nanopipettes in different salts is shown in Fig.4.5. Conductances for positive voltages are shown with an asterisk, while negative voltages are represented with circles. The rectification mentioned in the last section can be observed at low salt

concentrations, where conductances for negative voltages are systematically larger than for positive voltages.

The positive and negative conductances of each nanopipette are simultaneously fitted. The fit is performed on the logarithmic scale for both axes so that larger conductance values don't have a bigger weight in the fit. The fits are performed using the linear (Fig. 4.5, dashed lines) and non-linear (Fig. 4.5, continuous lines) conductivity models presented in Sec.1.5.3 and the surface charge model presented in the previous section. To perform the fit, we use the following expression:

$$G = \frac{2\pi \cdot g(c) \cdot d \tan(\alpha)}{4 + \pi \tan(\alpha)} + \frac{\pi\kappa \cdot |\sigma(c)| \cdot \mu_+ D^*}{L^*} \tag{4.3}$$

This is the same expression as Eq. 1.13, but adding a parameter κ for the surface conductance term. The values $L^* = 150$ nm and $D^* = 45$ nm have been used for the length and diameter of the conical nanopipette tip dimensions for the surface conductance term, as explained Sec.1.5. This dimension corresponds to the sensing region at the tip of the nanopipette. Similar values have been used in previous works [29]. The values of L^* and D^* have been estimated from the TEM images (Fig.4.1). d and κ are the free parameters of the fit. The parameter κ accounts for surface charge variations that can occur in the nanopipette tip region and inaccurate estimation of the surface conductance in conical geometries. The d and κ parameters obtained for the linear and non-linear models are shown in Tab.4.1.

From the fits in Fig.4.5, it can be observed that, except for $MgCl_2$, the non-linear conductivity models give a better fit of the conductance data. Especially at the higher salt concentrations, where it reproduces the non-linear behavior of the conductance data. On the other side, the surface conductance term reproduces the behavior at low salt concentrations, where the surface charge determines the conductance.

The diameters estimated with the fits are lower for the linear model, as the linear model overestimates the conductivity; hence, we obtain smaller diameters from the fits to compensate for the conductivity overestimation. The values obtained for κ_{NL} and κ_L also indicate that surface charge conductivities are slightly overestimated in our model. This could be due to either an overestimation of the surface charge $\sigma(c)$ or due to an overestimation of the surface conductance of conical geometries.

4.3.3 Surface vs. bulk conductance

Conductance curves for all salts are plotted in Fig4.6a, showing the effect of pH at low slat concentrations. NaCl and $MgCl_2$, measured at pH = 9.0, show a higher surface conductance as the quartz has a higher nega-

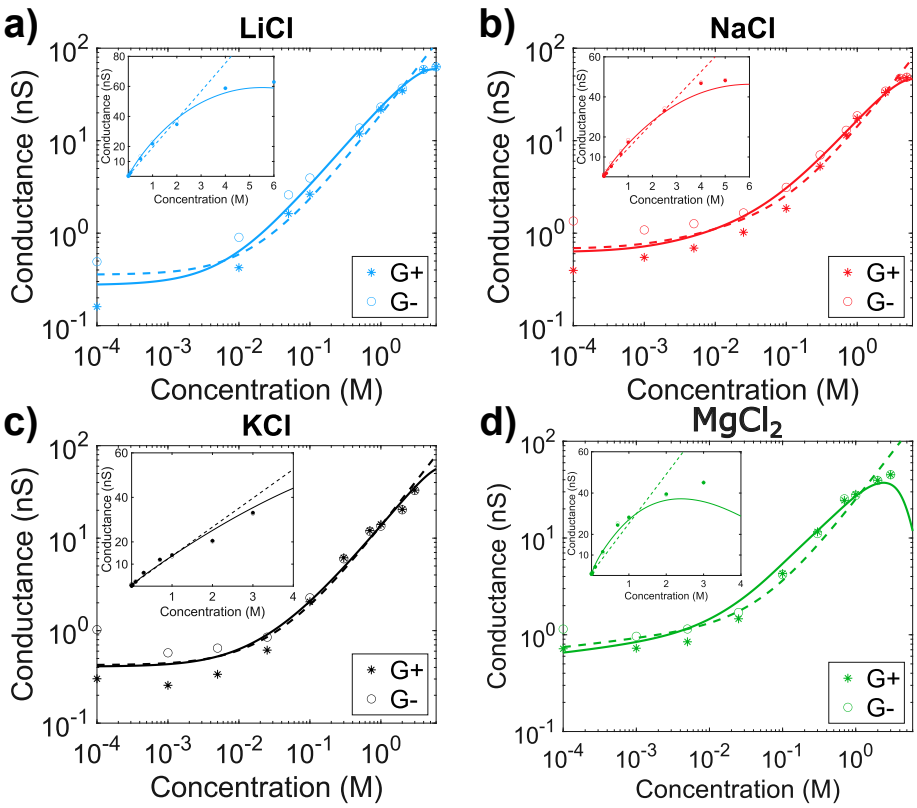


Figure 4.5: **Nanopipette conductance with concentration.** Conductances of nanopipettes at positive (G_+) and negative voltages (G_-) in LiCl, NaCl, KCl, and $MgCl_2$. Continuous lines are fit with the nonlinear (g_{NL}) and the linear (g_L) conductivity models. The conductance at 0 mM, where the buffer is only 10 mM Tris, is plotted at a concentration value of 10^{-4} on the x-axis.

tive charge for higher pH. On the contrary, LiCl and KCl were measured in buffers at pH = 7.5, where the higher concentrations of protons in the buffer neutralize the chargeable sites of quartz, reducing its surface charge. Therefore, the surface charge and surface charge conductance are lower. In Fig4.6b, the conductance of the nanopipettes in NaCl and LiCl (pH 9.0 and 7.5, respectively) has been decomposed into surface and bulk conductance. It can be observed that while for the lower pH, the bulk conductance overtakes the surface at a salt concentration of about 4 mM, for the higher pH, this does not happen until almost 33 mM. This is due to the larger Dukhin length (Eq.1.8) for the higher pH value. The Dukhin length is proportional to the surface charge; therefore, it is larger for the higher pH.

Table 4.1: **Conductance global fit parameters.** d_{NL} (in nm) and κ_{NL} (unit less) parameters obtained using the non-linear conductivity model. d_L and κ_L obtained using the linear conductivity model

	d_{NL} (nm)	κ_{NL}	d_L (nm)	κ_L
$LiCl$	39 ± 7	0.17 ± 0.09	20 ± 7	0.4 ± 0.2
$NaCl$	22 ± 6	0.41 ± 0.12	13 ± 4	0.61 ± 0.18
KCl	15 ± 5	0.31 ± 0.10	11 ± 4	0.38 ± 0.13
$MgCl_2$	28 ± 5	0.27 ± 0.07	15 ± 4	0.45 ± 0.13

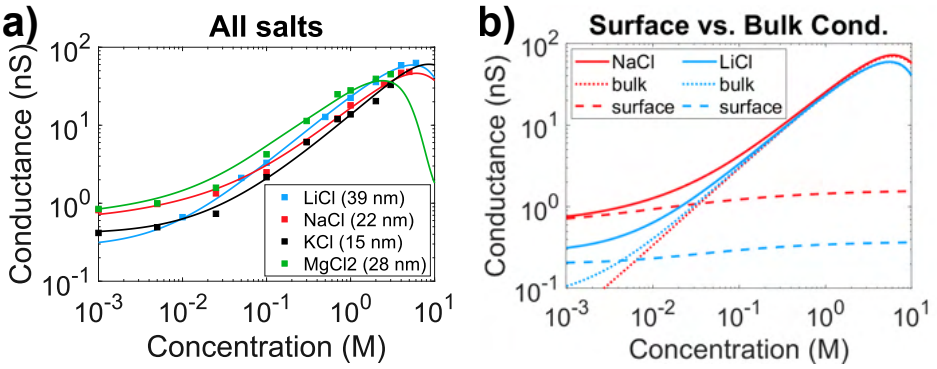


Figure 4.6: **Surface Conductance.** a) Nanopipette conductance for LiCl, NaCl, KCl, and $MgCl_2$. The points are the mean value of the conductance at positive and negative voltages. The continuous lines are the fits using the nonlinear conductivity models. b) Decomposition of the nanopipette conductance into the bulk and surface conductance terms for NaCl and LiCl

4.3.4 Diameter estimation at different salt concentrations

When performing experiments, we usually perform conductance measurements at one concentration and not over a broad concentration range. Therefore, it is important to be able to estimate the nanopipette tip diameter from conductance measurements at one concentration. In this section, we use the conductance model presented in the previous sections to predict the diameter for different salt concentrations to see if we obtain a similar diameter value independent of the concentration. We use Eq. 1.13 to estimate the diameters, with the surface charge model presented in Sec. 4.3.1. The diameters are calculated in four ways to explore which one better reproduces the diameter from the global fit to all concentrations. First, we estimate the diameter from the conductances at positive (V+) and negative voltages (V-) and using the non-linear conductivity model (g_{NL}) (orange and pink squares in Fig.4.7). Second, we calculate the diameters with the mean conductance between the positive and negative voltages, using the

non-linear model g_{NL} (red circles in Fig.4.7) and the linear model g_L (blue circles in Fig.4.7). The dashed black line represents the diameter estimated from the global fit using the non-linear conductivity model.

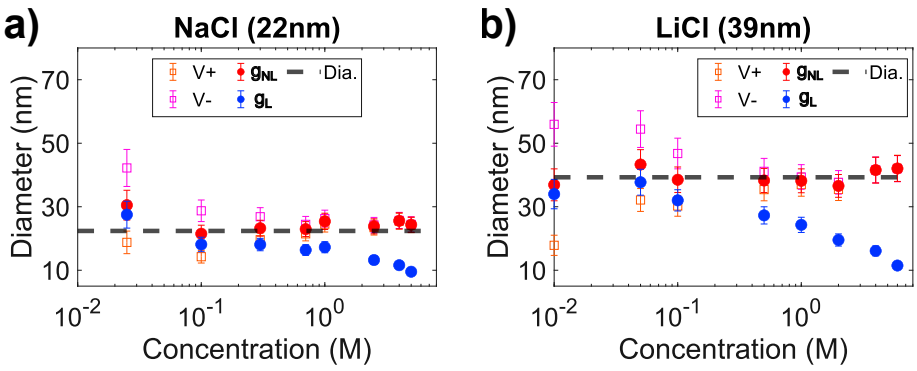


Figure 4.7: **Nanopipette tip diameter estimation.** Estimation of the tip diameter at different concentrations in different ways. 1) Pink squares (V_+) are diameter estimations from conductance at positive voltage using the nonlinear conductivity model. 2) Orange squares (V_-) from the conductance at negative voltages. 3) Red circles (g_{NL}) are diameters estimated with the mean conductance between positive and negative voltages and the non-linear conductance model. 4) Blue circles (g_L) are diameters estimated with the mean conductance between positive and negative voltages and the linear conductance model. a) Nanopipette with NaCl. b) Nanopipette with LiCl

From both panels in Fig.4.7, it can be observed that at low concentrations, when estimating the diameter just from conductance measurements at positive (V_+) or negative (V_-) voltages, the diameter is underestimated at positive voltages and overestimated at negative voltages. This is due to the rectification effect (explained in Sec.1.4.2), which causes concentration inside the nanopipette to be higher than in bulk at negative voltages (Fig.4.7, square pink symbols at low concentrations), and the diameter is therefore overestimated. At positive voltages, where the concentration is lower than in bulk, the diameter is underestimated (Fig.4.7, square orange symbols at low concentrations). However, when we take the mean value of both conductances, the estimated value for the diameter is in agreement with the diameter from the global fit. At higher concentrations, the conductances at positive and negative voltages converge as rectification diminishes. Therefore, the same value for the diameter is obtained from V+, V-, or their mean (red circles in Fig.4.7). However, it can be observed that the diameter estimates of the non-linear model (g_{NL}) converge to the diameter estimated by the global fit. In contrast, the values of the linear model (g_L) continually decrease, indicating that the linear model overestimated the conductivity at high concentrations, and hence the diameter is

underestimated. The nonlinear conductivity model does not consider the reduction in ions' mobility when concentration is increased, so it overestimates the conductivity. For this reason, we can conclude that the nonlinear conductivity (g_{NL}) model better estimates the nanopipette diameter than the linear conductivity model. Conductance measurements at just one voltage polarity are enough at high concentrations to determine the diameter. However, at low concentrations, when rectification is significant, it is better to use the mean diameter from positive and negative voltage polarities.

4.4 Flicker noise in nanopipettes

This section studies the electric current power spectral density (PSD) of nanopipettes. Specifically, the flicker noise (Eq.1.1) is investigated as it is the primary noise source in our nanopipette, as will be seen in this section.

The PSD is computed for each voltage and salt concentration from the electric current measurements presented in this chapter. The PSD is computed by performing a Fourier transform of the electric current signal. Then, a mean box filter of the PSDs is performed using a box width that exponentially increases with frequency. This permits us to consistently compare the different PSDs over several decades in frequency. The electric current measurements were acquired at 250 kHz using a 30 kHz low-pass filter before digitalizing the data. According to the Nyquist–Shannon sampling theorem, the PSD frequency range is up to 125 kHz, as shown in Fig.4.8. The characteristic drop of low pass filters can also be observed over 30 kHz in the computed PSDs.

Flicker noise for different concentrations

We start by comparing the PSD of current traces acquired at the same voltage but at different concentrations. Fig.4.8 shows the PSD of current measurements performed at 1000 mV at different NaCl concentrations. As concentration is increased, it can be observed that flicker noise increases and becomes the predominant noise source in the range below 20 kHz. However, the frequency range in which the flicker noise is dominant decreases for the lower concentrations. Therefore, a region where white noise (frequency-independent shot and thermal noise) is dominant can be observed from 3 kHz to 30 kHz below 100 mM NaCl.

Flicker noise for different concentrations

The dependence of flicker noise with the applied voltage can be observed in Fig.4.9a, where the PSDs for several voltages applied to a nanopipette

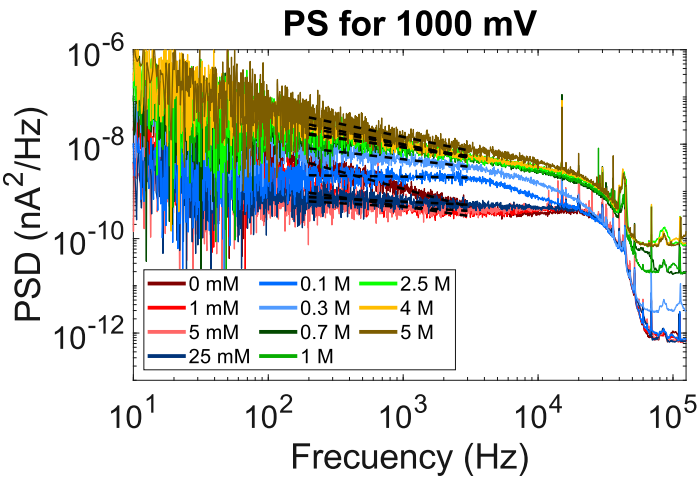


Figure 4.8: **PSD at different salt concentrations.** The PSDs correspond to electrical current measurements at 1000 mV in a nanopipette with a 22 nm diameter in NaCl. The concentration ranges from 0 mM (just 10 mM Tris at pH = 9.0) to 5 M. Data are from a nanopipette with a 22 nm tip diameter in NaCl.

containing 25 mM NaCl are observed. It can be observed that the flicker noise is dependent on the voltage magnitude but also on its polarity. For negative and positive polarities, the flicker noise increases as voltage increases. Fig.4.9b shows the magnitude of electric current vs. the magnitude of voltage for the electric measurements of Fig.4.9a. As rectification is still significant at 25 mM, the currents of negative voltages are higher. Surprisingly, it can be observed that although the absolute value of current at 0.75 and 1 V is higher than at -0.5 V (Fig.4.9b), the flicker noise is much higher for the negative voltage polarities (as seen in Fig.4.9a).

Noise parameters

As presented in Sec.1.3, Flicker noise is described by the equation (Eq. 1.1):

$$S_{I,1/f} = \frac{\alpha_H I^2}{f^\beta N_C} \Rightarrow \log(S_{I,1/f}) = -\beta \log(f) + \log(\alpha_H I^2/N_C) \tag{4.4}$$

where $S_{I,1/f}$ is the PSD of the current I, and α_H is the Hooge parameter. To further study the flicker noise properties that were presented in the previous sections, we now calculate the β and α_H parameters that describe Flicker noise. For this, a fit of the sections of the PSD where flicker noise is predominant is performed. The fit is performed to the frequency range of $200 - 3000$ Hz. The fits are shown in Fig. 4.9a and Fig. 4.8 by black dashed lines. For the fits, a linear fit to the logarithm of the PSD is performed so

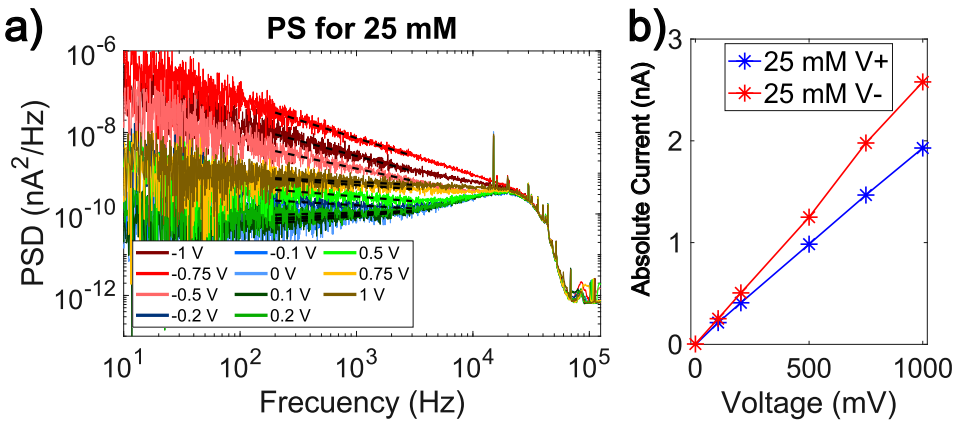


Figure 4.9: a) PSDs for several voltage conditions, for a nanopipette containing 25 mM NaCl. b) Magnitude of electric current vs. the magnitude of voltage for the electric current traces measured in a).

that the frequencies that were the PSD in orders of magnitude larger do not dominate the fit ($\log(S_{I,1/f}) = b \cdot \log(f) + a$). The slope of the fit is equal to the $b = -\beta$ in Eq.4.4, while the y-intercept is equal to $a = \ln(\alpha_H I^2/N_c)$.

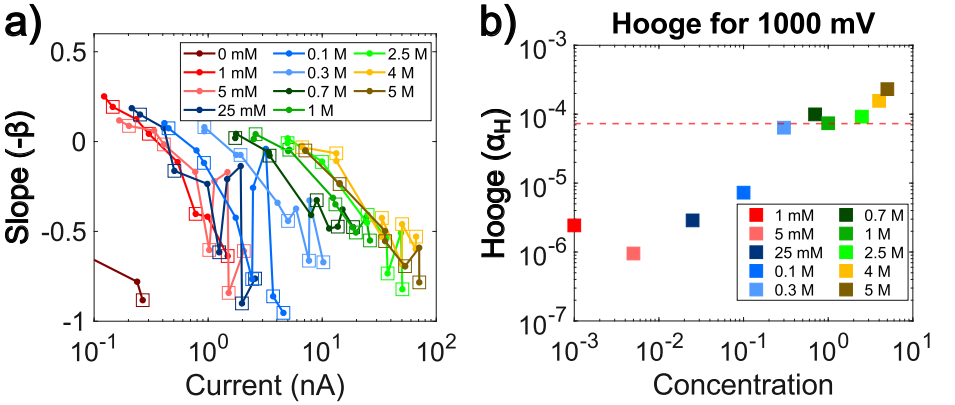


Figure 4.10: **Flicker noise parameters.** a) Slope values ($b = -\beta$) obtained from the linear fits of the PSD in the frequency range 200−3000 Hz. b) Hooge parameter (α_H) estimation for different concentrations for measurements performed at 1000 mV. The mean value of all the Hooge parameters is shown with the red dashed line $\alpha_H = 7.3 \cdot 10^{-5}$.

Fig.4.10a shows the slope ($b = -\beta$) values obtained by the linear fit vs. the absolute value of the current for the different concentrations and voltages. Moreover, slope values corresponding to negative voltages are identified with an extra square symbol around the circle symbol. It can be observed that for a given concentration, the slope values become more

negative with increasing current (and therefore voltage). This contrasts with previous results in nanopores, where flicker noise was found to have the same $\beta \sim 1$ value for different voltages[34]. There is a tendency for the flicker noise at the negative voltages to be higher (to have a bigger β value), even at high concentrations where the electric current is not dependent on the voltage polarity. This asymmetry in 1/f noise with voltage polarity was also observed in previous works in conical nanopores [109], and it is not well-understood [110]. Furthermore, from Fig.4.10a, it can be observed that for similar current values at different salt concentrations, the flicker noise is lower for the higher concentration (smaller β), where voltage is higher, indicating that the β parameter of flicker noise has a dependence on voltage, such that β increases with voltage.

From the y-intercept (a) of the linear fit, the Hooge parameter (α_H) can be estimated using the formula $\alpha_H = e^a N_c / I^2$. For this, the number of charge carriers inside the nanopipette needs to be estimated. This value is estimated by multiplying the charges per volume at each concentration by the volume of the nanopipette sensing region ($N_c = c \cdot V_{tip}$). We have chosen the first 150 nm of the tip as the sensing region of the nanopipette[29]. The estimated Hooge parameter for each concentration at 1000 mV is shown in Fig.4.10b, and the mean is shown as a dashed red line. From the mean value, we obtain $7.3 \cdot 10^{-5}$, which is in accordance with other Hooge parameters reported in the literature[111], for example, $1.1 \cdot 10^{-4}$ in solid-state nanopores [112]. However, it can be observed that the Hooge parameter is about one order of magnitude smaller at low salt concentrations. This is probably due to an underestimation of the number of charge carriers ($N_c = c \cdot V_{tip}$). At low concentrations, the concentration inside the nanopipette is higher than in bulk due to surface charge effects (rectification and Debye layer overlap), and N_c is underestimated.

4.5 Conclusions

In this chapter, we have presented the results of current measurements in nanopipettes over a broad range of concentrations at different voltages and for four different salts, LiCl, NaCl, KCl, and $MgCl_2$. Nanopipettes have asymmetric I-V plots, with negative voltages having a higher conductance due to the rectification effect. Rectification is significant for monovalent salts for concentrations up to 1 M, while for the case of divalent salts, it is only significant up to 100 mM. This shows the stronger shielding effect that divalent cations like Mg^{2+} have over monovalent. Divalent cations are more effective in suppressing surface charge-related effects.

The conductance of nanopipettes, in concentrations ranging from 1 mM to 5 M, was reproduced using a conductance model for the nanopipette that

considers its conical geometry and the negatively charged surface of the nanopipettes. The model was used to fit the data, considering the linear and nonlinear conductivity models for the salt solutions. The non-linear conductivity better reproduced the data at high concentrations, above 1 M. Furthermore, when estimating the diameter using the conductance at just one concentration, the nonlinear model converges to a diameter, while the linear model always gives smaller diameters. This suggests that the nonlinear conductivity model works better at high ionic strengths. This is relevant for the next chapters, in which we will estimate nanopipette diameters using the nonlinear model.

Regarding Flicker noise ($1/f^{\beta}$) in nanopipettes, we found that the β parameter increases as the voltage increases. This contrasts with previous results in nanopores where β was found to be voltage-independent [34]. Moreover, we observed that flicker noise is not symmetric and is higher at negative polarities; hence, for the same voltages, (β) is larger for the negative polarity. This asymmetric behavior for flicker noise in nanopipettes was also observed in previous works [109].

Understanding the different phenomena affecting the conductance and noise in nanopipettes is important for the next chapters, where nanopipettes will be used to study the translocation of DNA molecules in high salt concentrations of the three monovalent salts presented in this chapter, LiCl, NaCl, and KCl.

Chapter 5

Translocation of λ-DNA in high monovalent salt concentrations

The interaction between DNA and counter-ions in solution has a profound impact on its physicochemical properties, such as its persistence length, effective charge, and the intermolecular interactions with proteins and ligands [113, 79]. Theoretical approaches based on the Poisson-Boltzmann equation [114] and Manning's counterion condensation theory [115] ignore features such as the ion type/size, DNA's discrete charge distribution, and ion-ion interactions.

Nanopores have proven to be a versatile tool for detecting and studying single molecules [116, 117, 118]. In particular, DNA translocation experiments through nanopores show that the type of monovalent cation (Li^+, Na^+, and K^+) significantly affects translocation dwell times. It has been observed that the smaller the cation, the lower the DNA translocation speed [61]. This effect also depends on salt concentration, as higher salt concentrations produce lower translocation speeds.

This chapter presents experimental results on the translocation of λ-DNA through nanopipettes under high salt concentrations of 1, 2, and 4 M. The focus is on investigating how translocation parameters, including dwell time, mean current blockade, and charge blockade change for several experimental conditions. The conditions explored are voltage, nanopipette tip diameters, salt concentrations, and salt types. The chapter is organized as follows: first, the experimental protocols for conducting λ-DNA translocation experiments are detailed, followed by a description of the workflow used to analyze the data. Next, we investigate the charge blockade induced by DNA translocation events for experiments conducted at different applied

voltages and salt concentrations. Subsequently, the impact of salt type on translocation characteristics is studied, with a specific focus on how salt types influences current blockades and dwell times. A phenomenological scaling of dwell times is presented, enabling the identification of cation-specific effects. The chapter concludes with a summary and discussion of the findings.

5.1 λ-DNA translocation through nanopipettes

The DNA translocation setup described in Sec. 1.2.2 was used to perform the λ-DNA translocations experiments described in this chapter.

The λ-DNA molecule

λ-DNA is a 48,502 base pairs (bp) molecule. The contour length of a linearized λ-DNA molecule is $\sim 16.5\mu m$. λ-DNA is purchased from New England Biolabs in linearized form. λ-DNA has two 12 bp overhangs at its ends due to the linearization process. This means the 5' strands overhang the 3' strands by 12 bases. The overhangs are palindromic G-C-rich sequences called "sticky ends" as they can hybridize, forming a circular λ-DNA molecule or a concatemer of two or more λ-DNA molecules connected. Hence, a sample of λ-DNA may contain a portion of circularized λ-DNA or concatemers. Moreover, any sample containing DNA is prone to degradation over time. As a result, the sample also contains DNA fragments shorter than the λ-DNA molecule.

Experimental Protocol

A typical translocation experiment starts with preparing the microfluidic PDMS chip, which, in our case, contains six nanopipettes. Briefly, nanopipettes are fabricated by pulling quartz capillaries (0.2 mm ID and 0.5 mm OD) in the P-2000 pipette puller (Sutter instrument). After being pulled, the nanopipettes are cut and placed in a PDMS microfluidic chip, which is then sealed by plasma bonding the PDMS chip to a glass slide. Once the chip is ready, it is filled with the desired buffer. The plasma cleaner is used to make the inside of the microfluidic chamber hydrophilic to ease the filling of the nanopipettes with the buffer solution. Hydrophilicity helps to fill the nanopipettes. This step is crucial to ensure that air bubbles do not clog the nanopipettes, reducing also the 1/f noise. After filling the nanopipette and checking for nanopipettes with low noise and good tip size, λ-DNA is added to the cis chamber at a concentration of about 160 pM (see Fig.5.1). Further detail on the chip's design, preparation, and filling can be found

in Appendix A. The experiments are performed at 1, 2, and 4 M of KCl, NaCl, and LiCl, with $10mM$ of Tris, $10mM$ EDTA at pH 9.0. A high pH is chosen to make the quartz surface more negatively charged, reducing the DNA surface interactions. This can help to extend the nanopipettes' lifetime, as they take longer to clog during experiments. Measurements are taken at different voltages, 300-900 mV for each nanopipette. Experiments are performed at a controlled temperature of $25°C$. Ag/AgCl electrodes are used to detect the ionic current passing through the nanopipettes. The current signal is acquired using an Axopatch 200B (Axon Instruments) and filtered using an active 35 kHz low-pass external filter. The signal is finally digitized at 250 kHz using an acquisition card and saved to a computer using a Labview program.

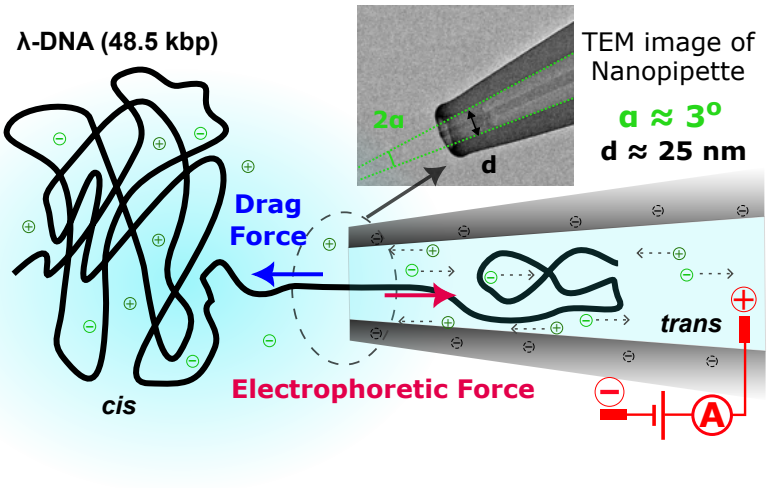


Figure 5.1: Schematic representation of a λ-DNA molecule, translocating through a nanopipette. During DNA translocation, the electrophoretic and the drag forces balance out. The inset shows the TEM image of a nanopipette.

For λ-DNA translocation experiments, the positive electrode is placed in the trans chamber, Fig.5.1, and the negative electrode is placed in the cis chamber. The DNA translocates from the cis chamber into the trans chamber across the nanopipette. During the experiments, the high electric field generated at the nanopipette's tip captures the DNA and then threads it through the nanopipette. This electric field is of the order of $10^5 - 10^7 V/m$, depending on the nanopipette's geometry and the applied voltage. The electric field is localized at the nanopipette tip, where the density of electric field lines is highest. The electric field decays rapidly on the cis side, where the field lines disperse into the bulk. The decay on the trans side is slower, as on this side, the nanopipette sections increase with a small angle of about 3^o, such that field line density decreases more slowly to the trans

side. As λ-DNA is a long molecule ($\sim 16.5\ \mu m$), the forces acting on it will be localized on the part of the molecule that is inside the first hundreds of nanometers inside the nanopipette ($100 - 200nm$). The remanding part of the molecule is in regions of smaller electric field. This is schematized in Fig.5.1, where it is shown that the only part of the molecule that is under tension is the section in the narrowest part of the nanopipette at the tip.

λ-DNA translocation events

A jump in the ionic current is observed when a λ-DNA molecule translocates through the nanopipette. This jump occurs because the DNA excludes a volume inside the nanopipette so that no ions can flow through the volume occupied by the DNA molecule. The volume exclusion causes an increase in the electric resistance, and hence, the current decreases ($I = V/R$). However, λ-DNA can translocate in different modes, corresponding to different folded configurations of the DNA molecule. In Fig.5.2, translocation events of different DNA folded configurations can be observed. Each translocation event is shown with a corresponding representation of the λ-DNA folded configuration that produced the observed current signature. The observed

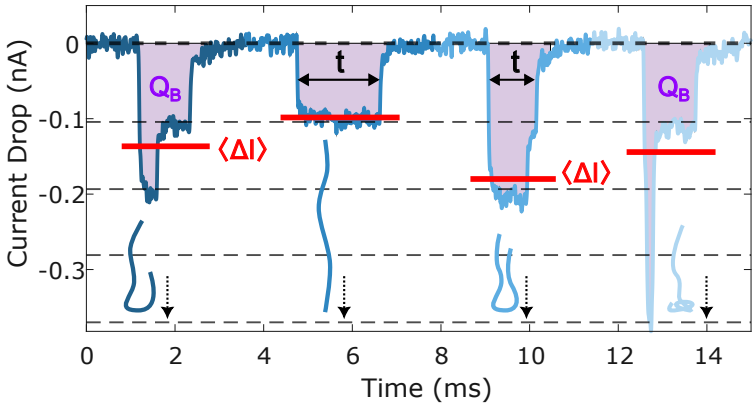


Figure 5.2: Different λ-DNA translocation events, showing the folded configuration that would produce the observed current signature at the bottom. For the translocation event, the charge blockade Q_B in purple, the mean current blockade $\langle \Delta I \rangle$, red line, and the dwell time t_D, in black, are shown.

current blockades occur at discrete levels, as shown by the dashed horizontal lines in Fig. 5.2. These levels correspond to the number of DNA duplexes entering the nanopipette during the translocation. Different parameters are identified for the translocation events in Fig.5.2. t_D is the time it takes for a molecule to translocate. $\langle \Delta I \rangle$ is the mean current blockades for a given translocation event. Q_B is the blocking charge defined as the area limited

by the translocation trace and the baseline across a translocation event. The charge blockade Q_B is defined as $Q_B = \langle \Delta I \rangle \cdot t_D$. The figure shows that translocations of more compact λ-DNA configurations have a higher $\langle \Delta I \rangle$ and a shorter t_D. However, as we will see in this chapter, Q_B remains constant across all translocation events at the same voltage.

5.2 λ-DNA analysis workflow

This section presents the workflow for analyzing the λ-DNA translocation data. Data are analyzed using a custom-made MATLAB code. The more important aspects of the code are explained here.

Event detection

During DNA translocation experiments, the electric current passing through the nanopipette is measured at any time. This signal is acquired at a 250 kHz sampling frequency. Hence, the time resolution of the raw signal is 0.004 ms. This signal must first be pre-treated to eliminate regions with high noise or where the nanopore was clogged. More information about the pre-treatment of the data can be found in Appendix D. The treated signal is then used for the subsequent analysis. The raw signal is first averaged using a box filter of 10 points to proceed with the detection of translocation events. This is done to reduce the data points for the translocation detection process. Experiments can take up to 20-30 min per voltage condition; recording at 250 kHz gives $\sim 4.5{\cdot}10^8$ data points. After the box filter, our time resolution is 0.04 ms, enough to detect λ -DNA translocations that are at least 0.7 ms long.

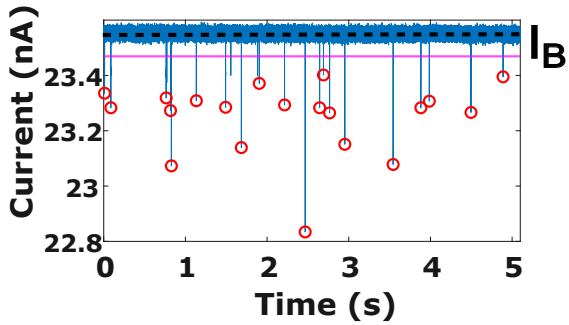


Figure 5.3: Current trace during a translocation experiment. The black dashed line shows the baseline current I_B, which is the current when no DNA is blocking the nanopipette. The pink line indicates the detection threshold for translocation events. Detected translocation events are marked with red circles.

To detect translocation events, we use an algorithm that detects current deviations from the baseline in the filtered signal I_B (black-dashed line in Fig.5.3) below a chosen threshold $\sim 0.05 - 0.2nA$ (pink line in Fig.5.3). The precise value of the threshold depends on the diameter of the tip and the experimental conditions, and it is taken as half the current blockade of the first level. The baseline current (I_B) value is measured locally, and its value changes if there is a drift in the current signal. When the signal deviates from the local baseline beyond the threshold and returns to the baseline after a few ms, it is detected as a DNA translocation event. The translocation events are then cut out of the signal with a few ms before and after the event for further analysis. For all detected translocation events (red dots in Fig.5.3) the local baseline current before the event $\langle I_B \rangle$ is calculated and subtracted to the event current, $I_{\text{event}} - \langle I_B \rangle$ so that for the subsequent analysis, the baseline of all events is centered at zero (as shown in Fig.5.2). The raw, and not the box filter signal, is saved in this step. Finally, the current traces for each event are used to calculate the dwell time t_D, charge blockade Q_B, and mean current blockades $\langle \Delta I \rangle$ for each event as shown in Fig. 5.2.

Single λ-DNA event selection

After detection, the translocation events must be filtered to separate those corresponding to a single λ-DNA molecule from the other events, such as DNA fragments or concatemers of multiple λ-DNA molecules.

To reliably separate the single DNA events from the rest, we use a filtering Bayesian-based procedure that uses the charge blockade $Q_B = \langle \Delta I \rangle \cdot t_D$ as a key parameter to classifying events. The charge blockade is a good parameter for separating translocation events by the molecule's size, as it remains invariant independently of the folded configuration of the DNA molecule. Translocations of linear λ-DNA molecules, like the second example in Fig. 5.2, will have longer t_D, but smaller $\langle \Delta I \rangle$. If the molecules translocate in a more compact configuration, as the third example in Fig. 5.2, t_D will be shorter, and $\langle \Delta I \rangle$ will be bigger. However, the Q_B's of the translocation events are distributed around a constant Q_B value for translocations at the same voltage. Therefore, Q_B is a good parameter to separate translocation events from λ-DNA molecules from fragments or concatemers.

The steps of the selection algorithm are explained in Fig. 5.4. First, all detected events are plotted on a t_D versus $\langle \Delta I \rangle$ plot, Fig. 5.4a. Events with the same Q_B will fall on a $1/x$ curve on this plot. It can be observed that a significant portion of the translocations are contained in a cloud around a $1/x$ curve. An initial seed (Q_B^0) is chosen for the value of Q_B^0, ($Q_B^0 \sim 0.19$ pC

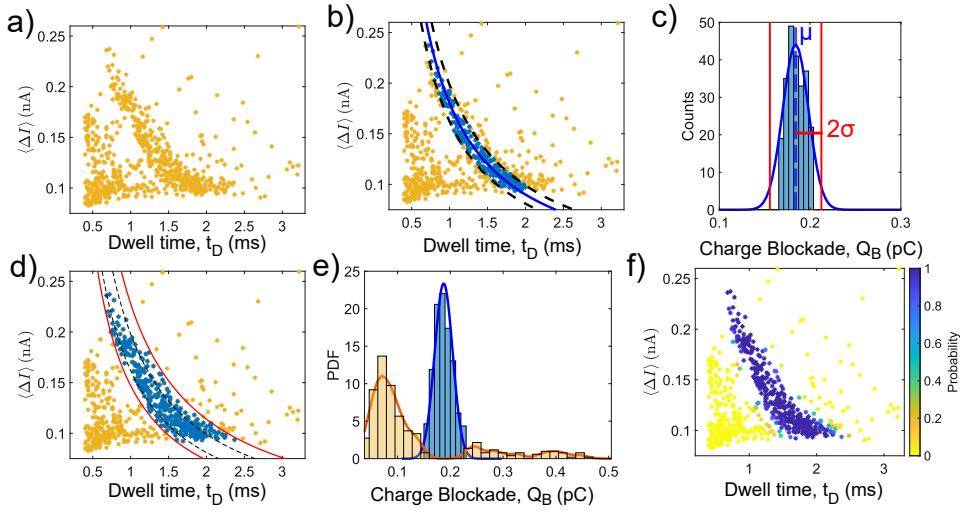


Figure 5.4: **Event selection algorithm.** Steps of the algorithm used to select single λ-DNA translocations from translocation events due to fragments or concatemers. a) All detected translocation events. b) Pre-selection of single λ-DNA events, which have Q_B values between $Q_B^0 \cdot (1-s)$ and $Q_B^0 \cdot (1+s)$. c) Gaussian fit to Q_B histograms of the pre-select events, blue events shown in b. The blue continuous line represents the fit. From this section, we calculated σ for the next selection. The black lines are at $\pm 2\sigma$ of the Gaussian mean value (μ). d) Second selection using parameters (σ and μ) of the Gaussian fit. e) The Gaussian fit of this second selection (blue continuous line) determines the single λ-DNA PDF. For the non-selected events, the PDF is determined using a kernel estimator (orange continuous line). f) Probability of events to belong to the single λ-DNA type. The probabilities are calculated from the PDF distributions in e.

in Fig. 5.4b), together with a spread (s) (typically $s \sim 0.1-0.25$). All events falling in the region delimited by $Q_B^0 \times (1-s)$ and $Q_B^0 \times (1+s)$ are pre-selected as single-DNA events, Fig. 5.4b. The confidence range, defined by the spread s, is chosen to contain most points clustered around the initial seed Q_B^0 that are visibly separated from the rest. We then fit the Q_B distribution of the pre-selected set of events with a Gaussian distribution (blue line, Fig. 5.4c). The mean value of the Gaussian function (μ) and standard deviation (vertical red lines in Fig. 5.4c are at $\mu \pm 2\sigma$) define a new region delimited by the values $Q_B^{\text{lower}} = \mu \cdot (1-2\sigma)$ and $Q_B^{\text{upper}} = \mu \cdot (1+2\sigma)$ (red lines Fig. 5.4d). The dashed lines in Fig. 5.4d show the limits from the pre-selection in the first step. Points within this new region are selected as single λ-DNA translocations and fitted again to a Gaussian distribution. This Gaussian is normalized and it is now considered the probability distribution function (PDF) of single λ-DNA events (Fig. 5.4e, blue Gaussian). A kernel density estimator is then used for the nonselected events outside the blue region

to estimate their PDF (Fig. 5.4e, orange PDF). We then use a Bayesian procedure to assign a probability for each point to belong to the blue PDF, according to their Q_B value (Fig. 5.4f). Finally, a random selection using the calculated probabilities is performed for the final selection of single λ-DNA events. The final selected events are shown in blue in Fig.5.5a.

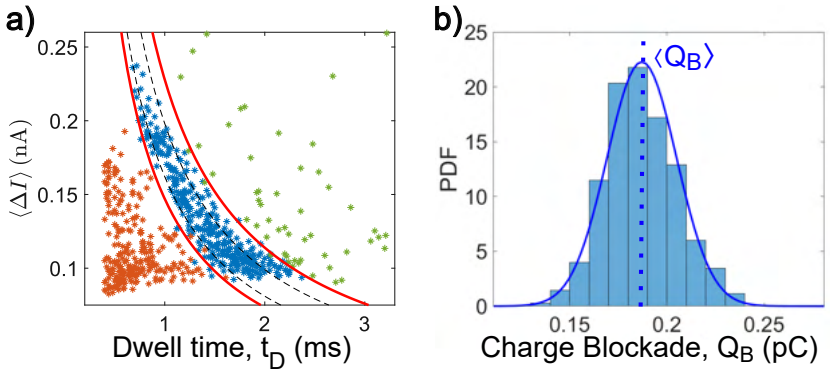


Figure 5.5: a) Final selection of λ-DNA translocation events after the selection algorithm. b) Q_B histogram of the single λ-DNA translocation fitted with a Gaussian function (blue continuous line), from which we derive mean $\langle Q_B \rangle$ for the given experimental condition.

The nonselected events are further subdivided into orange and green, depending on whether they fall to the left or right of the blue set (Fig. 5.5a). The orange points correspond to translations of fragments shorter than λ-DNA, and the green to the translocation of concatemers longer than λ-DNA. As shown in Fig. 5.5a, a significant portion of the blue points ($55-70\%$) are single λ-DNA events, whereas the rest are fragments (orange points, lower Q_B) or conatemers with more than one DNA (green points, larger Q_B). The Q_B histogram of the single λ-DNA events is shown in 5.5b, and has a Gaussian distribution. The histogram is fitted with a Gaussian distribution to determine the mean $\langle Q_B \rangle$. The different types of events can be observed in Fig. 5.6a, where translocation events of single λ-DNA are shown with blue colors, fragments in orange/red, and concatemers in green colors. In Fig. 5.6b, the histograms of the charge blockade of the different types of events are shown, and it can be seen that fragments (orange) have smaller Q_B, than single λ-DNA events, which have smaller Q_B than concatemers.

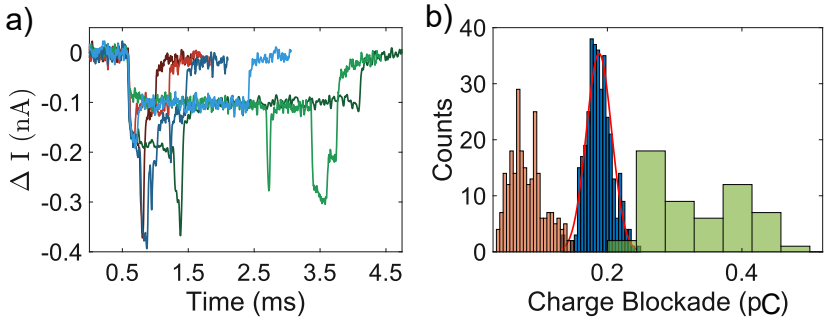


Figure 5.6: a) Examples of the three different types of translocation events are shown. Translocation events of DNA fragments in red, single λ-DNA molecules in blue, and concatemer translocation green. b) Q_B histograms of the three different types of translocations.

Mean parameters of single λ-DNA translocations

Once the translocation events of a single λ-DNA have been selected, the mean charge blockade and the mean dwell time of all events are determined. Furthermore, the current blockade of the different levels is estimated. A Gaussian fit is performed to the Q_B histogram to determine the mean charge blockade $\langle Q_B \rangle$, as shown in Fig. 5.5b. Similarly, the dwell time histogram is fitted with a Gaussian function to estimate the mean dwell time $\langle t_D \rangle$ (Fig.5.7 top, dashed line). The current blockades of each level (ΔI_n) are determined by fitting the current blockade histogram with a sum of multiple Gaussian, as shown in Fig.5.7 right, each Gaussian corresponding to one of the levels.

When conducting λ-DNA translocation experiments, translocation events are acquired at several voltage conditions for each nanopipette. The analysis described in this section is performed on each voltage condition. The results presented in this chapter come from the analysis of nanopipettes measured at 1, 2, and 4 M of NaCl, KCl, and LiCl salt types. For each nanopipette, we performed measurements at several voltages. Only those voltages containing at least $\sim$ 100 single λ-DNA translocations events are considered for further analysis. To estimate the diameter of the nanopipettes, we used Eq. 1.13 with the geometry parameters estimated from the TEM images in Chapter 4 (more details are found in Sec. 4.3.4). Several voltage measurements are performed for each nanopipette; hence, the nanopipette diameter is computed by taking the mean value of the diameters obtained for the different voltages. These diameters usually do not differ more than $\sim 1-3$ nm.

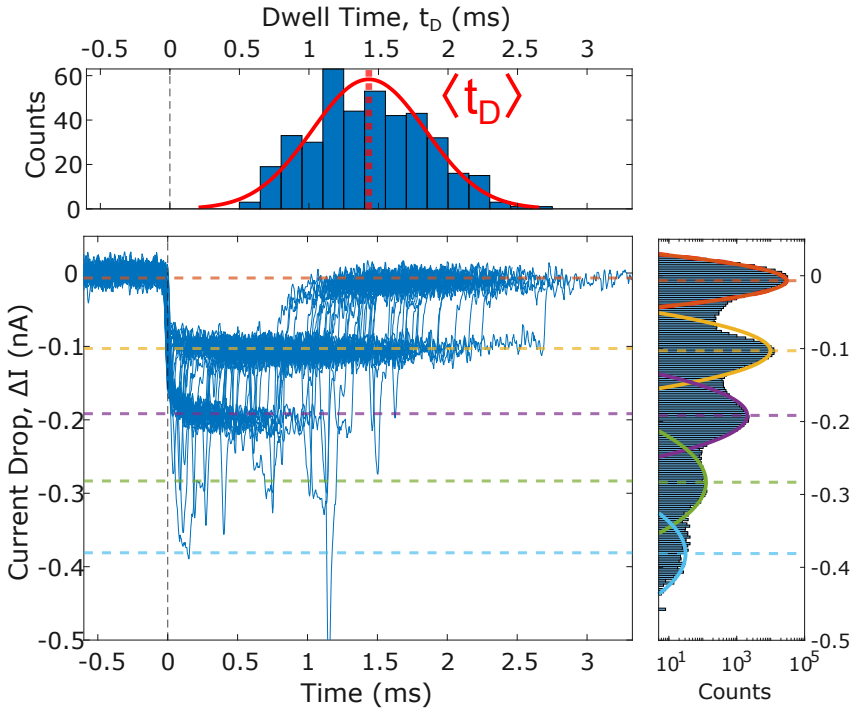


Figure 5.7: **Analysis of single λ-DNA events.** From the selected single λ-DNA events, the mean dwell time $\langle t_D \rangle$ is calculated by fitting a Gaussian function to the dwell time histogram (top). The current blockade (ΔI_n) histogram is fitted using a multiple Gaussian fit to extract the mean current block of each level (right).

5.3 Investigating charge blockade

In the previous sections, we have seen how the blockade charge $Q_B = \langle \Delta I \rangle \cdot t_D$ is constant for translocation events at a given voltage and can be used to select single λ-DNA translocation events. Q_B remains unchanged because λ-DNA translocates at a constant velocity, independently of the folded configuration in which it translocates. Therefore, for a U-type configuration (as exemplified in Fig. 6.1) where two DNA duplexes are blocking the nanopipette at any time, the t_D will be half that of a linear translocation (see Fig. 6.1), but $\langle \Delta I \rangle$, will be double.

This section studies whether the mean charge blockade $\langle Q_B \rangle$, of λ-DNA events is constant for different voltage values. When the voltage increases, $\langle t_D \rangle$ decreases proportionally to $1/V$, while the mean current blockade $\langle \Delta I \rangle$ increases. Charge blockade is given by $Q_B = \langle \Delta I \rangle \cdot t_D$, so for it to remain unchanged when voltage is changed, the decrease in t_D needs to be compensated with a corresponding increase in $\langle \Delta I \rangle$.

To see the effect voltage has on Q_B, we first compare the results of

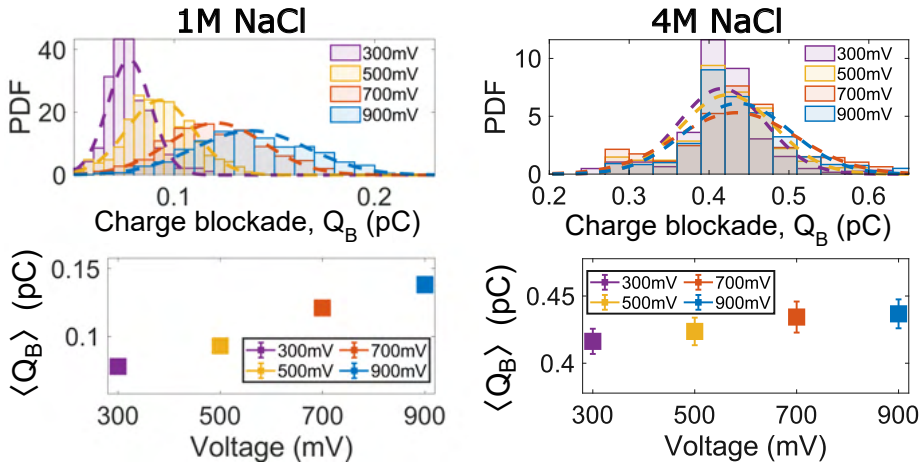


Figure 5.8: **Charge blockade.** Top. Charge blockade distributions for 300, 500, 700, and 900 mV. The distributions have been normalized to have an area of 1. The dashed lines are Gaussian fits to the distributions. Left a nanopipette with a $d = 36$ nm and 1 M NaCl, and right a nanopipette with a $d = 15$ nm and 4 M NaCl. Bottom. The mean $\langle Q_B \rangle$ estimated from the Gaussian fits as a function of voltage.

two nanopipettes at 1 and 4 M NaCl. The Q_B distributions for different voltages are shown in Fig. 5.8 top. It can be observed that, for the 1 M concentration, the mean of the distributions increases as the voltage increases. In comparison, for the 4 M concentration, the mean of the Q_B distributions remain constant with voltage.

The mean value of the Q_B distributions $\langle Q_B \rangle$, are plotted against voltages in Fig. 5.8 bottom. From these plots, it can be observed that while for 1 M, the relation between Q_B at 900 and 300 mV is $Q_{900}/Q_{300} = 1.8$, for 4 M it is $Q_{900}/Q_{300} = 1.0$. This indicates that $\langle Q_B \rangle$ is a constant quantity at the higher 4 M concentrations, while it is not constant at the lower 1 M concentration. All the translocation events for different voltages can be plotted onto a scatter plot of t_D vs. $\langle \Delta I \rangle$ (Fig.5.9). The figure shows the $\langle Q_B \rangle$ value for each voltage with a line of the corresponding color, while the mean value of the $\langle Q_B \rangle$'s taken over the four voltages is shown with the black dashed line. While for 4 M, all events fall along a curve of constant Q_B (Fig. 5.9 right panel), for the 1 M case, deviations are observed, as the Q_B value increases with voltage (Fig. 5.9 left panel).

The effect of Q_B not being constant comes from t_D and $\langle \Delta I \rangle$ not having an opposing behavior when voltage changes. The current blockade for a level n is given by $\Delta I_n = -n(V/R) \cdot (d_{DNA}/d_{Pore})^2$ (Eq. 1.24), is proportional to voltage, and the relation $\Delta I_n/I_B$ should be voltage-independent. Where I_B is the baseline current (see Fig. 5.3), and R the

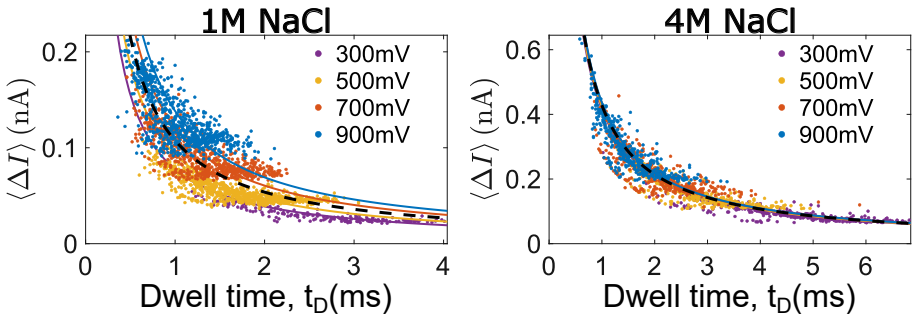


Figure 5.9: $\langle \Delta I \rangle$ **vs.** t_D **scatter plots.** a) Scatter plot showing the $\langle \Delta I \rangle$ vs. t_D of all translocation events for 300, 500, 700, and 900 mV. Data are from a nanopipette with a 36 nm tip diameter in 1 M NaCl. b) Same scatter plot for a nanopipette of 15 nm tip diameter in 4 M NaCl.

resistance of the nanopipette. The current blockades of the different levels, ΔI_n, are calculated by the multiple Gaussian fit to the current drop histogram (shown in Fig. 5.7). To observe where the increase in Q_B comes from, we plot the $\langle t_D \rangle$ for the different voltages, Fig. 5.10 top. We can observe how the $\langle t_D \rangle$ follows a $1/V$ curve (dashed lines in Fig. 5.10 top). However, we can observe differences in the current blockades at different voltages. While for 4 M, the current blockades ($\Delta I_n/I_B$) all collapse onto the same line for different voltages, this is not the case for the 1 M concentration. At 1 M, we observed a higher current blockade for higher voltages. For 4 M $(\Delta I_1/I_B)_{900}/(\Delta I_1/I_B)_{300} = 1.0$ while for 1 M $(\Delta I_1/I_B)_{900}/(\Delta I_1/I_B)_{300} = 1.5$. Hence, the increase in Q_B for 1 M is because of an increase in the current blockade, while the $\langle t_D \rangle$ follows the usual $1/V$ behavior. Therefore, there is a voltage effect in the current blockade during translocation, such that the relation $\Delta I_n/I_B$ is not voltage independent at the lower 1 M concentration.

Plotting the relation of Q_B at 900 mV and 300 mV (Q_{900}/Q_{300}) vs. nanopipette tip diameter for different concentrations of NaCl Fig.5.11, it can be observed how this relation decreases with increasing concentration. The nanopipettes with a 1 M concentration have a higher ratio Q_{900}/Q_{300}, and this ratio decreases when concentration increases. However, there is not a clear relation with nanopipette diameter. Although only data in NaCl were presented in this section, the same behavior is observed for charge blockades for LiCl and KCl.

The increase in the Q_B is caused by an increase in $(\Delta I_1/I)_{900}/(\Delta I_1/I)_{300}$ for the 1 M concentration. We believe this is caused by the salt concentration being lower at the nanopipette tip for higher voltages. These are called concentration polarization effects. The presence of the highly negatively charged DNA inside the nanopipette during translocation enhances the

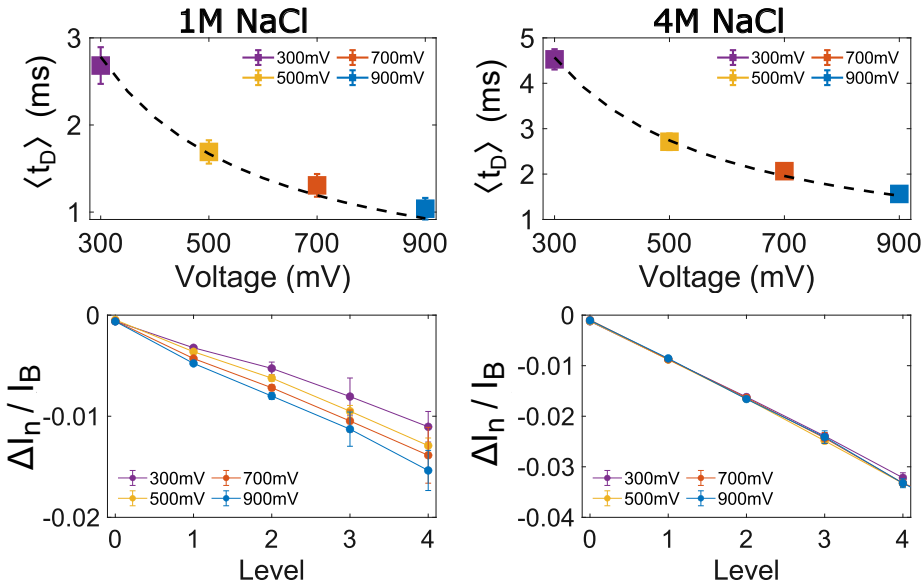


Figure 5.10: **Dwell times and current blockades.** Top. Mean dwell times $\langle t_D \rangle$ as a function of voltage for 1 and 4 M NaCl concentrations. Dashed lines are fitted with a $1/V$ function. Bottom normalized current blockades ($\Delta I_n / I_B$) for the different levels as a function of voltage for 1 and 4 M NaCl.

nanopipette's charge asymmetry. This slightly decreases salt concentration inside the nanopipette during DNA translocation(when a positive voltage is applied). This is the same effect that causes rectification at lower concentrations in nanopipettes. Here, the effect is observed at higher concentrations due to the increased charge asymmetry caused by the presence of DNA inside the nanopipette. Concentration polarization is voltage-dependent; hence, the concentration decrease inside the nanopipette is larger for higher voltages, which increases $\Delta I / I_B$. The concentration polarization effect has been reported previously and is responsible for the significant difference in current traces between forward and backward DNA translocations when the voltage is inverted [119]. The increased screening of DNA's charge at 2 and 4 M reduces this effect, such that it is less significant at these higher concentrations.

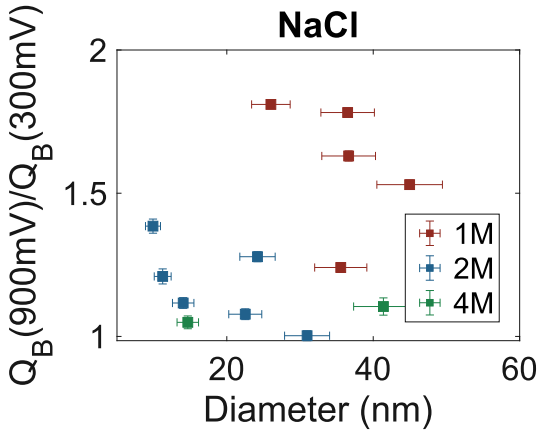


Figure 5.11: Ratio of charge blockade at 900 and 300 mV, Q_{900}/Q_{300}, plotted versus nanopipette diameter, for 1, 2, and 4 M concentrations of NaCl.

5.4 Cation size effect on current blockades

Here, we study the effect on the current blockade fraction for the first level $\Delta I_1/I_B$, when one DNA strand blocks the nanopipette for different salt types. We focus on the current blockade of the first level as it is the easiest to determine, the other levels being proportional to it (Eq.1.24). According to Eq.1.24, the current blockade at high salt concentrations for just one DNA strand blocking a nanopore is given by:

$$-\frac{\Delta I_1}{I_B} = -\frac{\Delta G_1}{G} = \left(\frac{d_{DNA}}{d_P}\right)^2 \tag{5.1}$$

with the crystallographic diameter of the DNA helix $d_{DNA} = 2.0$ nm, and d_P the nanopore diameter. It can be observed that the equation is independent of salt concentration and type.

In Fig. 5.12, the conductance drop percent ($\left|\frac{\Delta G_1}{G}\right| \cdot 100$) vs. the nanopipette diameter is plotted for the three monovalent salts. All the data correspond to measurements at 700 mV, so voltage effects on the current blockades are avoided. In Fig. 5.12, all points at 1, 2, and 4 M for a given salt type are plotted together. (In Fig. C.1 shown in Appendix C, it can be observed that, independently of the concentrations, all the points fall onto the same curve.) From Fig. 5.12, it can be observed that the conductance drop is bigger for LiCl than for NaCl and KCl. Performing a fit with Eq.5.1, with d_{DNA} as a free parameter, the diameter of DNA can be estimated for each salt type. The diameter values are recollected in Table 5.1. Using Eq.1.24, we obtain a reasonable estimate of the diameter, and the current blockade dependence with the tip diameter of the nanopipette

is correctly reproduced. However, the simple model given by Eq. 5.1 does not explain the differences in current blockade between the different monovalent salts, as the DNA crystallographic diameter does not vary between different monovalent ions. Although some works have reported changes in DNA's crystallographic diameter for different monovalent salts LiCl, NaCl, and KCl, these differences were of ~ 0.02 nm [120] and indicated a smaller radius in LiCl. Therefore, a change in the crystallographic radius can not explain the observed difference. To understand what could be causing this difference between the salts, a more complex model to describe the charge blockade during translocation needs to be considered.

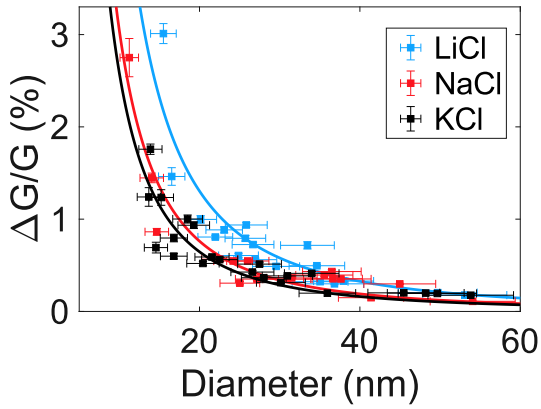


Figure 5.12: Conductance drop percentage ($\left|\frac{\Delta G_1}{G}\right| \cdot 100$) vs. nanopipette diameter, for different salt types, for λ-DNA translocations at 700 mV.

Table 5.1: **DNA diameter estimation.** d_{DNA} values obtained by fitting Eq.1.24 to the date in Fig.5.12

	LiCl	NaCl	KCl
d_{DNA} (nm)	2.3 ± 0.2	1.7 ± 0.1	1.6 ± 0.1

Improved model for translocation current blockades

Although Eq.1.24 gives a good estimation of the DNA diameter and correctly reproduces the current blockade dependence with diameter, the equation is based on an incorrect conceptual model. The equation is based on the idea that when DNA translocates, it decreases the number of ions inside the nanopipette. Hence, there is a conductance drop during the translocation of the DNA molecule. However, by doing some back-of-the-envelope calculations described in the next paragraph, one can see that this state-

ment is incorrect. When DNA enters the nanopipette, it brings a cloud of counterions that shield the negative charge of the DNA backbone. It is only at very high salt concentrations that the presence of DNA inside the nanopipette reduces the number of ions inside it.

For a 1 M monovalent salt concentration, the ion density per volume is 1.2 ions/nm^3, 0.6 cations/nm^3, and 0.6 anions/nm^3. If we consider that DNA has $2e^-$ negative charges every base pair step of 0.34 nm and a diameter of 2.0 nm for the DNA helix, we obtain for DNA a charge density per volume of 1.7 e^-/nm^3. To conserve charge neutrality, when DNA enters into the nanopipette, an opposite charge cloud (described by the Debye length) of cations comes into the nanopipette, shielding the DNA's charge, as seen in the scheme in Fig.5.13a. The cloud is an extra charge that comes into the nanopipette due to the presence of DNA and is of 1.7 e^+ cations per nm^3 of DNA volume inside the nanopipette. Therefore, for DNA to reduce the number of ions inside the nanopipette, the ion concentration must be higher than this value of 1.7 e^+/nm^3. If this occurs, DNA will exclude more ions due to volume exclusion than the cations it brings into the nanopipette due to its negative charge. The concentration at which DNA excludes the same number of ions as the number of cations it brings into the nanopipette occurs for a concentration of 1.4 M. For higher concentrations, the ions excluded due to volume exclusion exceed the extra cations that shield DNA's charge. Although we have to consider the case of a nanopipette here, this calculation holds equally for the case of nanopores.

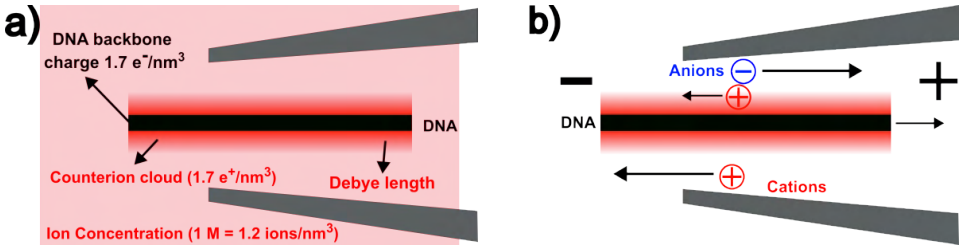


Figure 5.13: a) Scheme of the ion concentrations inside the nanopipette during DNA translocation. Note that the values of 1.7 e^+/nm^3 and 1.7 e^-/nm^3 for the counterions cloud and the DAN backbone charge are per volume of DNA inside the nanopipette. b) The scheme shows the reduction of cation mobility during DNA translocation due to the electrofriction caused by the DNA moving in the opposite direction from the cations.

Kesselheime et al. [121] showed, using atomistic simulations of DNA translocation in KCl, that the ion concentration inside the nanopipette is higher when DNA translocates. They showed that current blockade during DNA translocation occurs because DNA inside the nanopore reduces the mobility of ions, mainly cations close to DNA in the shielding cloud.

Therefore, the current blockade is not caused by a reduction in the number of ions in the nanopore but rather by a reduction in the mobility of the cations near DNA, which leads to a conductance drop. They found that the crossover concentration for which the number of ions excluded from the pore due to the volume of the DNA exceeds the number of extra ions in the counterions cloud was 1.2 M. This value is close to the value of 1.4 M of our back-of-the-envelope calculations. A scheme of this phenomenon, where cations near the DNA have lower mobility, is shown in Fig. 5.13b. The reduction in mobility is caused by the interaction of DNA with the cations around it. As cations and DNA move in opposite directions, the movement of DNA generates 'electrofriction' on the ions, reducing their mobility [122].

Considering the bigger current blockade observed for LiCl in contrast to NaCl and KCl. The stronger interaction of DNA with the Li^+ cations, caused by the higher charge density due to a smaller ionic radius (see Table 5.2), could cause a higher reduction in the mobility of the cations. This mobility reduction would be lower for Na^+ and K^+ cations, which have a smaller charge density. Hence, this would cause a larger current blockade for translocations in LiCl than in NaCl and KCl. This salt-type dependent effect will be further discussed in the next sections.

Table 5.2: Ionic radius of monovalent ions (in pm) [54].

	Li^+	Na^+	K^+
Ionic Radius (pm)	90	116	152

5.5 Phenomenological scaling for dwell times

In the previous section, we presented results indicating that smaller cations interact more strongly with DNA. Therefore, the smaller the cations, the greater the reduction in mobility when they move through the nanopipette in the presence of DNA. However, a stronger DNA-ion interaction will also affect DNA dynamics. In this section, we will investigate the effects of salt type on DNA dwell times.

We start by introducing a simple model that relates dwell time (t_D) to the different experimental conditions in our experiments. The DNA velocity during translocation is determined by a balance between the electrophoretic force pulling on DNA and the drag force. As is shown in the scheme in Fig. 5.14, these forces balance out, and as a result, translocation occurs at a constant velocity. It has been shown that DNA translocation happens at a constant velocity up to the last section of the molecule when the DNA tail enters the nanopipette, where translocation speed increases [123]. However,

since λ-DNA is a very long molecule ($\sim 16.5\ \mu m$), the speed-up effect of the last part of DNA occurs in a small fraction of the total translocation time.

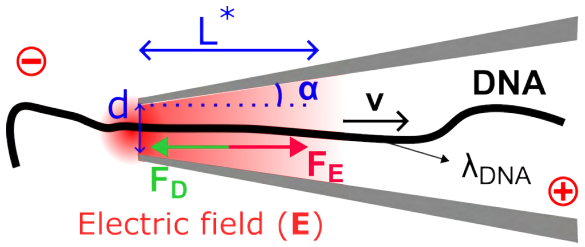


Figure 5.14: Schematic representation of the forces involved during DNA translocation through a nanopipette. More details of the model are described in the text.

The electrophoretic force acting on DNA is given by $F_E = E\lambda_{DNA}L^*$, where E is the electric field, and $\lambda_{DNA}L^*$ is the effective charge of the DNA inside the nanopipette. The electric field at the nanopipette tip can be approximated by $E = 2tan(\alpha)V/d$ (see Eq. 1.22), where V is the applied voltage, d is the nanopipette tip diameter, and α is the nanopipette angle (see Fig. 5.14). λ_{DNA} is the DNA's effective charge per unit length, and L^* is the length of the nanopipette's region where the electric field is strongest (the parameters are shown in Fig. 5.14). DNA's effective charge depends on the ionic conditions, as higher ion concentrations will shield DNA's charge more effectively. To account for this effect, we consider a charge per unit length inversely proportional to the square root of concentration, in analogy with Debye length. Therefore, $\lambda_{DNA} = \lambda_0/\sqrt{c/c_0}$, where λ_0 is the charge per unit length of DNA at a concentration $c_0 = 1$ M. The electrophoretic force is balanced by a hydrodynamic drag force on the DNA due to its fast movement inside the nanopipette, plus the drag of the DNA coil in the cis-side outside the nanopipette [124]. We assume that the drag force is given by $F_D = \gamma v$, where v is the translocation speed, and $\gamma = \eta R^*$ is the drag coefficient with η the viscosity and R^* length parameter that describes the friction acting onto the DNA. Considering the balance of forces $F_E = F_D$:

$$F_E = F_D \rightarrow E\lambda_{DNA}L^* = v\eta R^* \rightarrow v = \frac{E\lambda_{DNA}L^*}{\eta R^*} \tag{5.2}$$

using now that the translocation velocity is given by $v = L_{DNA}/t_D$, where L_{DNA} is the length of the DNA molecule, we can rewrite the dwell time as:

$$t_D = \frac{\kappa\eta}{E\lambda_{DNA}L^*} \rightarrow \frac{t_D}{\eta} = \frac{\beta}{\sqrt{c/c_0}V/d} \tag{5.3}$$

using $\kappa = R^*L_{DNA}$ is a parameter with units of $[m^2]$. β is a parameter given by $\beta = \kappa/(2tan(\alpha)\lambda_0 L^*)$, which has units of $[m^2/C]$. This equation

relates all the parameters that can vary in the experiment: the voltage (V), the nanopipette tip diameter (d), and the salt concentration (c), with its corresponding change in viscosity (η).

Panels a, c, and e in Fig.5.15 show the $\langle t_D \rangle$ for nanopipettes of variable tip sizes at different salt concentrations and monovalent salt types (KCl, NaCl, and LiCl). From Fig.5.15a, c, and e, it can be observed that translocation dwell times follow two trends for a given salt type. First, for a fixed salt concentration, dwell times increase with the tip diameter of the nanopipette (light to dark color in Fig. 5.15). Second, dwell times increase with salt concentration for a fixed tip diameter (red for 1 M, blue for 2 M, and green for 4 M).

We searched for a phenomenological scaling relation to better compare the dwell time results obtained under different salt types independently of the other experimental conditions (salt concentration and nanopipette diameter). Scaling all data into one master curve allows us to extract the specific effect of each cation. Different quantities varying across the experiments are used to perform this scaling. These are the viscosity of the buffer (η), the salt concentration (c), and the size of the nanopipette diameter (d). The scaling collapses all data points onto a master curve when the dwell time is divided by the viscosity, and the voltage is divided by the nanopipette diameter and the square root of the salt concentration, as seen in Eq. 5.3. Table 5.3 shows the viscosity values used for the scaling. The rescaled data are shown in Fig. 5.15 b, d, and f.

Table 5.3: **Dynamic viscosity at 25^oC (298 K) of the different salt solutions.** The viscosity values are in ($mPa \cdot s$). The viscosity of pure water at 25^oC is 0.89 $mPa \cdot s$. The last column shows the cation radius in pm. Viscosity data are from [125].

Salts	1 M	2 M	4 M	Cation Radius (pm)
LiCl	1.02	1.18	1.58	90
NaCl	0.98	1.09	1.41	116
KCl	0.89	0.90	0.96	152

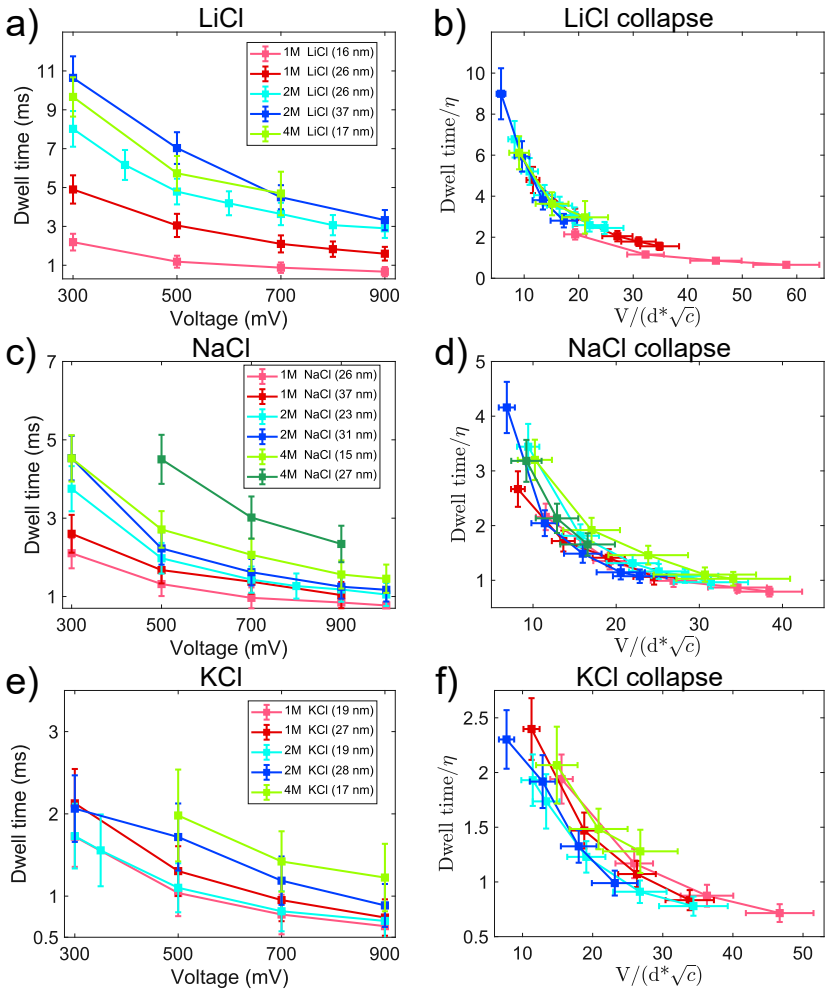


Figure 5.15: a) Dwell times for different salt concentrations and tip diameters in LiCl. b) Dwell time data after the phenomenological collapse described in the text. c) and d) are the same but in NaCl. e) and f) are the same but in KCl.

By plotting all the rescaled data onto the same plot, we can now observe the effect of the cation size on the translocation dwell times as shown in Fig. 5.16. It can be seen that the smaller the cation size (Table 5.3, last column), the bigger the effect on the dwell time. Dwell times follow the $Li^+ > Na^+ \sim K^+$. Performing a fit with the algebraic function Eq. 5.3, with β as the free parameter, we obtain the β values shown in Table 5.4. Recalling the equation $\beta = \kappa/(2tan(\alpha)\lambda_0 L^*)$, the parameters α and L^* are common for all the nanopipettes independently of the salt type. Therefore, the higher β value for LiCl is probably caused by two factors. The smaller Li^+ cations cause a higher screening of DNA's charge, reducing λ_0. In addition, κ could also depend on the cation type, as smaller cations interact

stronger with DNA and hence could cause a higher electrofriction on DNA. However, it is difficult to determine from translocation experiments to which extent each factor contributes to the increase in β. We will further discuss the effect of charge reduction or electrofriction on dwell times in Chapter 7.

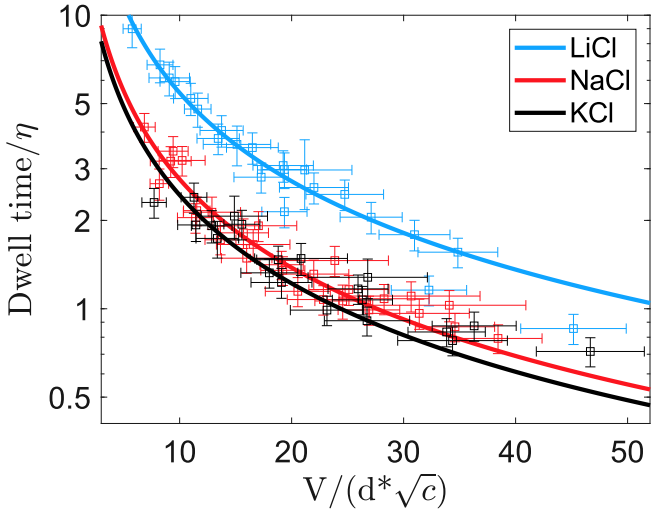


Figure 5.16: **Dwell times collapse.** Translocation dwell times of LiCl, NaCl, and KCl after the phenomenological collapse. The continuous lines correspond to fits with Eq. 5.3 to obtain the β values in Table 5.4.

Previous works combing experiments and simulations estimated a value of $\kappa = 62\ \mu m^2$ for 1 M KCl. By assuming $L^* = 150$ nm, and $\lambda_0 = -0.58$ e/bp in 1 M KCl, which are values reported in the literature for the sensing length of nanopipettes and the effective charge per length of DNA [29, 50], we can estimate κ from the β value obtained for KCl. We obtain a value of $\kappa = 65 \pm 6\ \mu m^2$ for KCl in accordance with the previous value of 62 μm^2 [61].

The effect found for the dwell times correlates with the one presented in the previous section for the current blockade. The stronger interaction between Li^+ and DNA causes a higher reduction in the mobility of Li^+ cations around DNA. However, this interaction also causes a charge reduction/friction on DNA, such that the dwell time of DNA in LiCl increases.

Table 5.4: β values obtained from fitting the scaled dwell times with Eq. 5.3 for each salt type.

	LiCl	NaCl	KCl
β	54.4 ± 1.7	27.6 ± 1.4	24.4 ± 2.1

5.6 Cation size effect on charge blockade

This section briefly shows the effects of concentration and salt type on charge blockade ($Q_B = \langle \Delta I \rangle \cdot t_D$). Figure 5.17a compares the Q_B values for translocations at 700 mV, in 1 M concentration in different salt types. It can be observed that translocations in LiCl have a larger Q_B value. The longer dwell times (t_D) in LiCl cause a larger Q_B. Translocations in NaCl show a slightly bigger Q_B than in KCl, as predicted from the previous results, where NaCl had a slightly larger t_D. Fig. 5.17b shows the effect of concentration on Q_B for KCl. Higher concentrations cause larger ΔI and longer t_D; hence, a significant increase in Q_B is observed with increasing concentration. From both panels in Fig.5.17, it can be seen that charge blockade is not dependent on the tip diameter size of nanopipettes, remaining relatively constant for a big range of diameters. This indicates that although dwell times increase with tip diameter, this effect is compensated by a smaller current blockade of larger nanopipettes so that the charge blockade remains constant.

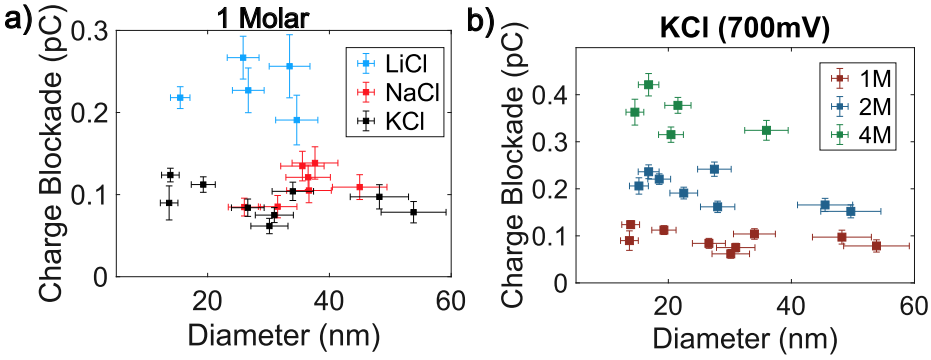


Figure 5.17: **Ion type and concentration effect on $\langle Q_B \rangle$.** a) Charge blockade ($\langle Q_B \rangle$) at 700 mV for 1 M LiCl, 1 M NaCl, and 1 M KCl as a function of the nanopipette tip diameter. b) Q_B of 1, 2, and 4 M KCl at 700 mV as a function of the tip diameter. results for NaCl and LiCl can be seen in Fig. C.2 in Appendix C.

5.7 Conclusions

In this chapter, we showed that although for a given voltage charge blockade (Q_B) is constant, this is not always true for different voltages. For the three salt types, while at 4 M, concentration Q_B is constant, at 1 M, Q_B increases with voltage. The increase of Q_B is caused by an increase of the relative current blockades, $\Delta I/I_B$, which increases with voltage.

Additionally, it was shown how the current blockade ($\Delta I/I_B$) and dwell

times (t_D) depend on the cation type used for the experiments. Translocations in LiCl have longer t_D and larger $\Delta I/I_B$. Li^+ cations, with a higher charge density, interact more strongly with DNA due to their smaller size, and therefore, they slow down DNA more effectively than Na^+ and K^+ cations. However, this effect is reciprocal, as larger $|\Delta G/G|$ are observed in LiCl, indicating that Li^+ cations suffer a higher mobility reduction when DNA translocates through the nanopipette.

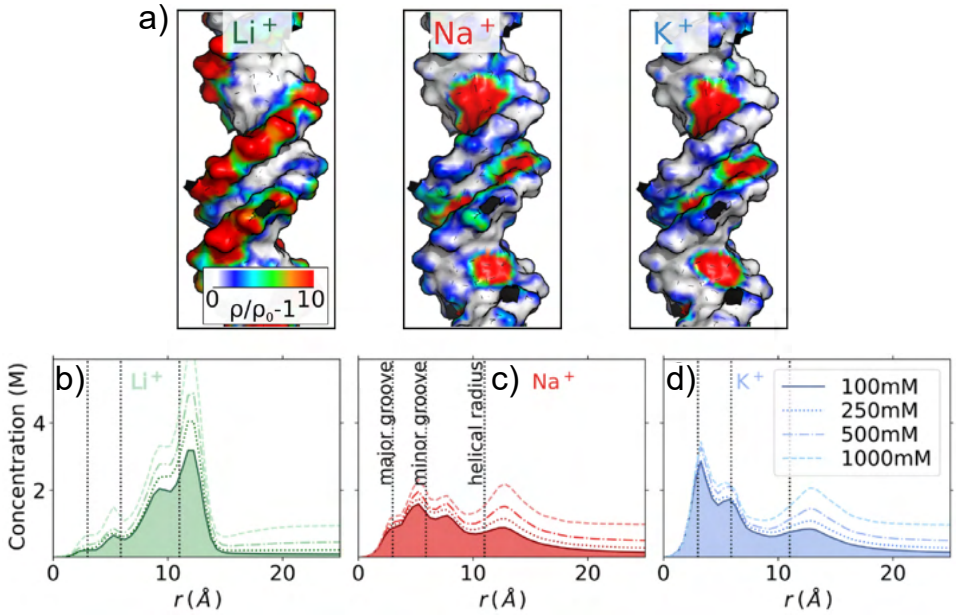


Figure 5.18: **Ion distributions around DNA.** a) three-dimensional ion distributions obtained by simulations and projected on the DNA surface. red regions indicate high cation density. b), c), and d) radial concentration profiles for 100–1000 mM bulk salt concentration.Image from [120].

The differences observed for monovalent cations in this chapter correlate with simulation results from previous works [120, 126]. In these works, it was shown that the dissociation rates of Li^+ cations from the sites at the phosphates of the DNA backbone were about one order of magnitude smaller than for Na^+ and K^+, $\sim 1\ ns^{-1}$, compared to Li^+'s, $\sim 10\ ns^{-1}$ [126]. This difference in dissociation rates and binding affinities can be observed in the ion distributions around DNA shown in Fig. 5.18a. The smaller Li^+ cation preferably binds to the DNA phosphate groups of the backbone, situated at the outside of the helix as shown in Fig.5.18a and b. In contrast, Na^+ and K^+ cations preferably bind to the major and minor grooves, indicating the importance of charge density and not only total charge in determining the complexation of ions around DNA. These effects can not be described

by traditional theoretical approaches like the Poisson-Boltzmann equation [114] and Manning's counterion condensation theory [115].

Ions play a crucial role in mediating DNA binding with proteins and RNA. They reduce the energetic costs by shielding the molecules from electrostatic repulsion. Therefore, understanding how the different cations interact with DNA and how they screen DNA negative charge is crucial to understanding biomolecule interactions.

Chapter 6

Study of λ-DNA configurations during translocation

In this chapter, we focus on analyzing λ-DNA configurations during translocation and how these configurations depend on experimental conditions such as voltage or salt concentration. We show that the initial 3D configuration of the DNA at the time the nanopipette captures the molecule determines the configuration that is measured during translocation. In the chapter's last part, we investigate the dwell time dispersion of different confined configurations.

6.1 Introduction

Let us start by accurately defining two terms of the DNA translocation process through a nanopore, which will be extensively used in this chapter. The terms are the "initial 3D configuration" (I.C.) and "confined configuration" (C.C.).

As shown in the previous chapter, translocation events have various discretized current levels corresponding to the number of DNA duplexes inside the nanopore, as shown in Fig.6.1. When talking about C.C., we will refer to the configuration that is measured during translocation, which corresponds to a sequence of these discretized current blockaded levels. The C.C. can hence be labeled using a code, as shown in Fig. 6.1. The numbers in the code correspond to the number of DNA duplexes that block the nanopipette during the translocation process Fig. 6.1. The C.C. can be subdivided into folded and unfolded configurations. For example, an unfolded confined configuration (U.C.C) occurs when only one duplex blocks the nanopipette

during the whole translocation, and it will have a 010 code (second event in Fig. 6.1). Folded confined configurations (F.C.C) occur when more than one duplex blocks the nanopore at any time during the translocation event. Therefore, the DNA is folded onto itself when translocating through the nanopipette. 0210 and 0410 (first and third events in Fig. 6.1) are examples of F.C.C, although folded configurations can have much more complex codes.

On the other hand, the term "initial 3D configuration" (I.C.) or "free configuration" refers to the 3D shape in space that the DNA molecule has just at the moment when the DNA molecule gets captured by the nanopipette to start the translocation process. These I.C. are very diverse, as seen in Fig.6.2, and depend on parameters such as the elastic properties of DNA.

It is important to note that different DNA I.C. can produce the same DNA C.C. when the DNA molecule translocates. Moreover, for a given 3D shape of DNA in space, the measured C.C. will be different depending on the position at which the DNA molecule gets captured. If the molecule has no topological knots in its 3D shape and it gets captured at one end, it will produce a U.C.C. with a 010 code. However, if it is trapped at another position, it will be an F.C.C. with a 0210 code.

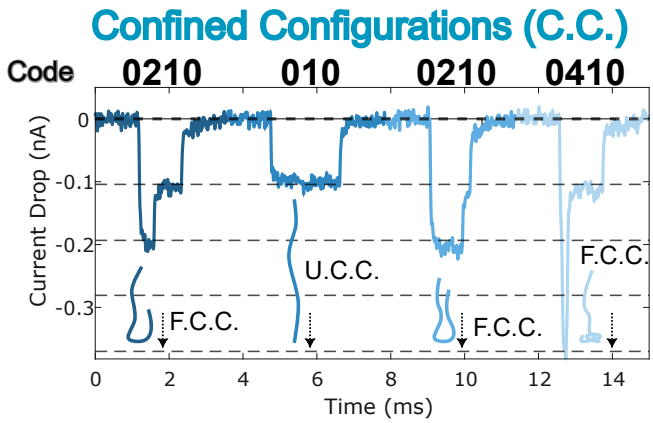


Figure 6.1: Examples of different confined configurations (C.C.) of λ-DNA during that are measured during translocation. On the top, the corresponding code for each translocation is shown.

For simplicity, when a DNA molecule is schematized during translocation, it is often given a specific shape before translocation. Examples are the linear shape or the U-shaped configuration shown in Fig. 6.1. The following section aims to provide a more realistic view of the translocation dynamics and explain why the various measured C.C. occur and how they are influenced by the I.C. of DNA molecules.

6.1.1 Dynamics of DNA translocation

Before translocation, λ-DNA molecules diffuse in the cis-chamber in a 3D space and are in equilibrium. Hence, the I.C. of the DNA molecules is well described using the elastic models such as the Worm-Like Chain(WLC) model and Freely Jointed Chain (FJC), presented in Sec. 3.1.2. Here, we simulate various I.C. of the λ-DNA molecule in 3D space using an adapted FJC model. The model is described in detail in Appendix E. Briefly, for the simulations, the λ-DNA is divided into short segments of 1 nm. The angles (ϕ and θ) from one segment to the next are randomly changed, but only small angle changes are allowed so the molecule has a smooth shape. The angle changes are weighted by the molecule's persistence length so that smaller persistence lengths (P) allow bigger angle changes. The simulations must fulfill a condition from the WLC model that the correlation between tangent vectors must follow an exponential distribution, given by $\langle \hat{t}(s) \cdot \hat{t}(0) \rangle = e^{-s/P}$, where $\hat{t}(s)$ represents the tangent vector at position s along the molecule. More details about the tangent vector correlation can be seen in Fig. E.1. Additionally, simulations of DNA molecules with different persistence lengths are shown in Fig. E.2. In Fig. 6.2, three different I.C. of λ-DNA can be observed. These I.C. have been simulated considering a persistence length of 50 nm and a contour length of 16500 nm, which are the values reported at 1 M concentration of NaCl [127].

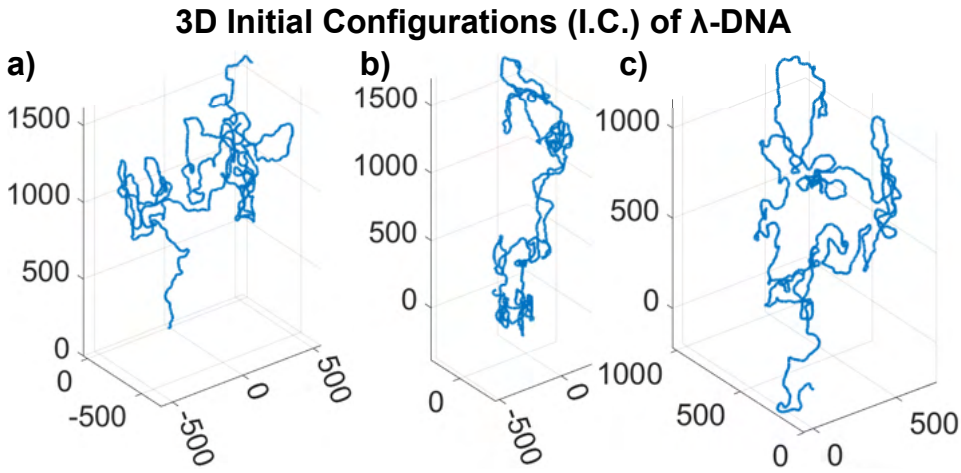


Figure 6.2: **Simulated 3D I.C. of a λ-DNA molecule.** The axes of the 3D plots are in nm. For the simulations, a persistence length of 50 nm [127] and a contour length of 16500 nm have been used.

When a segment of the λ-DNA I.C. is near the tip of the nanopipette, the electric field pulls on the DNA molecule, attracting the segment towards the nanopipette tip. Once the segment of the DNA molecule has been

captured, the electric force begins to translocate the molecule through the nanopipette. The DNA molecule undergoes a geometrical shrinkage to pass through the nanopipette, as the complete λ-DNA I.C. must pass through a confined section of about $\sim$ 30 nm at the nanopipette tip. The gyration radius (R_g) is a good measure of the size of the I.C., and it can be calculated by the equation:

$$R_g = \frac{\sqrt{\sum_i^N |r_i - \langle r \rangle|^2}}{N} \tag{6.1}$$

where r_i is the position vector of the segments, and N is the number of segments and $\langle r \rangle$ the mean position of the segments The gyration mean gyration radius of the λ-DNA molecules is about $\sim$ 450 nm, as shown in Fig.6.3a by simulating thousand λ-DNA molecules. To visualize the geometrical shrinkage, Fig. 6.3b shows the nanopipette size together with the x / y projection of an I.C. for a λ-DNA.

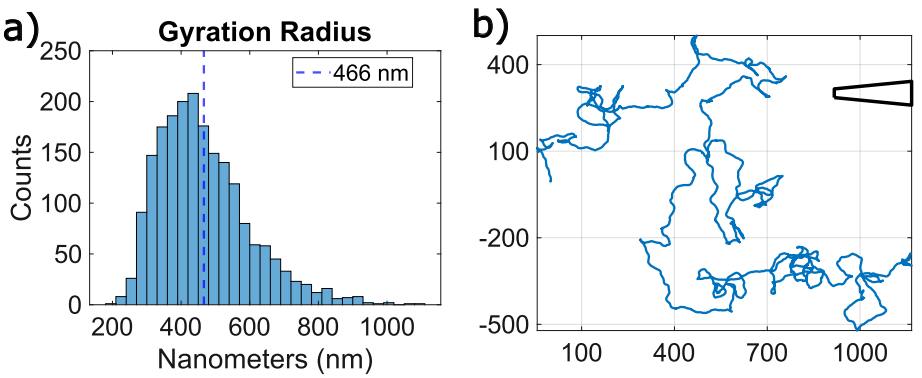


Figure 6.3: a) Gyration radius distribution of 1000 simulated λ-DNA I.C. In blue, the mean of the distribution is 466 nm. b) x/y projection of a 3D λ-DNA configuration. The axes are in nm. On the top right, the tip of a nanopipette is shown to visualize the size of the I.C. compared to the nanopipette size. The nanopipette has a diameter of 30 nm, is 250 nm long, and has a cone angle of 6°.

During translocation, the electric force constantly pulls on the DNA molecule, which is uncoiled to pass through the nanopipette. The force on the DNA is locally applied to the segment of the molecule inside the nanopipette. Therefore, when one end of the DNA falls into the pipette, it takes some time for the tension to propagate along the whole DNA molecule so that the force reaches the other end and pulls on it. This process is called tension propagation.

Two different phenomena can cause folded C.C. during translocation due to more than one DNA duplex blocking the nanopore. First, the DNA molecule can be folded onto itself while it translocates so that simultaneously, two or more DNA duplexes are inside the nanopipette. This mainly

occurs at the start of the translocation, as DNA is not always captured on one of its ends. Therefore, DNA is usually folded when the translocation starts so that two duplexes are inside the nanopipette. This folding does not occur during translocation, as the tension along the DNA molecule will unfold further folds present along the DNA molecule. The second phenomenon that can cause folded C.C. is the presence of knots in the I.C.. If topological knots are present in the DNA molecule, these knots cannot be undone by the tension propagation process and will translocate through the nanopipette. The knots will produce higher current blockades, as when knots translocate, there are multiple DNA duplexes inside the nanopipette. These knots are called equilibrium knots, as they occur when the DNA molecules are in equilibrium in solution. The occurrence of these knots is dependent on the molecule's properties, such as the persistence length and the contour length. Nanopores and nanopipettes have been used in previous works to identify different types of knots in DNA molecules [17, 128].

Timescales of dsDNA translocation

λ-DNA translocation, and generally dsDNA translocation is a non-equilibrium process [123, 129], as the dwell time is much shorter than the relaxation time of the λ-DNA molecule. Hence, λ-DNA molecules do not reorient during the relocation process. The relaxation time scale of a polymer is the Zimm relaxation (τ_Z) and is given by the equation [123]:

$$\tau_Z = \frac{0.3\eta \left(\sqrt{N} l_0\right)^3}{k_B T} \tag{6.2}$$

where η is the solution viscosity, N is the number of Kuhn segments, l_0 is the Kuhn length, k_B is the Boltzmann constant, and T is the temperature. For dsDNA, the Kuhn length is $\sim$ 80 nm. Therefore, the Zimm relaxation time of λ-DNA is $\sim$ 100 ms, whereas the dwell times, as shown in Cha.5, are in the range of 0.5-8 ms. This indicates that the λ-DNA molecules do not have time to reorient and considerably change their I.C. during the translocation process. Therefore, the initial I.C. of the DNA molecule can be considered 'frozen,' such that the main movement of the DNA molecule is caused by force applied by the electric field of the nanopipette.

6.1.2 Chapter summary

In this chapter, we start by analyzing the confined C.C. occurring during λ-DNA translocation. For the analysis, the current traces observed during DNA translocation are subdivided into levels corresponding to the number of DNA duplexes inside the nanopipette. The configuration can then

be written as a sequence of digits or a code, as shown in Fig.6.1. Each identified level will have a residence time. We then analyze the different levels' current blockades and residence times distributions and see if they depend on the applied voltage or salt concentration. The motivation is to determine whether faster translocations that occur "farther away from equilibrium" have different levels and residence time distributions due to the I.C. relaxation relocation processes.

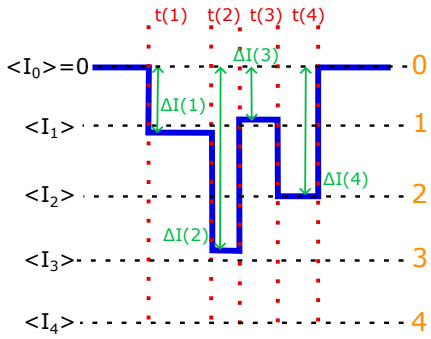
In the final section of the chapter, we present a biased diffusion model that has been used to model unfolded translations. We adapted this model to be used for any possible C.C. and studied the dwell time dispersion of configurations.

6.2 Configuration analysis

Previously, in Section 5.2, it was shown how to select traces corresponding to a single λ-DNA molecule. Here, we further analyze the λ-DNA translocation events to associate each current level during the translocation event with an integer number that indicates the number of dsDNA duplexes inside the nanopipette and the residence time of the level as schematized in Fig.6.4. The sequence of integer numbers will be called 'code' and describes the DNA C.C. during translocation. The code always starts and ends with a 0, zero duplexes inside the nanopipette, indicating that all translocation events start and then return to the baseline, as shown in Fig. 6.4. The numbers in between correspond to the sequence of DNA duplexes inside the nanopipette during the translocation process, as was shown in Fig. 6.1. For each level, the residence time and the level's current blockade are calculated, as exemplified in the table in Fig.6.4.

6.2.1 Levels detection and level assignation

A custom-made MATLAB algorithm is used to detect individual levels inside the translocation event and to calculate the level's residence time and current blockade. The algorithm also assigned a code to each translocation, as shown in Fig. 6.4. In this section, we explain the main characteristics of the level detection algorithm. The translocation events are first filtered using a digital low-pass filter and then further filtered with a box filter of 3 data points, which reduces the noise without losing much temporal resolution in our signal. The filtered signal is then shifted ± 0.06 ms, which is the time it takes for the current to jump from a level to a different level, ~ 0.06 ms. Then, the absolute value of the difference between the forward and backward-shifted signals and the non-sifted signal is computed (pink and yellow in Fig. 6.5). This difference is like 'calculating a derivative'; hence,



Code: 013120

k	1	2	3	4
Level, n(k)	1	3	1	2
Residence Time, t(k)	t(1)	t(2)	t(3)	t(4)
Blockade, ΔI(k)	ΔI(1)	ΔI(2)	ΔI(3)	ΔI(4)

Figure 6.4: **Configuration Analysis.** The current trace (in blue) is separated into different levels, depending on the current blockade **I(k)**. To each level we assign, the current blockade (**I(k)**), the residence time (**t(k)**) and an integer number (**n(k)**), that corresponds to the number of DNA duplexes. $\langle \Delta I_n \rangle$ are calculated by a multiple Gaussian fit to the current blockade histogram of all events, as explained in Sec. 5.2.

a peak in the derivative will happen when a jump between levels happens. Then, a threshold is used to detect all the peaks in the derivative over the threshold. The threshold is selected as half the current blockade of the first level and is the green horizontal line in Fig. 6.5. Once the peaks are identified for the forward and backward shifted case (blue and red asterisks in Fig. 6.5), the time at which the level jump happens is calculated by taking the mean value of the peak times. The vertical dashed lines in Fig. 6.5 are the times at which the level jumps happen. We then calculate the residence time for each level as the time between two jumps ($t(k)$ in Fig. 6.4) and the mean current blockade for each level ($\Delta I(k)$ in Fig. 6.4). With these two quantities, we fit the relocation event using vertical and horizontal lines, shown in purple in Fig. 6.5.

To compute the code for each translocation event, we need to assign an integer to each level of the translocation event. To do this, we used the level's mean current blockade, $\langle \Delta I_n \rangle$, which is calculated by doing a multiple Gaussian fit to the current blockade histogram of all events, as explained in Sec. 5.2 and shown Fig. 5.7. For each detected level, we calculate the probability of belonging to each fitted Gaussian, depending on the level's current blockade $I(k)$, by performing a Bayesian classification. The Gaussian of the multiple Gaussian fit with the highest probability determines to which level $n(k)$ the current blockade $I(k)$ is assigned.

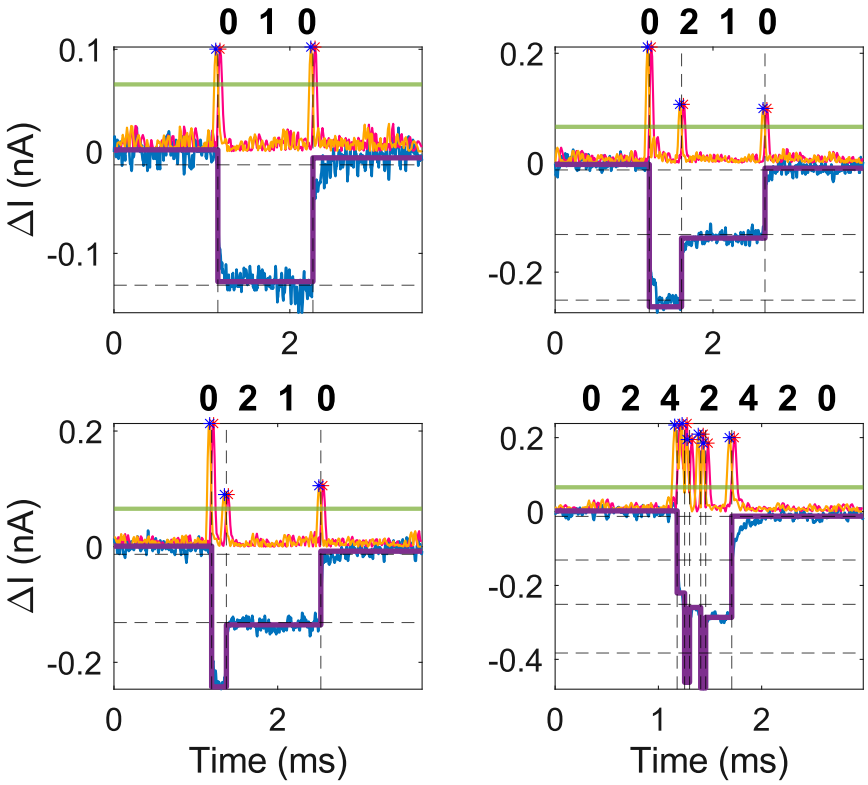


Figure 6.5: **Confined Configuration analysis**. The figure shows four randomly selected λ-DNA translocation events. In yellow and pink, the "derivatives" used to find the times when a level jump occurs are shown. In green is the threshold for detecting the peaks in the derivative. The red and blue points correspond to the found peaks over the threshold. On top of each event, the code assigned to the translocation is shown. For each event, the times at which level jumps are detected are shown with vertical dashed lines. The horizontal dashed lines show the mean current blockade, $\langle \Delta I_n \rangle$, calculated for each level by performing the multiple Gaussian fit of the histogram of all the events as explained in Fig.5.7.

After the analysis, every translocation event has been subdivided into different levels, each with a mean current blockade, a residence time, and an integer corresponding to the number of DNA duplexes, as shown in Fig.6.4. The sequence of integers assigned to each level then further defines a code for each translocation event that describes the C.C. of the λ-DNA molecule during the translocation event.

6.2.2 Code degeneracy

It is important to note that errors in assigning the levels can occur not only because of misassignments of the algorithm but also due to the limited resolution of the nanopipettes over the DNA configuration. Nanopipettes with

smaller tip diameters have a higher signal-to-noise ratio, and therefore, level assignment will be better. Although these nanopipettes will have shorter dwell times, reducing the sampling points acquired per translocation event.

In order to correctly assign a level, the DNA stands need to fill the sensing region of the nanopipette. For nanopipettes, the sensing length (L_S) is about 100-200 nm. If, for example, three duplexes of DNA translocate, but these three duplexes only fill 20 nm ($L_K = 20$ nm) of the nanopore (Fig. 6.6a), the current blockade of the level will be between levels 1 and 3 and maybe misassigned as level 2, as shown in Fig. 6.6a. In this case, the configuration 1-3-1 DNA stands will be misassigned as a 01210 code instead of a 01310. Only when the length of the knot or the DNA fold is of the order of the sensing length ($L_K \approx L_S$) will the code be correctly assigned, as shown in Fig. 6.6b. This generates a degeneracy in the code assigned to configurations since the same configuration can be assigned two different codes. Moreover, C.C. can be misassigned with a code that does not correspond to the real configuration.

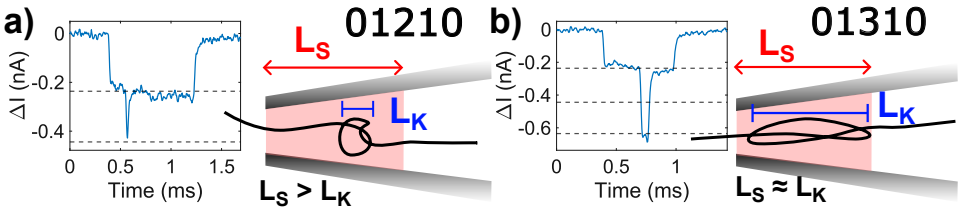


Figure 6.6: **Code Degeneracy.** Example of code degeneracy in a DNA translocation event. The same DNA C.C. can give different codes because a level is misassigned in a). The sensing region (L_S) of the nanopipette is shown in red. In blue is the knot's or fold length (L_K). The horizontal dashed lines indicate the different levels $\langle \Delta I_n \rangle$.

Misassignment detection

As shown in Fig. 6.4, level misassignment is a problem, as it can spoil our code assignment because higher levels will be classified as lower levels. However, there is an easy way to quantify how often this happens, helping us estimate how good our code assignment is. The code of a correctly assigned DNA molecule must satisfy the following statement, 'between two uneven numbers of the code, the code cannot contain an even number' (this is without counting the 0's at the start and end of the code). This statement comes from the possible geometries that DNA can adopt during translocation. If a DNA molecule has one end inside and one end outside the nanopipette, it can only achieve this by having an uneven number of duplexes inside it. This statement implies that codes such as 01210 or 03410

are misassignments, as these configurations are not possible. Important to note is that in an event such as 0140, the 4 is not between two uneven numbers, as it happens at the end of the code. To estimate the success of our level's classification, the percentage of events that do not satisfy this statement can be calculated, and this will give us an estimation of how many events were misassigned. However, it needs to be pointed out that this method is not perfect, as the misassignment of two levels in the same event can cause the misassigned event to fulfill the statement. However, we observed that the assignation of two levels inside the same event does not occur very often in our analysis.

6.2.3 Knots detection

As mentioned previously, multiple DNA duplexes inside the nanopore can occur due to the folding of the DNA onto itself or knot formation. Folding frequently happens at the start of translocations, as the nanopipette does not usually capture DNA at its ends. Because tension is applied to the DNA molecule during DNA translocation, further folds will unfold due to the tension propagation along the molecule. However, the tension cannot unknot the knots occurring in the DNA molecules, so these knots produce levels with a higher current blockade during the translocation event. These knots can be detected from the code assigned to each translocation. Knots are detected with the following rule: a DNA molecule has a knot when a section of the code assigned to each translocation (including the 0's at the start and the end) obeys the rule $n_1 k n_2$, with $k \geq \max(\mathrm{n}_1, \mathrm{n}_2) + 2$. This would correspond to a simple knot, such as 01310 (like in Fig. 6.6b) or 0150. Additionally, more complex knots can happen, which satisfy the rule $n_1 k_1 k_2 n_2$, with k_1 and $k_2 \geq \max(\mathrm{n}_1, \mathrm{n}_2) + 2$, for example, translocation events like 015310 or 03510 [17].

6.3 Effect of voltage on levels residence times and level transitions

In this section, we explore the effect of voltage and, therefore, translocation velocity on the level residence times and the transitions between the different levels. The aim is to investigate whether the higher translocation velocity at higher voltages affects the C.C. that occurs during DNA translocation and, therefore, the transition probabilities between levels.

With all the λ-DNA translocation events at a given voltage condition, a 2D-histogram of the conductance drops ($\Delta I/V$) vs. level dwell times of all the events can be plotted, Fig. 6.7. The conductance drop ($\Delta G = \Delta I/V$) is used to normalize the plots; as voltage increases, the current blockades

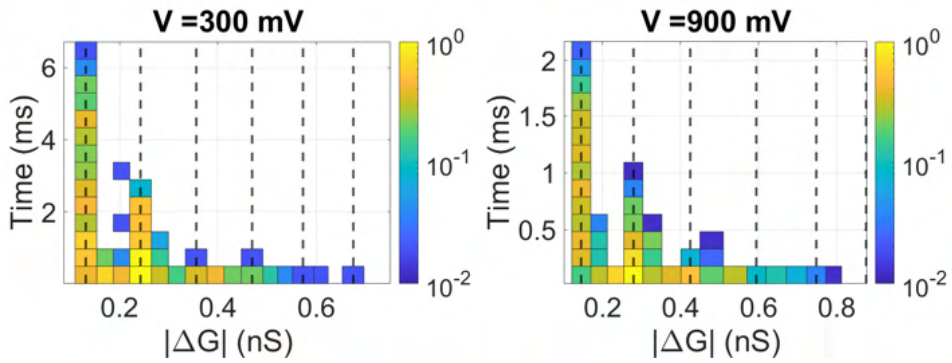


Figure 6.7: 2D-histogram of current blockade vs. dwell time of all calculated levels for a 300 mV (left) and 900 mV (right) voltage. Data are from a nanopipette with a 22 nm diameter and 2 M NaCl. The histograms are represented in a color code and have been normalized by the highest value so that the scale is from $0-1$. The color scale color is in a log scale. The mean dwell times are 3.75 ms (300 mV) and 1.18 (900 mV). 2D-histograms for 500 and 700 mV can be found in Fig.E.3 in App. E.

as was shown in Sec. 5.3. The 2D-histograms also show the mean current blockades for the levels 1 to 6, $\langle \Delta I_{1-6} \rangle$. Additionally, the counts of each histogram box have been normalized so that they go from 0 to 1 on a log scale. The 2D-histograms shown in Fig. 6.7 look similar, the only difference being in the timescale (y-axis). The slight effect of voltage on conductance drops (x-axis) was described in Sec. 5.3, where it was shown that a higher voltage slightly increases the ratio $\Delta I/I$, and therefore ΔG. However, voltage's main effect is reducing the levels' residence times. As DNA translocates faster at higher voltages, the residence times of the different levels are reduced. However, the voltage does not change the distribution of conductance drops, indicating that voltage does alter the C.C. in which DNA is translocating and that the different configurations only translocate faster at higher voltages.

The 2D histograms in Fig. 6.7 also show that the residence time of levels with higher conductance drops is shorter and that these levels are less frequent. This indicates that bigger knots involving multiple duplexes are less probable and very punctual on the DNA molecules, as they do not occupy a significant fraction of the DNA molecule dwell time.

The effect on the level's residence times can better be observed by plotting the residence time distributions corresponding to each level. This is equivalent to plotting the histogram along the vertical dashed lines in Fig. 6.7. These histograms are shown in Fig. 6.8, where it can be observed that as voltage increases, it shifts the residence time distributions of the different levels to lower times. Moreover, for a given level, the residence times are

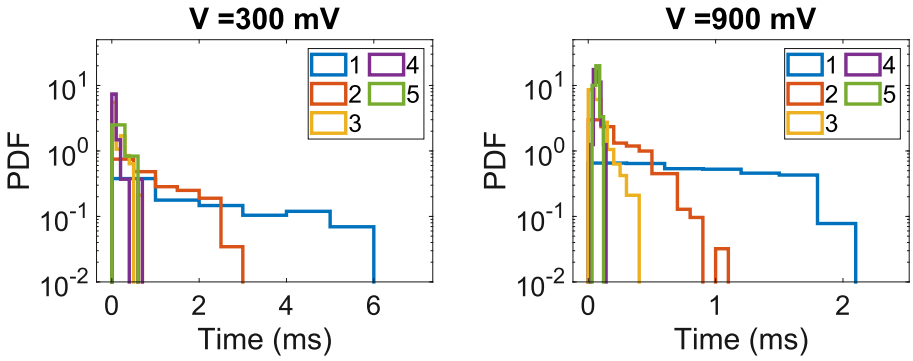


Figure 6.8: **Residence time distributions.** Residence time distributions of different levels. Data are from a nanopipette with a 22 nm diameter and 2 M NaCl. Plots for 500 and 700 mV can be found in Fig.E.4 in App. E.

always larger for the lower voltage.

In summary, the voltage does not change the C.C. in which DNA translocates through the nanopipette; it only increases their translocation velocity, reducing the residence time of the different levels.

6.3.1 Transition-level matrix

In the previous section, it was shown that voltage does not affect the conductance drop distributions and that it only changes the residence time distributions. However, voltage could have another effect. It could change the kinetics between the different levels, changing transition probabilities between levels.

The transition matrix describes the stochastic process defined by the transition probabilities between the different levels. Having assigned a code to each translocation event, one can calculate the number of transitions between the different levels. The number of transitions is then recorded in a matrix. The matrix is a $6x6$ matrix ($\boldsymbol{T}$), where the rows and columns are enumerated from 0 to 5. The numbers 0 to 5 correspond to the levels of the code. Although levels above 5 can occur, these are rare, so they are not considered for the transition matrices. The matrix's position T_{nm} contains the number of transitions from level m to level n, where n is the row number and m is the column number. The diagonal elements of the matrix T_{nn} are equal to 0, by definition, as there are no transitions between the same levels. Note that T is not a transition matrix, as it does not contain probabilities. The matrix, $\boldsymbol{T}$, which contains the number of transitions between different levels, can then be normalized by columns $\boldsymbol{T^C}$, to contain probabilities. The normalization is done by dividing all the elements of a column by the sum of all the elements of that column, such that $\sum_n T^C_{mn} =$

1. The matrix normalized by columns, $\boldsymbol{T^C}$, then contains, in each column m, the probabilities that a transition takes place from the level m, to a level $n = 0 - 5$.

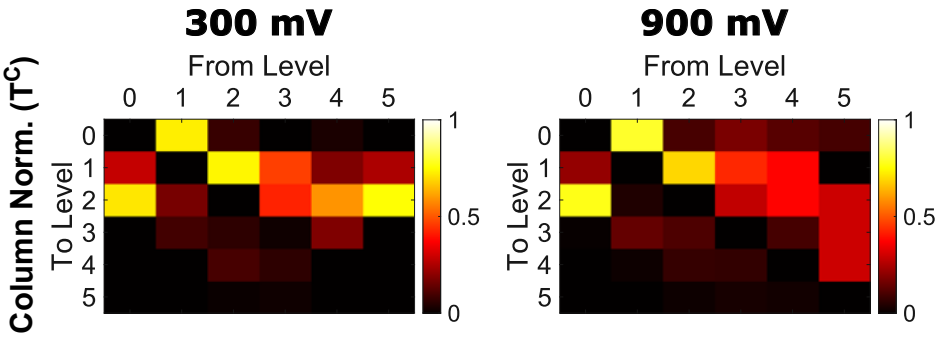


Figure 6.9: **Transition Matrix.** Transition matrix ($\boldsymbol{T^C}$) for 300 and 900 mV. Data are from a nanopipette with a 22 nm diameter and 2 M NaCl. $\boldsymbol{T^C}$ for 500 and 700 mV can be found in Fig.E.5 in App. E.

The transition matrix for 300 and 900 mV are shown in Fig.6.9. It can be observed that the matrices for 300 and 900 mV are very similar, indicating that the transitions between the same levels are equally probable for both voltages. It should be pointed out for $\boldsymbol{T^C}$ that because higher levels such as 4 and 5 do not occur frequently, these levels do not have good statistics, which is why they look different for 300 and 900 mV.

Looking at the transition matrices, several interesting facts about λ-DNA translocation can be observed:

1. The most probable way to start a translocation is by λ-DNA being folded since the matrix entry $\boldsymbol{T^C_{20}}$ is the largest of the column.

2. The most probable way to end a translocation is by having 1 strand inside the nanopipette, as $\boldsymbol{T^C_{01}}$ is the largest entry in its column.

3. When 2 DNA duplexes are inside the nanopore, the most probable is to have 1 strand after. The entry $\boldsymbol{T^C_{12}}$ is the largest of its column. Moreover, the most probable configuration when translocating is 0210 (in Sec. E.3, the probability of 0210 and 0120 events is shown).

4. While the transition from 2 to 1, $\boldsymbol{T^C_{12}}$ has a high probability, the transition from 1 to2, $\boldsymbol{T^C_{21}}$ is very improbable. This indicates that when one end of the molecule is inside the nanopipette, it rarely occurs that the other end falls into the nanopipette (in Sec. E.3, the probabilities of 0120 and 0210 can be seen).

5. The entry $\boldsymbol{T^R_{42}}$ is much more probable than the other elements in its row. Moreover, $\boldsymbol{T^C_{24}}$ has a higher probability than $\boldsymbol{T^C_{23}}$ and $\boldsymbol{T^C_{25}}$. This indicates the presence of a subpopulation of circular λ-DNA in the sample, as circular λ-DNA will have configurations that contain only even numbers, like 02420 or 0240. These will be shown in the next section.

The facts presented here are common to all voltage conditions, indicating that they are general facts about DNA translocation.

6.3.2 End-to-end time analysis

For the analyzed events, the time between the crossing of each end of the molecule across the nanopipette can be calculated using the code. We call this time the end-to-end time (t_e). Having decomposed the translocation events into levels, given by the event code, and knowing the residence time of each level, the t_e is calculated as: 't_e is the sum of the residence times of odd levels'. The end-to-end time is calculated for all translocation events that do not contain level misassignments. t_e can be better visualized in Fig. 6.10a, where we can see that only the residence time of the odd levels contributes to t_e. As we will see in this section, some interesting facts can be found by analyzing the end-to-end time distributions.

Circular λ-DNA

The λ-DNA molecule has sticky ends with two complementary overhangs of 12 bases on each end. These overhangs can hybridize and form a circular λ-DNA molecule. This does not happen frequently, as λ-DNA is a long molecule, so circularization is rare. However, fact number 5 in the previous section points to a subpopulation of circular λ-DNA. Further evidence of the presence of circular DNA in the sample can be found by looking at the end-to-end time histogram as circular λ-DNA molecules will have an end-to-end time of 0 ms, as the two ends are hybridized to each other (see the bottom configuration 0240 in Fig. 6.10a).

Fig. 6.10b shows the end-to-end time histogram (in blue) of all translocation events at 900 mV. In red, translocations with end-to-end 0 ms time are shown. The large number of short end-to-end times is caused by the events having 0 ms end-to-end time. The abnormal amount of events with 0 ms end-to-end time indicates the presence of a population of circular DNA in the sample.

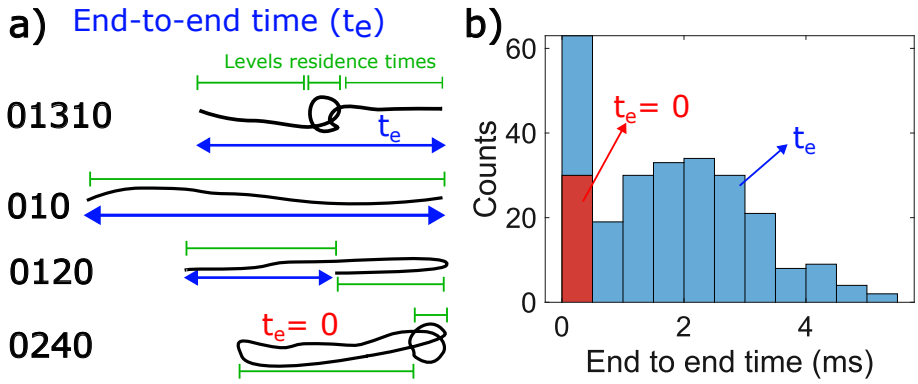


Figure 6.10: a) End-to-end time (t_e) for different DNA configurations. t_e is shown in blue for linear DNA molecules and in red for circular molecules. The residence times of the different levels are shown in green. b) End-to-end time distribution (t_e) all translocation events of a nanopipette with a 27 nm diameter, in 4 M NaCl and at 900 mV. In red, translocation events of circular molecules, $t_e = 0$, are shown.

Molecules ends

Another interesting fact can be found by looking at the histogram of the end-to-end time normalized by the dwell time (t_e/t_D). t_e is divided by its translocation time. This analysis does not consider circular λ-DNA molecules and can be easily excluded as $t_e = 0$. In Fig. 6.11a, the t_e/t_D histogram is shown in blue, while the histogram expected if all positions along the molecule were equally probable to be captured by the nanopipette is shown in red. From the histograms, it can be observed that λ-DNA is captured far more often from near its ends, as in this case $t_e/t_D \sim 1$ (as seen in Fig. 6.11b). On the contrary, if λ-DNA would be captured near its middle position, $t_e/t_D \sim 0.1$ (Fig. 6.11b).

This fact indicates the molecule ends are probably more likely to be found outside the DNA coil (the DNA I.C.) and that the ends are more mobile than other parts of the molecule; therefore, λ-DNA molecules are captured more often by its ends.

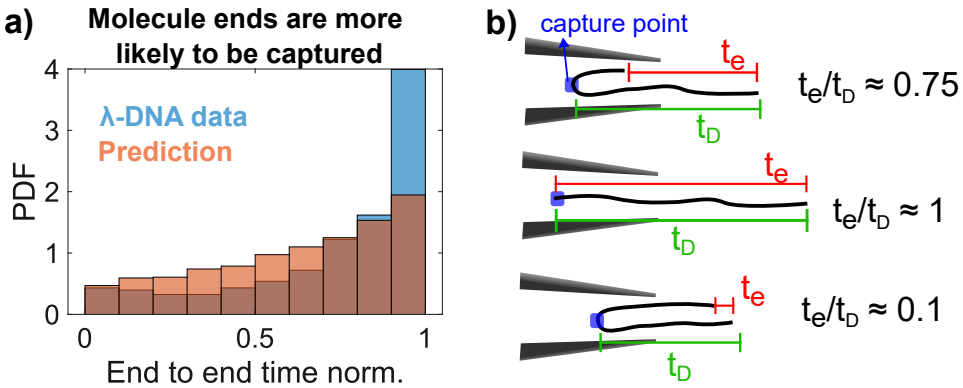


Figure 6.11: a) The normalized end-to-end time histogram, t_e/t_D, not considering circular λ-DNA molecules, is shown in blue. In red, the histogram expected if all positions along the molecule would be equally likely to be captured. b) Translocations with different folding configurations showing the t_e/t_D expected for each case.

6.4 Effect of salt concentration on confined configurations

In the previous section, it was shown that voltage does not influence the C.C. measured during translocation. We now investigate if the measured C.C. depend on the salt concentration, as salt has been shown to influence DNA elastic properties so that it could influence the initial I.C.. Three different nanopipettes with similar diameters and different salt concentrations (1, 2, and 4 M) are shown to present the effect of salt concentration. Results for more nanopipettes are shown in Appendix E. In Fig.6.12(left), a random section of 30 λ-DNA translocation events is shown for the three different concentrations. The mean current blockade $\langle \Delta I_{3-4} \rangle$ of the 3° and 4° levels are shown with blue and red horizontal lines, respectively. The percentage of events that have a current blockade bigger than the 3° and 4° levels is shown in the legend. From these plots, it can be observed that higher current blockades occur for higher salt concentrations. In particular, it can be seen that there is a significant increase at 4M (Fig.6.12c, bottom panel), in comparison with 1 and 2 M (Fig.6.12c, top and middle panels).

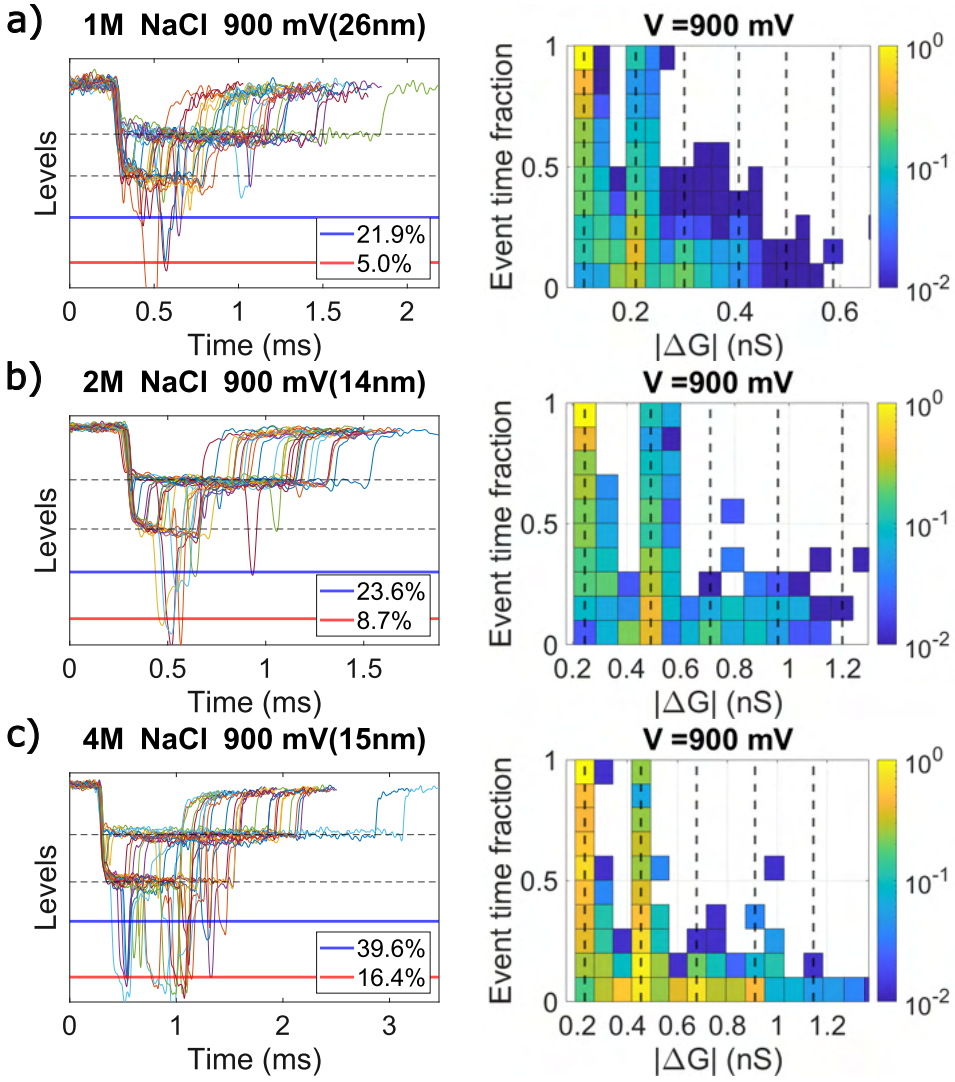


Figure 6.12: **Effect of salt concentration on confined configurations.** a) 1 M NaCl, b) 2 M NaCl and c) 4 M NaCl. On the left, 30 random translocation events are shown, indicating the mean current blockade, $\langle \Delta I_{3-4} \rangle$, of the 3$^{\text{rd}}$ and 4$^{\text{th}}$ levels shown in blue and red lines. The legend shows the percentage of events with current blockades that are larger than the 3$^{\text{rd}}$ and 4$^{\text{th}}$ levels thresholds. On the right, 2D-histograms of the conductance drop ($\Delta G = \Delta I/V$) of the levels vs. their time fraction are shown. The histograms have been normalized to range from 0 to 1 by dividing by the largest value of the histogram. The color scales are logarithmic.

From the 2D-histograms, Fig. 6.12 right, it can be observed that the counts for higher levels increase for 4 M, with respect to 1 and 2 M. In these histograms, the level's times have been normalized by the total translocation time so that the y-axis of the histograms represents the fraction of time the different levels are observed. Such a representation permits us to compare the shorter translocation events with the longer ones that occur for higher concentrations. The larger probability of higher levels at 4 M can be correlated with an increase in knot probability during DNA translocation, which is presented in Table 6.1 (more results can be seen in Appendix E, Sec. E.3). The fact that the knot probability increases with concentration indicates a direct salt effect on the 3D initial configuration. DNA's persistence length has been shown to decrease with increasing salt concentration. Consequently, the DNA's I.C. will be more densely packed at higher concentrations, which increases the probability of having equilibrium knots so that more knots are measured during λ-DNA translocation.

Table 6.1: **Knotting probability.** λ-DNA knots probability, defined as #knots)/(# of translocation events), for nanopipettes at different NaCl concentrations, at 900 mV.

	1 M NaCl	2 M NaCl	4 M NaCl
Knot prob.	0.20	0.22	0.44

6.4.1 Conclusions of voltage and concentration effects

The last sections showed how voltage does not significantly affect the configurations measured during λ-DNA translocation, while salt concentration does. This was attributed to the fact that C.C. depends on the DNA I.C. when translocation starts. Voltage does not affect the I.C., as the electric force only acts on the molecules during the translocation. Hence, voltage only increases the translocation velocity but does not affect the configurations. However, salt concentration changes the DNA environment, as increasing the salt concentration increases the screening of DNA's negative charge. Moreover, lowering the persistence length [127]. Although translocation velocity is reduced, higher current blockades are observed during λ-DNA translocation at higher salt concentrations. Moreover, the knot probability also increases with salt concentration. This indicates that salt concentration influences DNA's I.C. before translocation, as reducing the persistence length causes more density-packed I.C. with a smaller gyration radius. The more densely packed I.C. will present a higher knot probability; therefore, a larger number of higher current blockades are observed when the salt concentration is increased.

6.5 Dwell-time dispersion of different configurations

This section investigates the dwell times dispersion observed for different C.C. of λ-DNA. Fig. 6.13 shows a scatter plot of mean current blockade $\langle \Delta I \rangle$ vs. dwell time t_D of all translocation events. The continuous black line indicates the events' mean charge blockade Q_B. Events with unfolded C.C. (with 010 codes) are those with longer dwell times and lower mean current blockades. An example of a 010 configuration is shown in Fig. 6.13. On the contrary, more compact configurations have shorter dwell times and bigger current blockades. For example, a 0135310 configuration is shown in Fig. 6.13. From the scatter plot, a larger dispersion in dwell times (black bars in Fig. 6.13) can be observed for the linear configurations compared to the more compact ones. This section explores the origin of the different dwell time dispersions observed in Fig. 6.13.

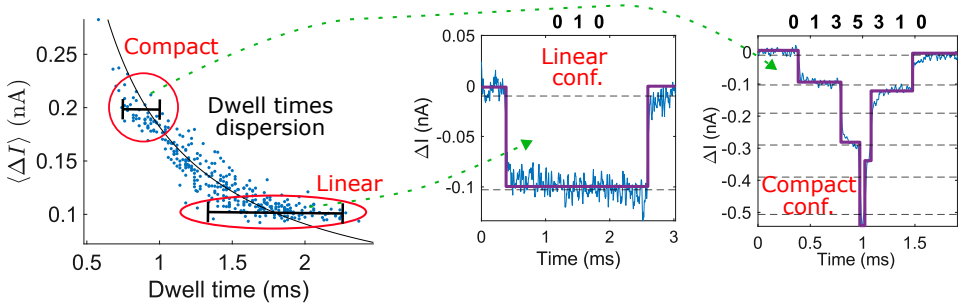


Figure 6.13: **Dwell time dispersion for different configurations.** The scatter plot on the left shows the dwell time dispersion depending on the configuration. An example of a linear and a compact configuration are shown on the right.

6.5.1 Modeling DNA translocation through nanopores

One simple model to describe DNA translocation through nanopores is to use a 1D-biased diffusion model [130]. In this model, the translocation process is modeled as the nanopore undergoing a 1D-biased Brownian motion along the DNA, in which DNA is modeled as a rigid rod with no configurational entropy, as exemplified in Fig. 6.14. Note that instead of considering the DNA's movement, we consider that that nanopore moves over the DNA, with a velocity that is given by the DNA translocation velocity. This can be done as the translocation velocity is the relative velocity between the DNA and the nanopore. The nanopore has a drift velocity (v) and a diffusion coefficient (D) and moves along the DNA molecule that has a contour length (L_c). In this model, the probability density function (PDF) $P(x,t)$ (per

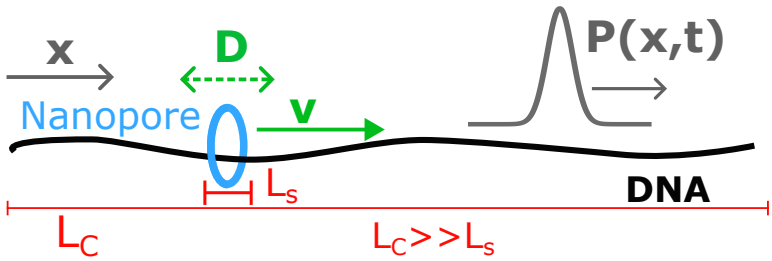


Figure 6.14: **1-D biased diffusion model**. The model considers the nanopore making a biased Brownian motion described by the velocity (v) and the diffusion constant (D) over the total length of the DNA molecule (L_c). $P(x,t)$ is the probability density function per unit length of finding the nanopore at a position (x) on the DNA at a time t.

unit length) of finding the pore at position (x) along DNA can be obtained by solving the Fokker-Planck (FP) equation, Eq. 6.3, with proper boundary conditions. The FP equation is:

$$\frac{\partial P(x,t)}{\partial t} = D\frac{\partial^2 P(x,t)}{\partial x^2} - v\frac{\partial P(x,t)}{\partial x} \tag{6.3}$$

where D and v are the diffusion constant and the drift velocity of the nanopore moving along the DNA molecule. To solve the FP equation, the boundary conditions are $P(x,0) = \delta(x)$ and $P(L,t) = 0$. The first is to assume the initial capture of DNA into the pore at $t = 0$, while the second is the absorbing boundary once the DNA has been completely translocated from the cis to the trans side [131, 68]. With these boundary conditions, the solution of the FP equation is:

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}}\left(\mathrm{e}^{-(x-vt)^2/4Dt} - A\mathrm{e}^{-(x-2L-vt)^2/4Dt}\right) \tag{6.4}$$

where $A = exp(vL/D)$. However, in translocation experiments, $P(x,t)$ is not experimentally measurable, as we lack the exact position of the nanopore on the DNA. However, from experiments, we can measure the distribution of times it takes for the nanopore to complete the 1D random walk over the DNA molecule, which is the dwell time (t_D) distribution. This distribution is given by the first-passage density function (FP-PDF) defined as:

$$F(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{L} P(x,t)\mathrm{d}x \tag{6.5}$$

$F(t)$ has the physical meaning of being the probability per unit time that the nanopore passes the adsorbing edge at $x = L$, where L is the combined length of the DNA molecule and nanopore length or sensing region $L =$

$L_c + L_s$. The result for the first-passage probability density function, $F(t)$, is:

$$F(t) = \frac{L}{\sqrt{4\pi D t^3}} \mathrm{e}^{-(L-vt)^2/4Dt} \tag{6.6}$$

In a typical translocation experiment, the molecules' contour length (L_c) is much longer than the nanopore length or sensing region, $L_c >> L_s$, of the nanopipette. For λ-DNA, the contour length is $L_c = 16.5\ \mu m$, while the sensing length of a nanopipette is approximately $L_s = 100 - 200\ nm$; therefore, we can assume $L \approx L_c$.

Having calculated the FP-PDF, one can calculate the mean dwell time ($\langle t_D \rangle$) expected for the dwell time distribution and the variance of this dwell time distribution. These are given in Eq. 6.7 and 6.9.

$$\langle t_D \rangle = \int_0^\infty t F(t) \mathrm{d}t = \frac{L}{v} \tag{6.7}$$

$$\left\langle\ t_D^2 \right\rangle = \int_0^\infty t^2 F(t) \mathrm{d}t = \left(\frac{L}{v}\right)^2 + \frac{2DL}{v^3} \tag{6.8}$$

$$\mathrm{var}\,(t_D) = \left\langle t_D^2 \right\rangle - \langle t_D \rangle^2 = \frac{2DL}{v^3} = \frac{2D\langle t_D \rangle}{v^2} \tag{6.9}$$

One has to note that this model only applies for linear/unfolded translocation, with the configuration 010, as we have considered the length to $L \approx L_c$ to be the λ-DNA contour length. However, for all the other non-linear configurations, the length of the molecule is smaller than the contour length (L_c), as the molecule is folded or contains knots. These configurations will have a shorter length, and in the next sections, we will see how to overcome this so that the biased diffusion model can be applied to all C.C. beyond the linear case.

Effective diffusion

Using the fluctuation-dissipation theorem, we can estimate a value for the diffusion coefficient (D) in our λ-DNA experiments. The fluctuation-dissipation theorem is given by:

$$D = k_B T/\gamma \tag{6.10}$$

where γ is the friction coefficient of DNA when translocating through the nanopore. Recalling the model presented in Chapter 5, during DNA translocation, the electrophoretic and drag forces balance out $F_E + F_D = 0$, such that translocation occurs at a constant velocity. The driving force is given by $F_E = E\lambda l_s$, where l_s the sensing length of the nanopipette where the field is concentrated, and λ the DNA effective charge per length. The electric field at the tip of the nanopipette is $E = 2Vtan(\alpha)/d$, where V is

the applied voltage, d is the tip diameter, and α is the cone angle of the nanopipette. The driving electrophoretic force is opposed by the drag force $F_D = -\gamma v$. Hence, the fluctuation-dissipation theorem can be rewritten as:

$$D = \frac{k_B T v}{F_E} = \frac{k_B T v d}{2V tan(\alpha) \lambda l_s} \tag{6.11}$$

Considering a sensing length $l_s = 100\ nm$, a voltage $V = 500$ mV, a DNA effective charge $\lambda = 0.4$ e/bp, and a nanopipette parameters $d = 30$ nm and $\alpha = 3^o$, the electrophoretic force can be estimated to be $F_E \approx 12.6$ pN. Using a typical velocity for λ-DNA translocation $v = 4\ \mu$m/ms, the diffusion constant can be predicted using Eq. 6.11, and is $D \approx 1.3 \cdot 10^{-3} \mu m^2/ms$. Considering this value for the diffusion constant, for a λ-DNA molecule of 16.5 μm, using Eq. 6.9, the dwell time variance for linear translocation events would be $\text{var}\,(t_D) \approx 10^{-3}$ ms^2. However, experimentally, the measured variance for linear translocations is orders of magnitude larger $\text{var}\,(t_D) \sim 1$ ms^2.

The significantly larger variance of the experimental data points to an additional mechanism, apart from diffusion, that broadens the dwell-time distribution. Several works have shown fluctuations in the translocation velocity of different DNA molecules [132, 133]. The fluctuations are due to DNA molecules' different I.C. before translocation, as shown in Fig. 6.15. As shown in Fig. 6.15a, I.C. where the DNA is, on average, nearer to the nanopipette (number 1) translocate faster and therefore have shorter dwell times, as seen in the distribution in Fig. 6.15b. On the contrary, I.C., where DNA starts stretched and, on average, further away from the nanopipette (number 3), has longer dwell times.

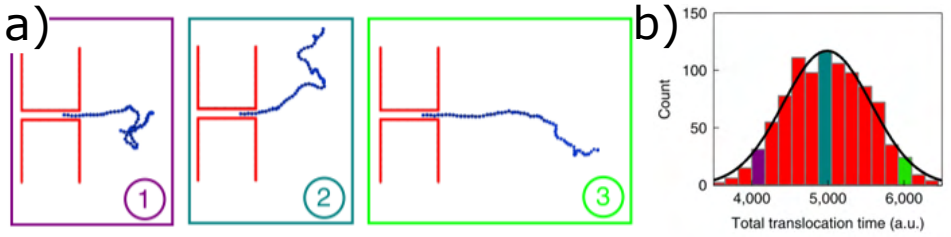


Figure 6.15: **Mean translocation velocity fluctuations caused by the I.C.** a) Three different initial I.C. for a DNA translocation. While in 1, the mean distance from the DNA molecule to the nanopore is small, in 2 and 3, it is larger. b) Dwell time distribution, obtained by simulating linear translocations with different I.C. The distribution shows the translocation time obtained for the I.C. 1, 2, and 3. Images from [123]

The fluctuations of the translocation velocity of different molecules broaden the dwell-time distributions. It is important to note that the dwell time vari-

ance is not caused only by the diffusion constant (D) but also by fluctuations in the molecule's velocities. The previously presented model considers a constant drift velocity. Therefore, if Eq. 6.9 is used to estimate the diffusion coefficient from the dwell times variances, the diffusion constant will be significantly overestimated.

Hence, from now on, we will consider an effective diffusion (D_{eff}). This effective diffusion gathers the effects of velocity fluctuations and the "real" diffusion constant.

6.5.2 Configuration dependent effective diffusion

In this section, we adapt the 1D-diffusion model to be used for all translocation events, independent of their C.C., and not only for linear translocations. We then use the adapted model to investigate the dependence of the effective diffusion for different configurations.

The 'unique' characteristic of linear translocations is that we know that the nanopore has performed the random walk over the whole molecule length, which is the molecule's contour length (L_c). For other C.C., we do not know the length of the molecule. This length will be shorter than the contour length as the molecule is translocating in a more compact manner. This is exemplified in Fig. 6.16, where it can be observed that the length over which the nanopore has to translocate (L_{mol}) is dependent on the configuration. For the case of more compact configurations, like 020 or 02130, the length L_{mol} is shorter. To adapt the model to all possible configurations, we need a way to compute L_{mol} to assign a length to non-linear configurations.

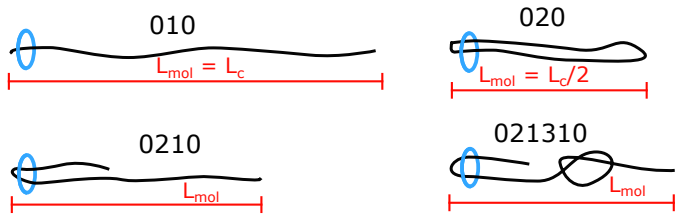


Figure 6.16: **Molecules effective length.** Schematic representation of four different configurations, showing how the effective length (L_{mol}) of the molecule is shorter for more compact configurations.

The analysis to calculate $\langle t_D \rangle$, var(t_D) and L_{mol} is described now. Figure 6.17a shows translocation events at different voltages on a $\langle \Delta I \rangle$ vs. t_D scatter plot. For each voltage, the translocation events are grouped by their $\langle \Delta I \rangle$ value, as shown in Fig. 6.17b. Each group contains a $\sim 9\%$ of all the translocations events, and there is a $\sim 4.5\%$ overlap between the groups so that there is a total of 20 groups. Only 10 groups are shown in Fig. 6.17b for

clarity, as the groups have overlap. For each of these groups, represented with the index i, we calculate the mean dwell time $\langle t_D \rangle_i$ and the mean current blockade $\langle \Delta I \rangle_i$, represented by the bigger colored squares in Fig. 6.17b. Additionally, the dwell time variance of each group is calculated $\mathrm{var}(t_D)_i$. These values are then used to calculate the drift velocity (v) and the effective diffusion, using an adaptation of the 1D biased diffusion model presented in Sec. 6.5.1.

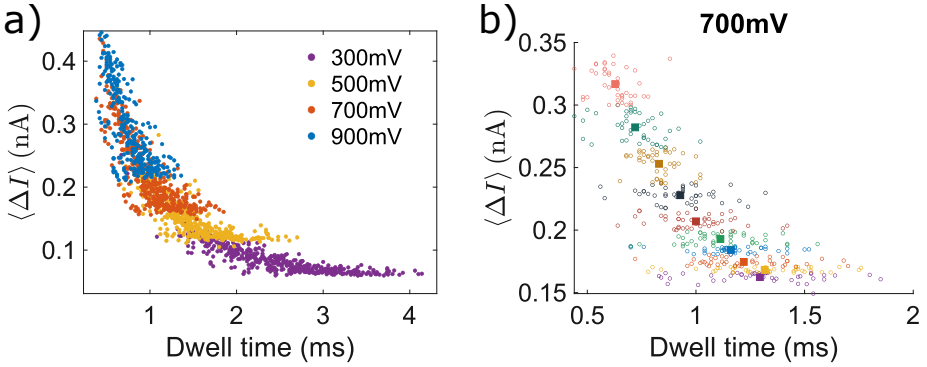


Figure 6.17: a) Scatter of mean current blockade ($\langle \Delta I \rangle$) vs. dwell time (t_D) for four different voltage conditions. b) Separation in groups for the 700 mV condition to calculate $\langle t_D \rangle_i$, $\langle \Delta I \rangle_i$ and $\mathrm{var}(t_D)_i$. Only 10 of the total 20 groups are shown. The big colored squares represent the calculated $\langle t_D \rangle_i$ and $\langle \Delta I \rangle_i$ of each group.

Drift velocity

In Fig. 6.18a, the $\langle t_D \rangle_i$ vs. $\langle \Delta I \rangle_i$ for the different voltages are shown. To compute the molecule's length for each of the groups $L_{\mathrm{mol,i}}$, the following formula is used, which assumes that molecules with the same current blockade ($\langle \Delta I \rangle_i$) have the same length:

$$L_{\mathrm{mol,i}} = L_c \cdot \frac{\langle \Delta I \rangle_{\mathrm{lin}}}{\langle \Delta I \rangle_i} \tag{6.12}$$

where $\langle \Delta I \rangle_{\mathrm{lin}}$ is the current blockade of the linear molecules (i.e. the 010 configuration). For these molecules, the length is equal to the contour length ($L_{\mathrm{mol}} = L_c$). The length, L_{mol} of molecules with a current blockade twice as big, for example, 020 configurations, the L_{mol} will be half the contour length $L_{\mathrm{mol}} = L_c/2$. As we work with λ-DNA, the contour length is 16.5 μm.

Figure 6.18b shows the molecule length, $L_{\mathrm{mol,i}}$ vs. the mean dwell time, $\langle t_D \rangle_i$, of each group. The plot shows that molecules with a shorter effective length, resulting from more compact configurations, have shorter dwell times. For each voltage, a fit is performed to calculate the drift velocity (v),

straight lines in Fig. 6.18b, using the following equitation:

$$L_{\text{mol,i}} = \langle t_D \rangle_i \cdot v \tag{6.13}$$

this is Eq. 6.7, but considering different molecule lengths. The points of a given voltage in Fig. 6.18b follow a straight line, indicating that drift velocity is independent of the confined configuration (C.C.).

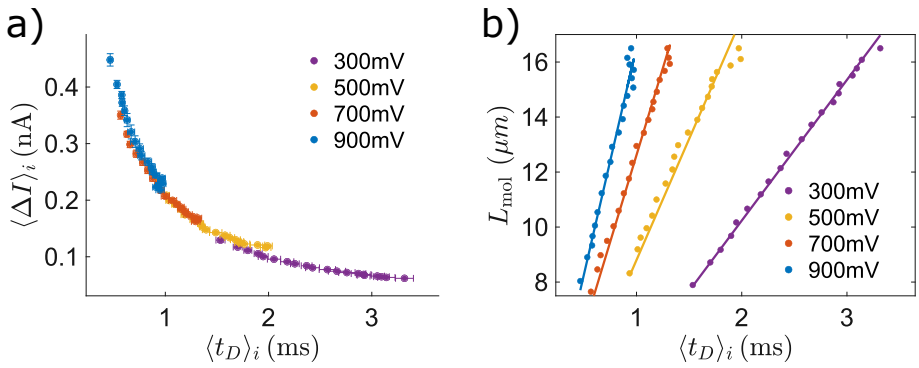


Figure 6.18: a) Plot showing the $\langle t_D \rangle_i$ vs. $\langle \Delta I \rangle_i$ of the four voltage conditions. b) $\langle t_D \rangle_i$ vs. the length of each group $L_{\text{mol,i}}$. Solid lines are fits with Eq. 6.13 to calculate the drift velocity (v).

Effective diffusion

The dwell time variance $\text{var}(t_D)_i$ vs. the mean dwell time are plotted in $\langle t_D \rangle_i$ Fig. 6.19a, where it can be observed that the dwell time variance ($\text{var}(t_D)_i$) does not follow a linear relation with the mean dwell time. According to Eq. 6.9, the relation would be linear if the diffusion constant, D, is independent of the C.C., as is the case for v. However, the fluctuation-dissipation theorem, Eq. 6.11, states that diffusion constant (D) is half if the electrophoretic force (F_E) is double. This would be the case for 010 and 020 configurations because, for the 020 configuration, the electrophoretic force is double, as two duplexes are inside the nanopore at any time. Hence, the diffusion constant is half that of the 010 configuration ($D_{020} = D_{010}/2$), as $D = \frac{k_B T v}{F_E}$ (Eq. 6.11). Therefore, we define a configuration-dependent diffusion constant of the form:

$$D_{\text{eff}} = \left(\frac{\langle \Delta I \rangle_{\text{lin}}}{\langle \Delta I \rangle_i}\right)^{\beta} D_{\text{lin}} \tag{6.14}$$

where D_{lin} is the effective diffusion of linear configurations. The exponent β accounts for the reduction of D_{eff} for the more compact configurations. $\beta \approx 1$ is what would be expected from the fluctuation-dissipation theorem, but

the velocity fluctuations that also contribute to D_{eff} are also configuration dependent, β should be bigger than 1 ($\beta > 1$). The relation between the variance ($\text{var}(t_D)_i$) and the mean dwell time ($\langle t_D \rangle_i$) is given by:

$$\text{var}\,(t_D)_i = \frac{2D_{\text{eff}} \cdot \langle t_D \rangle_i}{v^2} = \frac{2(\frac{\langle \Delta I \rangle_{lin}}{\langle \Delta I \rangle_i})^{\beta} D_{\text{lin}} \cdot \langle t_D \rangle_i}{v^2} \tag{6.15}$$

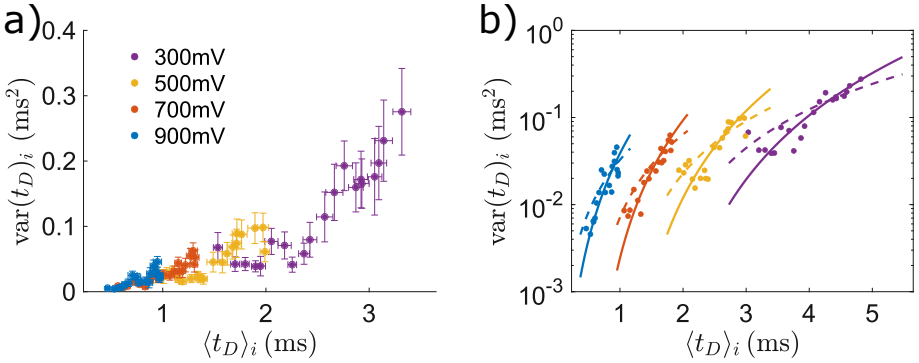


Figure 6.19: a) Plot showing $\text{var}(t_D)_i$ vs. $\langle t_D \rangle_i$. The error bars are computed using the bootstrapping method. b) $\text{var}(t_D)_i$ vs. $\langle t_D \rangle_i$ plot, with the y axis on a log scale. For clarity, the data of 700, 500, and 300 mV have been shifted 0.5, 1, and 1.5 ms to the right, so they do not overlap. Lines represent global fits with Eq. 6.15. The continuous line with β as a free parameter but the same β for the four voltage conditions. The dashed lines are fits with $\beta = 1$.

Using this equation, we can perform a fit to the variance data. For the fit, we take the velocity (v) calculated for each voltage in the previous fits and leave D_{lin} and β as free parameters. We perform a global fit to all the voltages such that the value of β is the same for all the voltages. The fitted data are shown in Fig. 6.19b on a log scale. Additionally, for clarity, the data for 700, 500, and 300 mV have been shifted 0.5, 1, and 1.5 ms to the right so that there is no overlap between the different voltages. Two fits are performed, one imposing $\beta = 1$ (dashed lines in Fig. 6.19b) and one with β as a free parameter (continuous lines). The fit with beta as a free parameter reproduces much better the data, as a bigger β value is required to reproduce the smaller variance for more compact configurations. For β, we obtain the value $\beta = 2.3 \pm 0.4$.

Finally, in Fig. 6.20a, we show the values obtained for the drift velocity (v) as a function of voltage. A linear dependence with voltage can be observed for the drift velocities. Figure 6.20b shows the values obtained for the effective diffusion for the different voltages as a function of the mean dwell time ($\langle t_D \rangle_i$). The dashed lines are for the case of $\beta = 1$, while the continuous lines are for $\beta = 2.3$. As observed from Fig. 6.20b, the effective

diffusion for linear configurations is the biggest and then decreases for more compact configurations. Moreover, D_{lin} increases with voltage. It is unclear what causes this increase in the diffusion constant, although some previous works have related it to the larger osmotic flow caused by higher voltages [131].

In summary, in this section, we have seen that the effective diffusion constant decreases with an exponent β, larger than expected from the fluctuation-dissipation theorem. This stronger decrease is probably because more compact confined configurations have smaller velocity fluctuations than extended (or linear) configurations. The different velocity fluctuations for different C.C. will be discussed in the next section. Moreover, the drift velocity was shown to be independent of the configuration and to have a linear dependence on voltage, which correlates with the previous results in this work since dwell times were shown to be proportional to $1/V$ in Chapter 5.

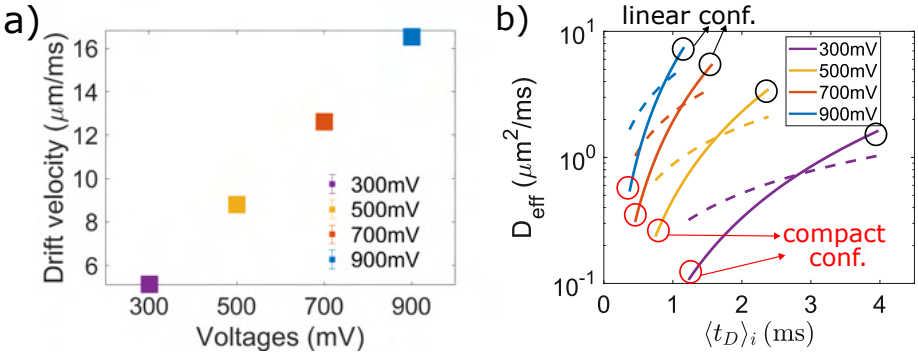


Figure 6.20: a) Dependence of the drift velocities (v), obtained from the fits with Eq. 6.13, with voltage. b) Effective diffusion constant as a function of the mean dwell time $\langle t_D \rangle_i$ for the different voltage conditions. Dashed lines are for the case of $\beta = 1$, while continuous lines correspond to $\beta = 2.3$.

6.5.3 Origin of configuration dependent velocity fluctuations

The previous section showed that the effective diffusion increases with an exponent β, larger than expected by the fluctuation-dissipation theorem. As shown in previous works, velocity fluctuations during DNA translocation are caused by fluctuations in the initial configuration of DNA [132, 133]. In these works, it was shown that I.C., where DNA is, on average, further away from the nanopipette when the translocation starts, has a slower translocation velocity, as was shown in Fig. 6.15. In this section, we show how more compact configurations, which correspond to translocations with a larger current blockade, have less dispersion in the mean average distance to the

nanopipette. This fact contributes to a decrease in the effective diffusion for more compact C.C., as was observed in the previous section.

To investigate this, we simulate different λ-DNA I.C. using the FJC model presented in Sec. 6.1.1 of this chapter. More information about the simulations can be found in Appendix E. For each of the simulated initial configuration (I.C.), the mean distance per segment (d_{conf}) to the tip of a nanopipette is calculated. The mean distance is given by:

$$d_{\text{conf}} = \frac{\sum_N^i \left(\left| \overrightarrow{d_i} - \overrightarrow{d_p} \right| \right)}{N} \tag{6.16}$$

where $\overrightarrow{d_p}$ is the vector that defines the position of the nanopipette tip, and $\overrightarrow{d_i}$ is the vector that defines the position of each segment i, as shown in Fig. 6.21a. Hence, $\overrightarrow{d_i} - \overrightarrow{d_p}$ is the vector from each segment to the nanopipette tip (Fig. 6.21a).

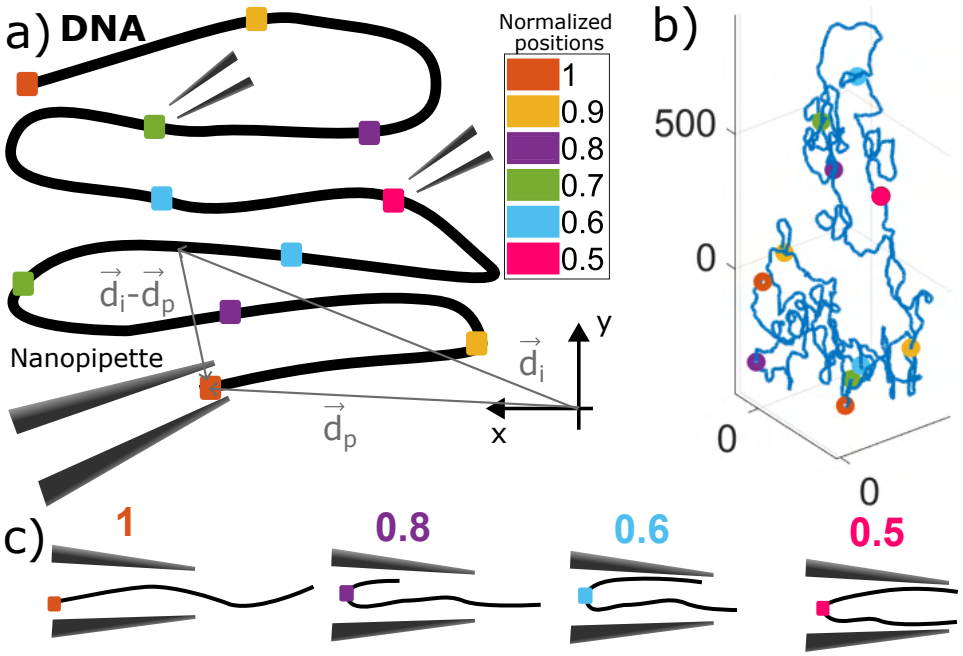


Figure 6.21: **λ-DNA simulations.** a) Scheme of the simulations. For each simulated DNA molecule, the nanopipette is placed at six positions along the molecule, which are given by the normalized positions 0.5 to 1. Positions 0-0.4 are symmetric with respect to 1-0.6. b) Simulated I.C. showing the normalized positions 0-1. c) Configurations when the DNA molecules are captured along different normalized positions 1, 0.8, 0.6, and 0.5.

For each simulated I.C., the nanopipette is placed at six positions ($\overrightarrow{d_p}$) along the DNA molecule's contour length, which is represented by the nor-

malized positions 0.5 to 1. The normalized position is the position along the molecule contour length where the nanopipette captures the molecule, divided by the molecule contour length. Hence, positions 0 and 1, 0.9 and 0.1, etc, are symmetric; therefore, they are represented by the same color in Fig. 6.21a. Because of the symmetry, d_{conf} is only computed from positions 0.5 to 1. In Fig. 6.21b, the positions on a simulated λ-DNA 3D I.C. are shown.

The position of the nanopipette along the DNA I.C. will cause different C.C., as can be seen in Fig. 6.21c. If we do not consider the presence of knots, if the nanopipette captures the molecule on its ends (normalized positions 0 or 1), it will be a linear translocation (a configuration 010). However, if the nanopipette captures the molecule in the middle (normalized position 0.5), the translocation will be a 020 configuration, with a larger current blockade than the 010 translocation. Therefore, translocations, where λ-DNA is captured by the nanopipette at positions 0.5-0.7, will produce more compact configurations during translocation, while if captured at positions 0.8-1, the configurations will be more linear.

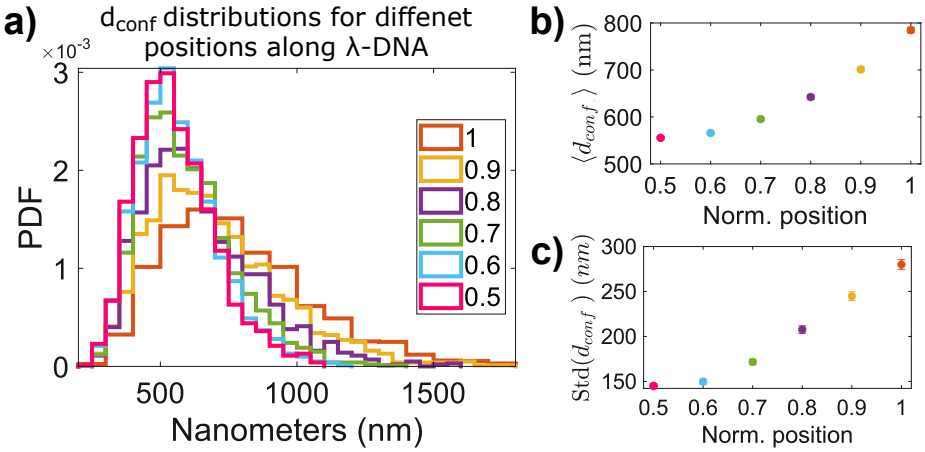


Figure 6.22: a) d_{conf} distributions obtained from simulations of 2000 λ-DNA molecules, modeled by the FJC model 3.1.2. d_{conf} is calculated for the six different normalized positions 0.5 to 1 along the molecule (see Fig. 6.21a). b) Mean value of the distributions in a). c) Standard deviation (Std) of the distributions in a).

The results from the simulations of different I.C. are shown in Fig. 6.22. In panel, a, the distributions of the mean distance per segment to the nanopipette tip (d_{conf}) are shown for the different normalized positions. It can be observed how when the DNA molecule is captured closer to the middle point (position 0.5), the d_{conf} distributions are narrower and have a smaller mean value. The mean value and standard deviation of the dis-

tributions are shown in Fig. 6.22 b and c as a function of the normalized position. These results indicate that the mean distances to the nanopipette tip and their dispersion are smaller for I.C. where the DNA is captured near the middle position. However, these I.C. lead to more compact C.C. during translocation, with higher current blockades, as observed in Fig. 6.21c. Therefore, more compact configurations will have smaller velocity fluctuations due to the smaller mean distance to the nanopipette. Hence, the effective diffusion will be smaller for the more compact configurations, which correlates with the decrease in effective diffusion observed in the previous section for more compact configurations.

6.5.4 Simulating dwell-times distributions using FP

To confirm the hypothesis smaller velocity fluctuations for more compact configurations are causing the decrease in the effective diffusion, we performed some simulations using the 1D-biased diffusion model presented in Sec. 6.5.1. However, instead of using a diffusion coefficient of about 0.1-10 μm^2/ms like was obtained for the effective diffusion (see Fig. 6.20b), we used the small diffusion of 0.0001 μm^2/ms predicted by the fluctuation-dissipation theorem. Instead, we allow for velocity fluctuations in the model, and we try to reproduce the var$(t_D)_i$ vs. $\langle t_D \rangle_i$ data in Fig. 6.19.

A summary of the conditions used in the simulation is shown in Fig. 6.23. To perform the simulations, we need three parameters: the diffusion constant D, the drift velocity v, and the length of the molecules L, depending on their configuration (as was shown in Fig. 6.16). These parameters are obtained in the following manner:

1. **Length (L).** As we have seen in the previous sections, the effective length of the molecules depends on their configuration during translocation (Fig. 6.16). To obtain the distribution of lengths for each voltage condition, we use the equation:

$$\ell_{\text{norm}} = \frac{L_{\text{mol}}}{L_c} = \frac{\langle \Delta I \rangle_{\text{lin}}}{\langle \Delta I \rangle} \tag{6.17}$$

   where $\langle \Delta I \rangle$ is the mean current blockade of each translocation event and $\langle \Delta I \rangle_{\text{lin}}$ is the mean current blockade of linear translocations. The histogram of ℓ_{norm} distributions for 700 mV is shown in Fig. 6.23a. For each voltage, the length value ℓ_{norm} is sampled from this histogram. The length of the molecule used in the simulations is then $L_{\text{mol}} = \ell_{\text{norm}} \cdot 16.5$ μm.

2. **Diffusion constant (D).** For the diffusion constant, we use a value as predicted from the fluctuation-dissipation theorem, 0.0001 μm^2/ms,

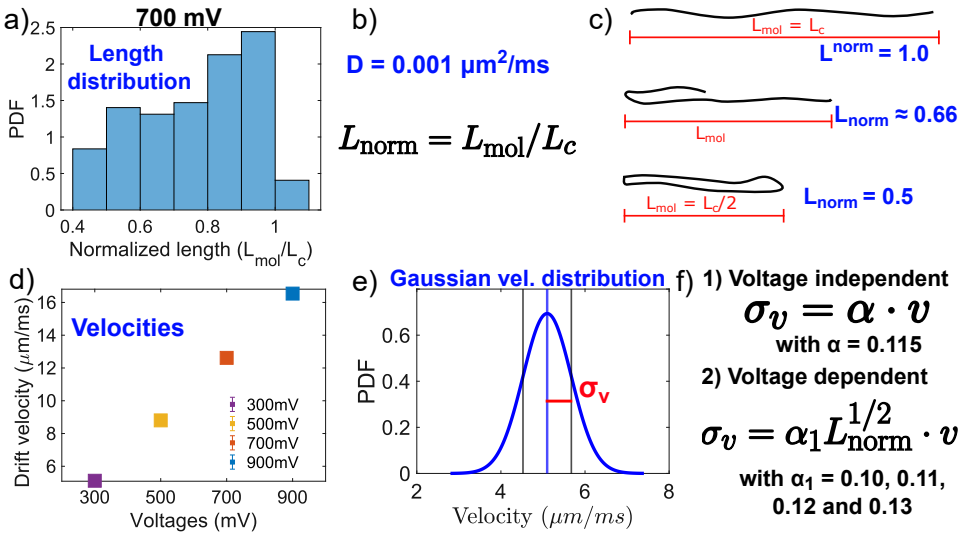


Figure 6.23: **Simulations with 1D biased diffusion model.** The figure shows how the three parameters v, D and $L_{\mathrm{mol}} = \ell_{\mathrm{norm}} \cdot 16.5$ μm are obtained. For the simulations, v and ℓ_{norm} are then obtained from the corresponding probability density functions (PDF), and then the dwell time is computed by the PDF given by Eq. 6.18.

as shown in Fig. 6.23b. As this is a small value, the contribution to the effective diffusion will come from the velocity fluctuations.

3. **Drift velocity (V).** For the velocity, we consider the drift velocities obtained in the previous analysis, Fig. 6.23d. We then consider a Gaussian distribution for the velocities, with a standard deviation σ_v (see Fig. 6.23e). Two cases are considered for the velocity fluctuations, models by σ_v (Fig. 6.23f). First, a voltage and configuration independent case, where the velocity fluctuations are given by $\sigma_v = \alpha \cdot v$. Second, a case where the velocity fluctuations depend on the voltage and the configuration, and therefore depend on L_{mol}. In this case, the $\sigma_v = \alpha_1 L_{\mathrm{mol}}^{1/2} \cdot v$, with $\alpha_1 = 0.10$, 0.11, 0,12, and 0.13 for 300, 500, 700 and 900 mV respectively. In this case, the velocity fluctuations increase with voltage. However, they decrease for more compact configurations (with a smaller L_{mol}), as σ_v is proportional to $L_{\mathrm{mol}}^{1/2}$. The α and α_1 parameters are chosen so that the simulations reproduce the data in Fig. 6.25.

To perform the simulations, we used:

$$F(t) = \frac{L_{\mathrm{mol}}}{\sqrt{4\pi D t^3}} \mathrm{e}^{-(L_{\mathrm{mol}} - vt)^2/4Dt} \tag{6.18}$$

where v and L_{mol} are picked from the corresponding density functions distributions as described and shown in Fig. 6.23, and then the dwell time is computed by the PDF given by Eq. 6.18. 10000 translocations are simulated with the 1D-biased diffusion model.

The simulated dwell-time distributions for the different values of ℓ_{norm} considered are shown in Fig. 6.25a for the case of 700 mV. The sum of all these distributions makes up the total dwell time distribution for that voltage is shown in black in Fig. 6.25a. In Fig. 6.25b, the simulated distributions (continuous lines) are compared with the experimental data (histograms). It can be observed that the simulations with the 1D biased diffusion model considering different lengths ℓ_{norm} and velocity fluctuations reproduce the experimental dwell time distributions well.

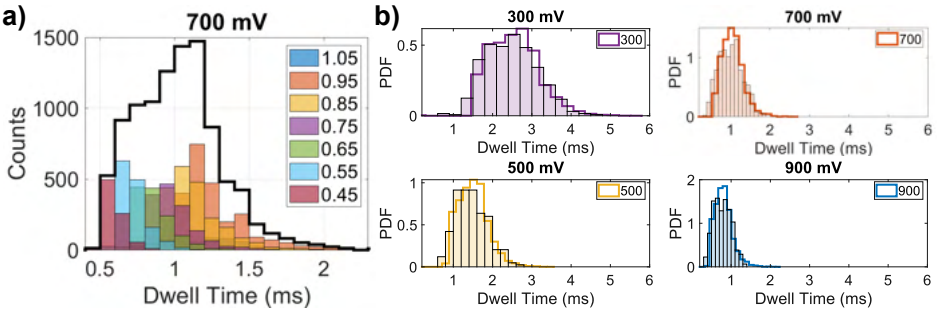


Figure 6.24: a) Dwell time distributions obtained for the different ℓ_{norm} used in the simulation. The total distribution, which is the sum of all the others, is shown with a black continuous line. b) Experimental dwell time histograms and histograms obtained by the simulation (continuous lines) for the different voltages. The simulation results shown here are for the case of voltage-dependent velocity fluctuations ($\sigma_v = \alpha_1 L_{\text{mol}}^{1/2} \cdot v$).

Finally, from the distributions obtained for the seven ℓ_{norm} values used in the simulation, shown in Fig. 6.25a, we compute the variance, var(t_D) and the mean dwell $\langle t_D \rangle$ for each ℓ_{norm} value. These are then compared to the experimentally obtained values in Fig. 6.25. The voltage and configuration-dependent velocity fluctuation case are shown in Fig. 6.25a, while the case of voltage and configuration-independent velocity fluctuations is shown in Fig. 6.25b. It can be observed that only when velocity fluctuation depends on voltage and on the configuration can the simulations reproduce the experimentally observed dwell time variances. This is in accordance with the results of the previous sections, where it was observed that the effective diffusion was larger for higher voltages. The larger velocity fluctuations at higher voltages would cause this. Moreover, it was also shown that more compact configurations have less dispersion in their I.C. Hence, the velocity fluctuations are smaller for these configurations.

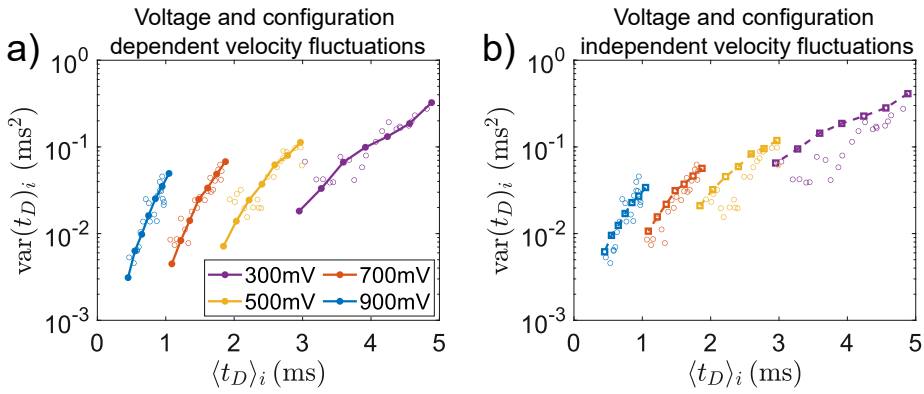


Figure 6.25: a)

In this section, it has been shown how considering a Gaussian distribution of the mean translocation velocities with a standard deviation ($\sigma_v = \alpha_1 L_{\mathrm{mol}}^{1/2} \cdot v$), we can satisfactorily reproduce the experimental dwell times distributions and also the experimentally observed dwell time variance for different configurations. In order to correctly reproduce the experimental data, we need to consider on one side, velocity fluctuations that increase with voltage, with different α_v values depending on the voltage. On the other side, we consider a decrease in σ_v proportional to $L_{\mathrm{mol}}^{1/2}$ for more compact configurations (configurations with smaller ℓ_{norm}).

6.6 Conclusions

In the first section of the chapter, a more realistic view of DNA translocation was discussed. Using the FJC model, different DNA I.C. for a λ-DNA molecules were simulated to illustrate their size, given by the gyration radius and the dispersion in gyration radius between the various 3D initial configurations.

A levels analysis was presented, by which each translocation event is divided into different levels with their corresponding residence times, current blockades, and number of DNA duplexes inside the nanopipette. Thanks to this analysis, it was shown that voltage does not affect the level distributions and the transition probabilities between the levels. The only effect of voltage is to reduce the residence times in the levels, as it increases the translocation velocity. Furthermore, by comparing data at different salt concentrations, it was observed that higher concentrations increase the occurrence of higher levels, which was correlated with a higher knot probability for higher concentrations. The conclusion of this analysis is that the C.C. measured during translocations are not affected by voltage, as they

depend on the DNA I.C. when translocation starts. As voltage does not affect the DNA when it is diffusing in bulk, voltage does not affect the initial 3D configurations. However, salt concentration affects the I.C., as DNA persistence length depends on salt concentration. Therefore, higher salt concentrations with shorter persistence lengths cause more densely packed 3D initial configurations. In the more packed I.C., the DNA chains are closer to each other, which increases the knot probability. This indicates that salt concentration influences the number of equilibrium knots present in long DNA molecules in salt solutions.

In the last section of the chapter, the dwell-times dispersion for different C.C. was investigated. A 1D-bias diffusion model was presented, which was previously used to model DNA translocation through nanopores [131]. It was discussed why the velocity fluctuation influences the model's diffusion constant; hence, the term effective diffusion is more appropriate. It was then shown how the effective diffusion of more compact C.C. decreases with an exponent β, larger than expected from the fluctuation-dissipation theorem. To investigate the effect of configurations on the effective diffusion, simulations of λ-DNA with the FJC model were performed. It was found that more compact configurations, which occur when DNA gets captured closer to the middle position, have a smaller average distance to the nanopipette, and the average distance distributions have a lower dispersion. Finally, we simulated 1D-bias diffusion considering velocity fluctuations and different molecule lengths. With the simulations, we were able to reproduce the experimental dwell time distributions and the dwell time variances.

Part III

DNA UNZIPPING WITH OPTICAL TWEEZERS

Chapter 7

Unzipping experiments of a long DNA hairpin in monovalent salts

This chapter presents force spectroscopy experiments of DNA in high monovalent salt concentrations. During the experiments, the two strands forming the DNA helix are continuously pulled apart by applying a mechanical force at the opposite ends of the molecule (see Fig. 7.1a), converting the duplex into a ssDNA strand. The process of separating the two hybridized strands is called unzipping, in analogy to a zipper. The two strands are connected by a small loop of ssDNA, forming a hairpin structure to prevent disengagement of the strands. The applied force is reduced to reform the hairpin so that the DNA strands hybridize again. This process is called the rezipping of the DNA hairpin. The mechanical unzipping experiments are performed on the mini-tweezers setup described in Chapter 2.

Optical tweezers is a single-molecule technique that provides complementary information to the DNA translocation experiments across nanopipettes presented in the previous chapters. The ability to apply and measure forces allows us to extract information about the molecule's stability, which is not accessible with relocation experiments. As presented in Chapter 5, DNA dwell times depend on the cation type/size. Here, we investigate if the cation type and size also influence the unzipping force of a DNA hairpin. The unzipping force of DNA hairpins has been shown to depend on the salt environment. Higher salt concentrations shield the negative charge of the DNA phosphate backbone more effectively, stabilizing the DNA hairpin so that a higher force is required to unzip it. In this chapter, we investigate how the cation type and the salt concentration influence the hairpin stability, correlating with the results in Chapter 5, where the dwell times were

shown to be larger for smaller cations. Finally, the unzipping data obtained in this chapter are discussed and compared with the results obtained in the previous Chapters 5 and 6.

7.1 Experimental Protocols

Mechanical unzipping experiments are performed in the mini-tweezers setup described in Chapter 2. For the experiments, a 3594 bp DNA hairpin is used. More details about the synthesis can be found in the Appendix B. Briefly, the hairpin is synthesized by digesting a λ-DNA with EcoRI restriction enzyme (New England Biolabs). The corresponding $\sim$3.5 kbp DNA fragment is then incubated with custom-designed oligos to form a DNA hairpin flanked by two 29 bp short DNA linkers with biotin- and digoxigenin-labeled opposite ends. To carry out the experiments, the molecular construct containing the DNA hairpin and the linkers is tethered between an AD bead and a streptavidin (SA) bead held on the tip of a micropipette by air suction (see Fig.7.1a). To achieve this, before the experiment, the hairpin construct is incubated with anti-digoxigenin (AD) beads so that the DNA hairpins bind to the AD beads through digoxigenin/anti-digoxigenin bonds. During the experiment, the two beads are brought close to each other so that the biotin-labeled end of the hairpin binds to the SA bead. The SA bead is held on the tip of a micropipette by air suction. The micropipette has a tip diameter of about a micron and should not be confused with the nanopipettes used for the electrical measurements.

The hairpin is mechanically unzipped and rezipped by repeatedly moving the optical trap in a controlled manner, as described in the pulling protocol in Sec.2.2.5. At least 10 unzipping/rezipping cycles are acquired for each molecule. For each salt concentration, a minimum of 5 molecules are measured. Experiments are performed at different concentrations of monovalent salt, NaCl, KCl, and LiCl. The buffer for the experiments is a 10 mM Tris buffer with 10 mM EDTA at a pH = 7.5. The ionic strength of this buffer is $\approx$ 28 mM of monovalent salt. The pH is lower than that of the translocation experiments in nanopipettes. We chose a lower pH because, from some test experiments, the pH seemed to be affecting digoxigenin/anti-digoxigenin or the biotin/streptavidin bonds, weakening them. Hence, the hairpin would undergo fewer pulling cycles before one tether broke. However, the change in pH does not affect DNA's charge. The phosphate groups of DNA have a very low pKa value of about pKa $\sim$ 2. Therefore, the oxygen (O) atoms of the DNA backbone's phosphate groups are fully dissociated at pH 7.5 and 9.0. The typical unzipping (blue) and rezipping (red) curves obtained during experiments are shown in Fig.7.1b. In these curves, the force applied on the DNA hairpin is plotted versus the distance moved by the optical trap.

It can be observed that the unzipping and rezipping processes happen at an approximately constant force (black dashed line) and have a characteristic sawtooth pattern. This pattern occurs because DNA base pairs open/close sequentially in groups of ∼10-100 bp.

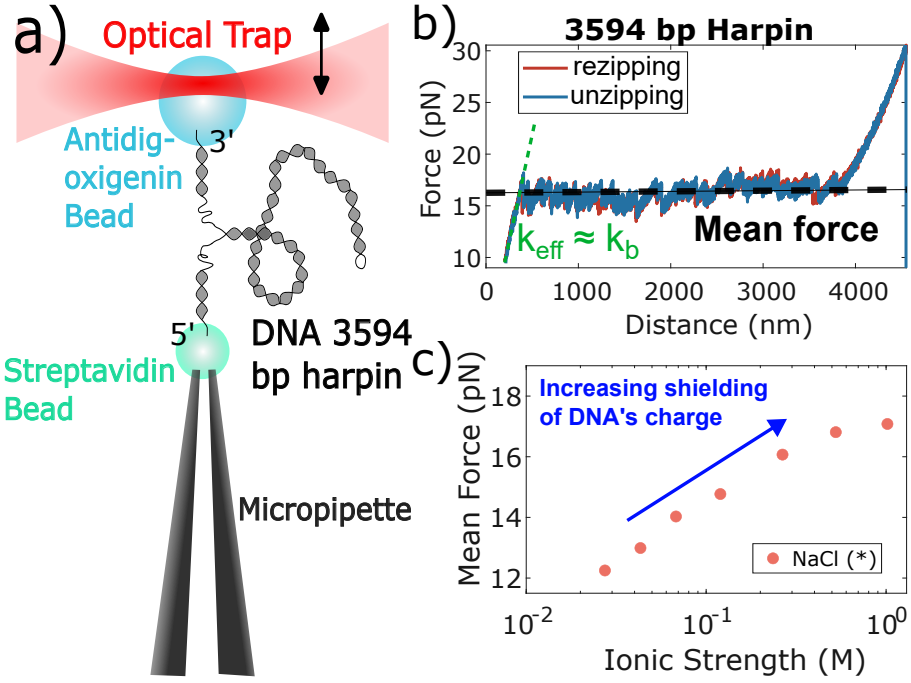


Figure 7.1: **DNA Hairpin unzipping.** a) Schematic representation of the unzipping of a long DNA hairpin. b) Unzipping and rezipping curves of a 3594 bp hairpin.c) Mean unzipping force plotted as a function of ionic strength for NaCl. Data are from previous works in our group [134].

Previous works have shown that in the range of 10 mM to 1 M NaCl, the mean unzipping and rezipping force increases with salt concentration [134] (Fig.7.1c). The force increase is due to the screening of DNA's negative charge. The two DNA strands repel each other because of their negative charge. At low salt concentrations, the repulsion is stronger, destabilizing the DNA hairpin. Hence, the force required to unzip/rezip is lower. As salt concentration increases, the shielding of the negative charge stabilizes the hairpin; therefore, the unzipping/rezipping force is higher. In Fig.7.1c, the force is plotted vs the ionic strength rather than the salt concentration. This is because, at low salt concentrations, one must account for the ionic strength of the ions of the buffer, the Tris, the EDTA, and the ions used to correct the pH. The buffer ions will contribute equally to shielding DNA's charge. Hence, the ionic strength rather than the salt concentration, is the relevant quantity in Fig. 7.1c.

Distance to molecular extension conversion

During unzipping/rezipping experiments, the distance from the micropipette to the trap (λ, shown in Fig. 7.2) is moved at a constant velocity, $\lambda(t) = \lambda_0 + vt$. The distance, λ, is the control parameter in the experiments and is measured by the position PSDs. The distance is first increased during the unzipping of the DNA hairpin and then decreased during the rezipping at the same speed, $\lambda(t) = \lambda_0 - vt$. As the trap distance increases (decreases), the force (f) exerted on the molecule increases (decreases), making the molecule unzip (rezip). The total distance λ between the micropipette and the center of the trap can be written as:

$$\lambda = x_b(f) + x_h(f) + x_d(f) + x_{ss}(f) \tag{7.1}$$

with $x_b(f)$, the displacement of the bead from the center of the optical trap; $x_h(f) = x_{h1}(f) + x_{h2}(f)$, the elongation of the handles of dsDNA; $x_{ss}(f)$, the extension of the ssDNA of the unzipped part of the hairpin; and $x_d(f)$, the average extension of hairpin's stem, that is defined as the extension of the hairpin stem projected along the force axis, being equal to the DNA-helix diameter ($\sim$ 2 nm). All the distances are shown in Fig.7.2. When the distance λ changes, it produces an extension variation in the force-dependent terms in Eq. 7.1. A small variation in extension ($\delta\lambda$) produces a variation in force (δf). The relation between these variations is the effective stiffness ($k_{eff} = \delta f/\delta\lambda$) of the system. The effective stiffens (k_{eff}) can be computed considering the stiffness of all the elements in the system (springs in Fig. 7.2). As all the springs are connected in series, the effective stiffness is given by:

$$\frac{1}{k_{eff}(f)} = \frac{1}{k_b(f)} + \frac{1}{k_h(f)} + \frac{1}{k_{ss}(f)} + \frac{1}{k_d(f)} \tag{7.2}$$

where $k_b(f)$ is the stiffness of the optical trap; $k_h(f)$ is the stiffness of the dsDNA handles; $k_d(f)$ is the stiffness of the hairpin stem; and $k_{ss}(f)$ is the stiffness of the single strands released during the unzipping process. The bead in the optical trap can be modeled by Hookes law, with a constant stiffness $f(x_b) = k_b x_b$).

During experiments with the mini-tweezers, the force-distance curve (FDC) shown in Fig. 7.1b is acquired. In order to compute the force-extension curve (FEC) corresponding to the extension of the ssDNA upon unzipping, the effective stiffness method is used [136]. In experiments with short dsDNA handles, the effective stiffness before the hairpin starts to unzip can be approximated by $k_{eff}^{-1} \approx k_b^{-1}$, i.e., other components can be neglected. In fact, the stiffness of the short dsDNA handles (29 bp) and the hairpin stem are large so that $k_h^{-1} \approx k_d^{-1} \approx 0$. Moreover, if force varies in a

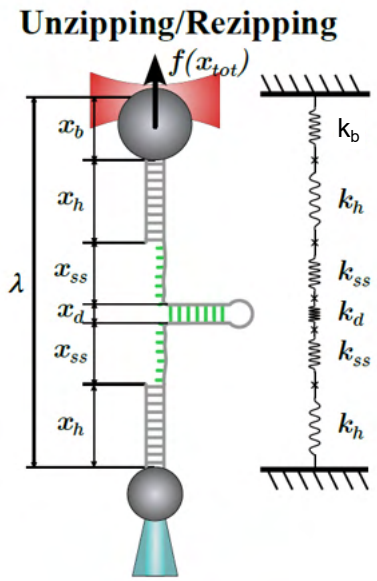


Figure 7.2: **Unzipping experiment at controlled extension**. By moving the optical trap, a force is applied to the molecular construct (details in text). The total trap-pipette distance is the sum of several contributions: the distance of the bead to the center of the trap (x_b), the extension of the dsDNA handles (x_h), the extension of the ssDNA of the unfolded molecule (x_{ss}), and the extension of the molecular construct (x_d). Image from [135].

small range during the experiments, the trap stiffness (k_b) remains constant. Therefore, the trap stiffness can be determined by performing a linear fit to the FDC before the first rip $f = k_b \cdot x_b$ (shown in Fig.7.1b, in green color). Where x_b is the displacement of the bead from the center of the optical trap. Note that before the first rip, the only contribution to λ comes from x_b. The FEC can then be obtained by computing $\lambda - x_b = \lambda - f/k_b$, where $\lambda - x_b$ is the extension of the DNA molecule. The analysis in this chapter will be performed on the FEC.

7.2 Unzipping experiments in monovalent salts

In order to compare with the results obtained in Chapter 5 for DNA translocation in LiCl, NaCl, and KCl, we carry out unzipping experiments at the same salt conditions (1, 2, and 4 M). The mean force (f_m) for each concentration is calculated by averaging the mean force of all the molecules measured at that condition. Fig. 7.3a shows the mean force for the different salts and concentrations. The f_m required to unzip the hairpin in LiCl

is higher at the three studied concentrations than NaCl and KCl, where the mean force is the same. The f_m in LiCl is $\sim$ 1 pN higher at 1 and 2 M, with respect to NaCl and KCl, while the difference increases to $\sim$ 2 pN at 4 M. An example of unzipping curves at 1 M for the three salts is shown in Fig.7.3b. For clarity, raw data was filtered using a box-car average to better appreciate the differences between the three salts. It can be observed that although the sawtooth pattern remains similar, unzipping/rezipping occurs at a higher force. This indicates that the Li^+ cations are more effective at shielding off DNA's negative charge, and therefore, the hairpin is more stable in solutions containing LiCl. Hence, a higher force is required to unzip the hairpin. The results correlate with the results from DNA translocation in Chapter 5, where the effect of Li^+ cations of dwell times was significantly bigger than Na^+ and K^+, which show very similar dwell times. The same effect is observed here for the mean force (Fig. 7.3a), with LiCl showing a different behavior than NaCl and KCl.

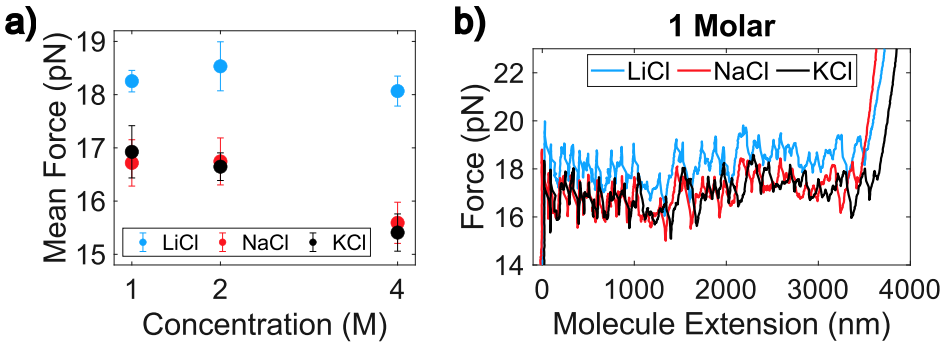


Figure 7.3: a) Mean unzipping/rezipping force for 1, 2, and 4 M of three monovalent salts, LiCl, NaCl, and KCl. Note that at such high salt concentrations, the salt concentration is approximately equivalent to the ionic strength. b) Unzipping curves at 1 M for the three salts.

7.2.1 Unzipping over a broad range of concentrations

The mean unzipping force of a DNA hairpin is determined by several factors, such as the stacking interaction between the bases, the base pairing, and the electrostatic repulsion between the DNA strands [137]. All these effects contribute to the free energy of base pairing. The free energies are described by the Nearest-Neighbor (NN) model and have been shown to increase for increasing ionic strength in the 10 mM - 1 M range in NaCl [79]. B-DNA is the most stable structure in this range; hence, the f_m increases as the electrostatic screening increases with salt concentration. The increase in f_m occurs because hydrogen bonding and stacking between bases are not

affected by salt concentration [137]; therefore, only electrostatic interactions determine f_m.

Surprisingly, the mean force in Fig. 7.3a does not follow the trend observed for NaCl in the salt range of 10 mM to 1 M (Fig.7.1). Here, a decrease in f_m is observed at the higher concentrations, especially for NaCl and KCl. This decrease is almost unappreciated for LiCl in the 1-4 M range, which seems to have a constant f_m. To better understand the trend observed at high salt concentrations, we take advantage of the high solubility of LiCl in water, $\sim$ 19 M at 25° [54]. The solubility of NaCl and KCl is much lower, $\sim$ 6 M and $\sim$ 4.5 M, respectively [54]. Therefore, we do unzipping experiments at 8 M of LiCl, and also at lower concentrations to confirm that LiCl follows the same trend as NaCl at lower concentrations. The results are shown in Fig.7.4a. The ionic strength is used instead of the salt concentration to account for the buffer ionic strength at low concentrations.

In Fig.7.4, it can be observed how f_m increases until 1 M for NaCl and KCl, and until 2 M for LiCl, and decreases for higher concentrations. A big drop of $\sim$ 4 pN is observed in the f_m when going from 4 M to 8 M LiCl (Fig.7.4a). In Fig.7.4b, the unzipping curves at 10 mM, 1 M, and 8 M are shown. The sawtooth pattern remains similar when concentration is changed. However, the mean force at 8 M decreases to a force value similar to that observed at 10 mM. The large decrease in f_m is the same behavior observed in KCl and NaCl for 4 M. The decrease in f_m at high concentrations indicates a destabilization of DNA due to the high salt concentrations, which will be discussed in the next sections.

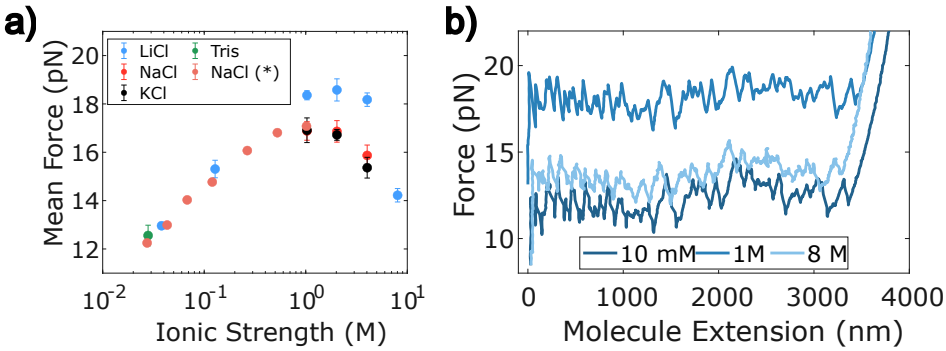


Figure 7.4: a) Mean unzipping/rezipping force plotted as a function of ionic strength for different concentrations of LiCl, KCl, and NaCl. The green points correspond to unzipping experiments performed in the 10 mM Tris and 10 mM EDTA buffer, with an ionic strength of $\sim$ 28 mM. $NaCl(*)$ data are from previous works in our group [134]. b) Unzipping traces at 10 mM, 1 M, and 8 M of LiCl.

Mean-force difference between AT and GC rich regions

Looking in detail at the unzipping curves (Fig.7.4b), it can be observed that the unzipping force is not constant over the hole molecule extension, the exact force is sequence dependent. Therefore the mean force unzipping of different regions is slightly different. This is caused by sequence-dependent effects, as base paring free energies are sequence-dependent. According to the Nearest-Neighbor (NN) model, base pair motifs containing G and C bases have free energies twice as big as motifs containing A and T [79, 134].

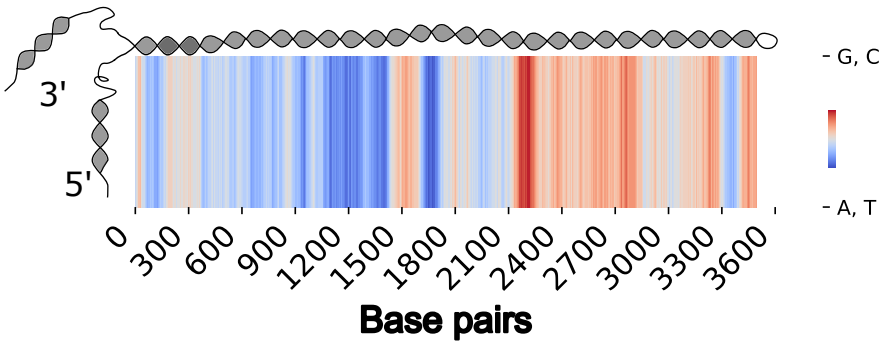


Figure 7.5: **Heat map of the 3594 bp hairpin.** The heat map shows regions with higher GC (red) or AT (blue) content, indicating the relative density of GC and AT base pairs along the hairpin.

The sequence of the hairpin used for the experiments contains regions rich in AT base pairs and regions rich in GC base pairs. A heat map of the hairpin is shown in Fig.7.5, where the regions containing a higher density of AT base paring are shown in more intense blue color, while regions with a higher density of GC base paring are shown in more intense red colors.

The AT/GC-rich regions can be correlated with parts of the unzipping curves. Figure 7.6 shows that AT-rich regions (in blue) unzip/rezip at a slightly lower force than GC-rich regions (in red). The regions correspond to a molecular extension between 500 and 1100 nm for the AT-rich region and between 2000 nm and 2600 nm for the GC-rich region. From the traces shown in Fig.7.6a and b, a decrease from 1.3 to 0.9 pN in the force difference can be observed at 8 M with respect to 10 mM concentration.

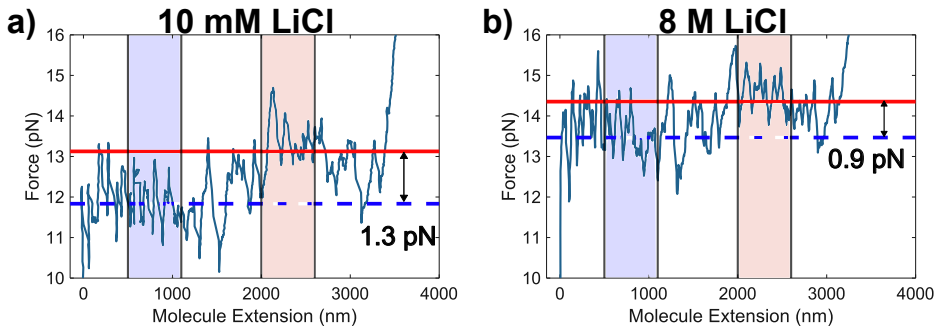


Figure 7.6: **Unzipping curves at high and low salt concentrations**. Unzipping curves at 10 mM a) and 8 M b) of LiCl are shown. The region between 500-1100 nm, shown in blue, is selected to compute the mean force of the AT-rich region (blue dashed line). The region between 2000-2600 nm extension (red) is selected to compute the mean force of the GC-rich region (red line). The force difference shown as a vertical black arrow corresponds to the mean force difference value over all the cycles acquired for the molecule.

To study the effect of concentration on the force difference between GC and AT-rich regions, the mean force difference ($\Delta F = F_{GC} - F_{AT}$) is computed at each condition. Fig.7.7 shows the force difference (ΔF) between the GC and AT region as a function of the ionic strength. ΔF decreases when ionic strength is increased. A decrease of ~ 0.4 pN is observed for the higher salt concentrations.

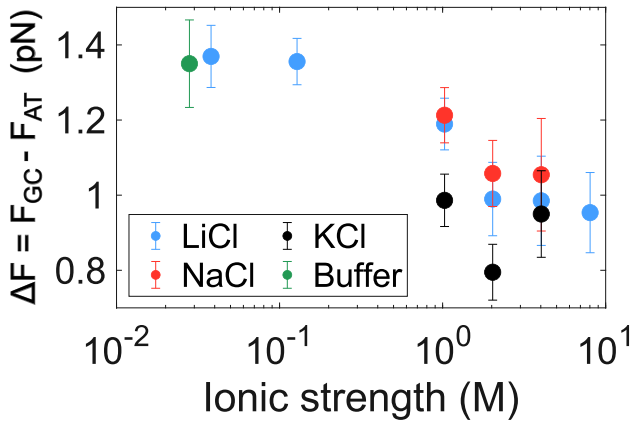


Figure 7.7: ΔF **over a broad concentration range.** The force difference between the GC and AT-rich regions plotted vs. the ionic strength of the solution.

Mean force decrease at high concentrations

This section discusses the mean force (f_m) decrease observed at high salt concentrations.DNA is commonly found in the B-DNA structure under a broad range of conditions, as this is its most stable structure under normal physiological conditions. However, as described in Sec.3.1.1, at very high concentrations ($\sim$ 4 M NaCl), the high screening of its negative charge causes a transition to a different structure, that is, Z-DNA. To our knowledge, DNA unzipping experiments in such high concentrations have not previously been reported. Studies of Z-DNA have mainly been performed using magnetic tweezers and electron microscopy [97]. Magnetic tweezers have been used to study Z-DNA by inducing the B-DNA to Z-DNA transition by applying torsion to the DNA, which generates negative supercoiling and the B to Z transition [138].

From the unzipping experiments, if the elastic parameters of ssDNA are known, the double-helix hybridization energy (ΔG_0) of the DNA hairpin can be calculated by the expression:

$$\Delta G_0 = -\int_0^{f_m} x_{ss}(f)df \tag{7.3}$$

where x_{ss} is the extension of the ssDNA as a function of the applied force. ssDNA has been shown to be well described by the WLC chain model, described in Sec.3.1.2 [139]. The persistence length (P) and base-to-base distance (b) of ssDNA have been previously calculated and are about $b = 0.7$ nm and $P = 1.2 - 0.76$ nm, as P decreases with increasing salt concentration [139]. Using the elastic parameters for ssDNA and Eq.3.4, that relates the extension of ssDNA ($x_{ss}(f)$) with the applied force, Eq.3.4 can be numerically inverted to compute the integral in Eq.7.3 [135]. Then, dividing by the number of base pairs of the hairpin, the mean free energy per base pair of the DNA hairpin can be calculated for each concentration, as shown in Fig.7.8 for NaCl. The value of $\sim$ 1.7 kcal/mol obtained at 1 M is in accordance with the value of $\sim$ 1.6 kcal/mol, which is the mean value of the base paring free energies reported for the NN model [79]. In Fig.7.8, it is observed that the free energy per base pair presents a minimum of 1 M, indicating that DNA has higher stability at 1 M. Above, although the electrostatic screening increases, the DNA's stability is reduced. This points to a change in DNA's structure above 1 M concentration, which reduces the stacking interactions between neighboring bases because of the rearrangement of the bases for the new structure. The rise per base pair in Z-DNA is 0.38 nm, compared to 0.34 nm for A-DNA (see Table 3.1). Hence, the distance between consecutive bases is greater, reducing stacking and base pairing stability [137]. In Z-DNA, the phosphate groups are closer together

than in B-DNA. This proximity is allowed by the higher screening due to the high salt concentrations.

We can't compute hairpin hybridization energy for LiCl and KCl as we lack the electric parameters P and b; to our knowledge, the elastic properties of ssDNA in LiCl and KCl have not been previously reported. However, we expect a similar behavior, as the parameters will not change significantly for other monovalent salts. The only difference is that in LiCl, the free energy minimum may be at 2 M, as the maximum of f_m is at this concentration.

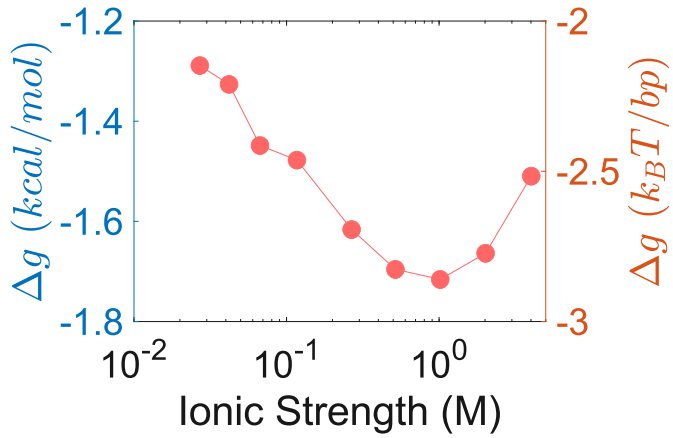


Figure 7.8: Base pairing energy per base pair as a function of salt concentration for NaCl.

The unzipping results presented in this chapter, where we observed a decrease of f_m over 1 M for NaCl and KCl and 2 M for LiCl (Fig.7.4a), imply the stability decrease for higher concentration in NaCl (Fig. 7.8). Although the base paring energy can not be computed for LiCl and KCl, a similar behavior for the base paring energy-free energies is expected for these salts. These results, together with the decrease in the force difference (ΔF) between the GC and AT-rich regions, indicate a structure of DNA.

Z-DNA formation has been found to be sequence dependent, with the propensity of forming Z-DNA being GC>CA>TA [94, 95]. Moreover, Z and B DNA can be present in the same DNA molecule, concerted by a B-Z junction, at which DNA's structure changes from B to Z (or Z to B) [97]. At this junction, to change the helix sense, a pair of bases does not hybridize and is extruded from the helix. From the results in this chapter, it is difficult to conclude exactly how the transition from B to Z DNA occurs. The transition could happen 'gradually' such that different regions of the molecule change to Z-DNA, while others remain as B-DNA so that increasing the salt concentration would increase the content of Z-DNA. However, it could also be that the DNA molecule changes from B-DNA to

Z-DNA when a certain salt concentration is reached. While the f_m in NaCl and KCl decreases gradually, suggesting a gradual change from B to Z DNA, for LiCl, the large decrease observed at 8 M indicates a sharper transition.

7.3 Discussion with previous nanopore results

The unzipping experiments with optical tweezers presented in this chapter show that the hairpin stability is reduced at high salt concentrations due to a change in DNA structure. This is relevant for nanopore experiments, as high concentrations of 4 M are usually used to reduce the DNA translocation speed and improve the SNR. Therefore, it is important to consider what is happening to DNA's structure at these high concentrations. Although nanopores are a very sensitive technique, it is not easy to use them to detect these slight differences in DNA's structure.

Having observed the results from optical-tweezers experiments, it is interesting to explore if the structural change at high concentrations is causing a signature that can be observed in the nanopipette translocation experiments. Determining changes in length or DNA diameter requires precise measurements, as the changes will be subtle. However, the increased screening of DNA's negative charge at higher salt concentrations and the structural change will substantially influence the elastic properties of DNA. It is known that dsDNA's persistence length decreases with increasing ionic strength due to the increased shielding of DNA's negative charge [139]. Additionally, Z-DNA has been found to be stabilized by DNA supercoiling, as its formation in supercoiled DNA relaxes the strain on DNA [140]. This suggests that Z-DNA is more flexible and, therefore, better at reducing strain in supercoiled DNA structures. The increased flexibility is due to Z-DNA's smaller diameter when compared to B-DNA. However, to our knowledge, no information about the elastic properties of Z-DNA has been reported.

As presented in Chapter 6, the initial conformation of DNA before translocation influences the current signatures observed during translocation. In this chapter, an increase in the knotting probability and in current drops higher than the 3° level was observed for translocations at 4 M concentration. An increase in DNA's flexibility due to a variation of the elastic parameters will influence these initial conformations, making them more compact. This is schematized in Fig. 7.9a, where the smaller persistence length P_a allows for a more dense packing of the DNA molecules, in contrast to Fig. 7.9b, where the persistence length (P_b) is larger, and therefore the conformation is more extended. The more compact conformations caused by the high salt concentrations increase the knotting probability. Hence, λ-DNA translocations at higher salt concentrations will have a higher percentage of bigger current drops caused by a higher probability of equilibrium

knots in the initial DNA conformation.

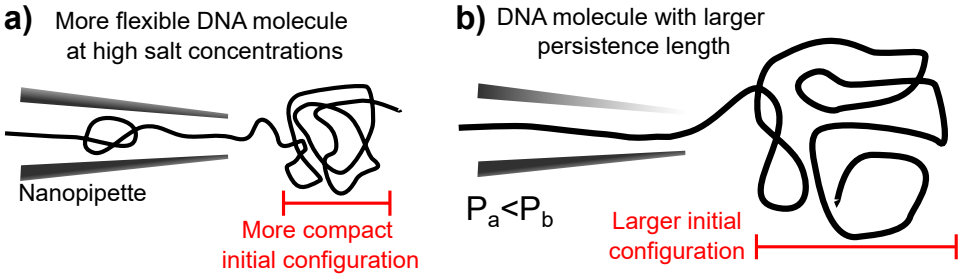


Figure 7.9: **Scheme of DNA molecules with different persistence lengths.** a) A more flexible DNA molecule with a smaller persistence length (P_a) will have more compact conformations. b) DNA molecules with a larger persistence length (P_b) will have larger, more extended conformations.

This effect with concentration is indeed observed from the translocation experiments. Fig. 7.10, shows 30 randomly selected translocation events at 900 mV for the three monovalent salts at 1, 2, and 4 M concentrations. The percentage of translocations that cross the 3$^{\text{rd}}$ level (blue) and the 4$^{\text{th}}$ level (red) are shown in the legend.

As concentration increases from 1 to 4 M, a significant increase in the number of events with higher current drops can be observed in Fig. 7.10. This increase indicates a higher presence of equilibrium knots in the initial DNA conformation for the 4 M concentration.

Surprisingly, a much higher occurrence is observed for 4 M LiCl, which is likely not solely due to knots in the initial conformation. Recent work combining experimental results and simulations has shown the formation of plectonemes during DNA translocation through nanopores [141]. Plectonemes are structures of supercoiled DNA that form due to the accumulation of torsional tension. Simulations show that during DNA translocation, the electro-osmotic flow inside the nanopore generates a toque on the DNA molecule that can cause the formation of plectonemes [142]. For plectonemes formation, simulations show that specific conditions of tension and torque onto DNA need to be satisfied. High torque is caused by a higher electro-osmotic flow and lower tension, as tension can unravel the plectonemes. These conditions are satisfied at higher salt concentrations. Therefore, at 4 M LiCl, the occurrence of plectonemes is significant increasing the percentage of translocations with higher current drops. This effect is much smaller for NaCl and KCl due to the faster dwell times, which indicate a higher tension on DNA during translocation, which reduces plectonemes formation.

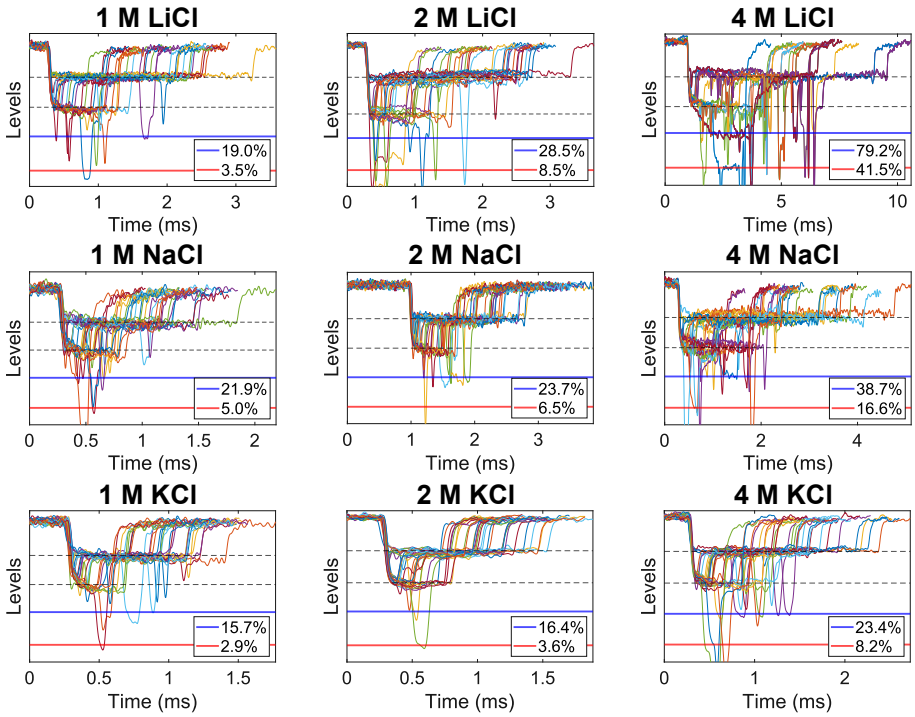


Figure 7.10: Plots showing 30 randomly selected λ-DNA translocations events for three different monovalent salts, LiCl, NaCl, and KCl. Results at 1, 2, and 4 M concentrations for each salt are shown. All the translocations are at 900 mV. The tip diameter of all the nanopipettes is similar ($18 - 26$ nm). The percentage of current drops exceeding the 3^{rd} level (blue), and the 4^{th} level (red) is shown in the legend.

7.4 Conclusions

In this chapter, we have presented unzipping experiments of a DNA hairpin at high salt concentrations. We have shown how the mean unzipping force depends on the cation type, having Li^+ a higher unzipping force than Na^+ and K^+ at the same concentration. These results connect to the dwell time results shown in Chapter 5, where dwell times in LiCl were larger. Dwell times increased with salt concentration for 2 M and 4 M in Chapter 5. However, for the mean unzipping force, a decrease was observed at 2 M, and 4 M. Additionally, the mean free energy per base pair was computed using reported elastic parameters of ssDNA in NaCl. A minimum is observed at 1 M, indicating that this is the concentration where the hairpin is more stable and that the hairpin is destabilized at higher concentrations. These results point to a structural change of DNA from B-DNA to Z-DNA, which has been previously reported for NaCl of ~ 4 M [98]. The result is particularly

important for the nanopore field, where high salt concentrations of 4 M are typically used to increase the dwell time and the SNR ratio of DNA translocation experiments.

Finally, translocation experiments show signatures of the structural change observed in the unzipping experiments. Looking at the λ-DNA translocation experiments, an increase in the proportion of translocations with higher current drops is observed at 4 M concentrations, with respect to the lower concentrations. This indicates that the stronger screening of DNA charge at 4 M and the structure change reduce DNA's persistence length, increasing its flexibility. Therefore, DNA adopts more compact conformations before translocation, increasing the knotting probability, which increases the portion of higher current drops in translocation events.

Part IV

COLLABORATIONS

Chapter 8

Single-molecule picometer resolution nanopore tweezers

This chapter results from a 30-day international internship at the University of Seattle in the lab of Jens Gundlach's. During this stay, I learned how to perform SPRNT experiments and how to build the SPRNT setup. Apart from learning the technique, I performed some experiments with the Hel308 helicase to learn the technique and experimental procedures. In addition, during my time in Seattle, we conducted some preliminary experiments with gp41, a helicase that had not previously been measured with this technique.

8.1 Introduction to SPRNT

SPRNT is an experimental technique that originates from nanopore sequencing, in which an enzyme controls the movement of DNA through a biological nanopore [67, 143, 56, 144, 145, 146]. However, in SPRNT experiments, the objective is not the sequencing of the DNA molecule but the study of the enzyme that controls the DNA movement.

The procedure to carry out SPRNT experiments is described in Fig. 8.1. More details about the experimental protocols can be found in Appendix F. Here, we briefly describe the more relevant aspects of the experimental protocols. First, a single Mycobacterium smegmatis porin A (MspA) nanopore [146] is isolated on a lipid bilayer such that it is the only connection between two salt reservoirs, Fig. 8.1a. To achieve this, a phospholipid bilayer is established by painting the phospholipids over a small aperture ($\sim$ 20 μm) that connects both reservoirs. After the bilayer has been established, when a voltage is applied across the membrane, no ion current is measured (Fig. 8.1a left panel). Then, MspA is added to the solution while a voltage is applied. The insertion of a pore into the bilayer is recognized by a jump

to a characteristic ion current. The jump depends on the voltage and ion concentrations (Fig. 8.1a central panel). From the electric current, we can determine whether the pore is inserted in a forward or backward orientation, as the MspA presents some rectification. Hence, a forward insertion like the one shown in Fig. 8.1a has a different current value than a backward insertion (opposite pore orientation). For the typical 0.5 M KCl conditions, the forward pore will have a current of 180 pA and the backward pore of 220 pA.

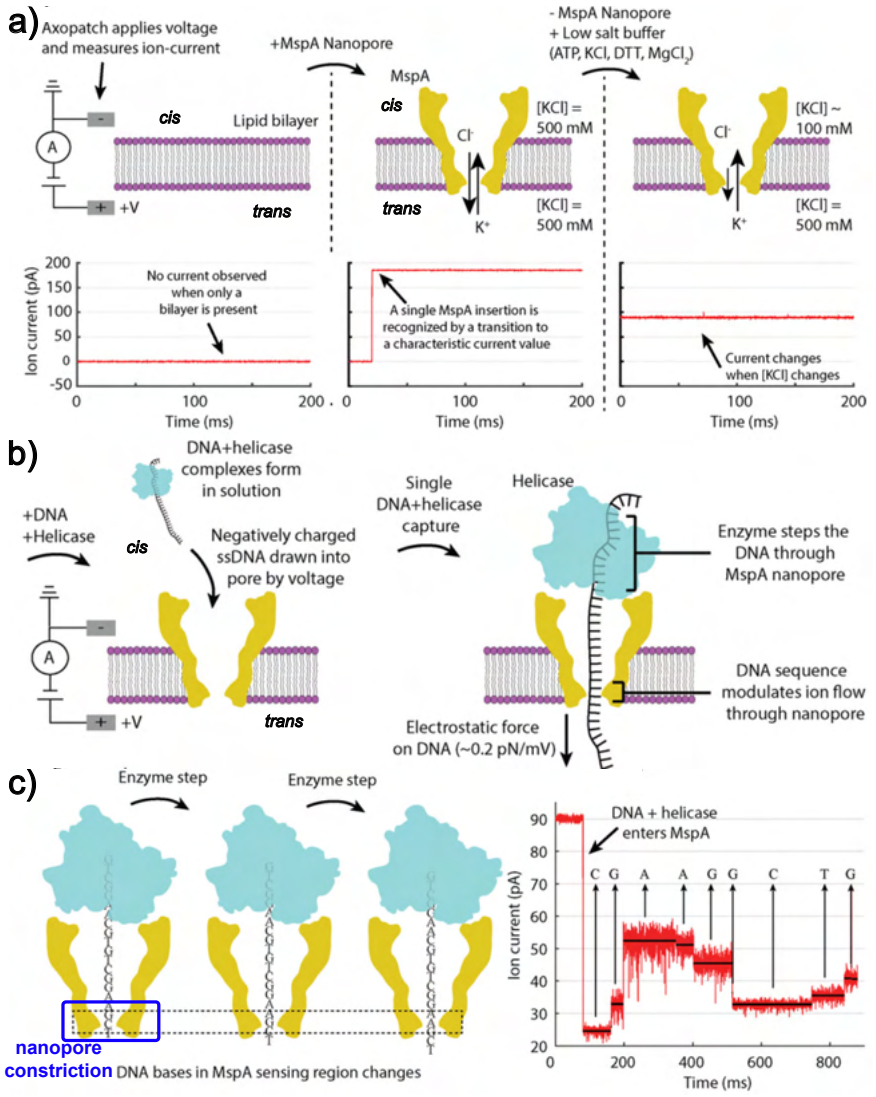


Figure 8.1: **Nanopore sequencing and SPRNT experimental setup.** a) Initial steps to establish a SPRNT experiment.

Figure 8.1: (Continued) First, a phospholipid bilayer is established, a voltage is applied across the membrane, and no ion current is measured. MspA is then added to the solution, and pore insertion into the bilayer is recognized by a jump to a characteristic ion current value that depends on the salt concentration, ion choice (normally KCl), temperature, and voltage. The buffer in the cis well can be exchanged freely, enabling flexibility in the enzyme's operating conditions. b) Enzyme and DNA are added to the solution, and DNA-enzyme complexes form. The voltage draws the negatively charged ssDNA towards the positive terminal, bringing the DNA through the pore until the enzyme rests on the rim of MspA. c) (left) An illustration of a SPRNT experiment. The enzyme takes discrete steps on the DNA, changing which bases are in the sensing region of MspA (dashed box). (right) SPRNT data of a helicase, with ion current states mapped to the DNA sequence in MspA. Figure adapted from [105].

Once a MspA nanopore is inserted in the membrane, the buffer in the cis-chamber can then be exchanged for the desired condition to measure the enzyme activity. This will change the current through the nanopore to a new current value (Fig. 8.1a right panel). The ion current is the main observable and is recorded throughout the whole experiment. When the MspA nanopore is in the desired conditions, the DNA molecules and the enzymes are flowed into the cis-chamber. During the experiment, the electric field at the nanopore will capture a single-stranded end of the DNA molecule (Fig. 8.1b, left panel). The electric field will then pull the molecules through the nanopore until the enzyme is reached, and the enzyme sits on the rim of the pore (Fig. 8.1b, right panel). From this point on, the enzyme controls the movement of the DNA molecules, and the molecule translocates as the enzyme steps along the DNA.

The bases of the DNA molecule that are in the narrowest part of the MspA nanopore, the constriction (the region of the pore with the smallest section, and hence with the highest ion current density and the highest sensitivity), determine the current flow through the nanopore (Fig. 8.1c). For the case of MspA, the constriction accommodates four ssDNA nucleotides. Hence, the measured signal is determined by the four nucleotides in the constriction. In the case of DNA sequencing, this enables the determination of the DNA sequence by measuring the ion current patterns. However, in SPRNT experiments, the sequence of the DNA molecule translocating the nanopore is known. Therefore, it is possible to infer the position of the enzyme on the DNA molecule and study the enzyme's position as a function of time as it moves over the DNA molecule. The duration of steps provides kinetic data about the motor enzyme. At the same time, the measured DNA sequence reveals the motor enzyme's exact position along the DNA, enabling analysis of how DNA sequence affects the motor enzyme

translocation over that sequence [147, 148]. During SPRNT experiments, the electrostatic force onto the DNA molecule is applied by the electric field at the pore constriction, leading to a force on the enzyme. This force can assist or hinder the enzyme movement along the DNA depending on where the enzyme introduces the DNA into the nanopore, assisted by the electric force, or pulls the DNA out of the pore while opposed by the electric force. The experiments shown in this chapter are all force-assisting experiments.

More details on the experimental protocols to perform SPRNT experiments can be found in App.E.

8.1.1 The MspA biological nanopore

MspA is an outer membrane protein that derives from Mycobacterium smegmatis. MspA has proven to be an ideal candidate for sequencing ssDNA [149]. This is due to the narrow pore constriction, which is just 1.2 nm wide [146]. This small size of the constriction can only fit four nucleotides. Hence, during SPRNT experiments, the signal is determined by the four nucleotides filling the MspA constriction at any time. A top view and cross-section of MspA are shown in Fig.8.2.

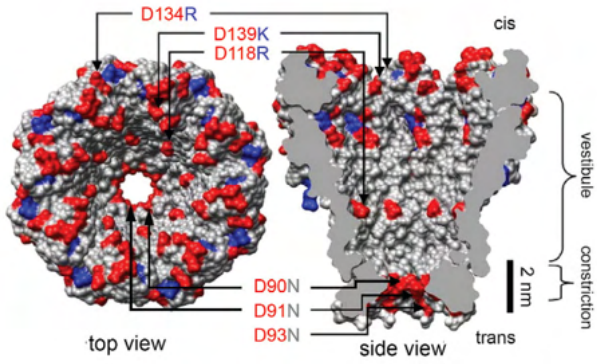


Figure 8.2: **Mycobacterium smegmatis porin A (MspA).** The figure shows Wild-type MspA. The negatively charged residues shown in red that were mutated to positive (blue) or neutral are shown. Image from [149]

In the first experiments with MspA, DNA was unable to translocate the pore. It was only after mutating the negative charge residues (D90N, D91N, and D93N), present in the nanopore constriction, to neutral residues that ssDNA could translocate through the nanopore. Further mutation of negative residues in the vestibule (D118R, D134R, and D139R) to positive ones significantly increased the DNA capture rate of the nanopore [146].

8.2 Experiments with Hel308

Hel308 is an ATP-dependent translocase that translocates on ssDNA in the 3' to 5' direction. Hel308 helicases are conserved among archaea and eukaryotes, including humans. In contrast to other enzymes that move along ssDNA, which perform steps on DNA 1 bp at a time, Hel308 was found to have substeps of less than 1 bp when moving along DNA. These sub-Amstrong measurements were performed with SPRNT. Hel308 has been extensively studied using SPRNT [20, 105]. Therefore, it was a great candidate to perform some first experiments while learning the technique.

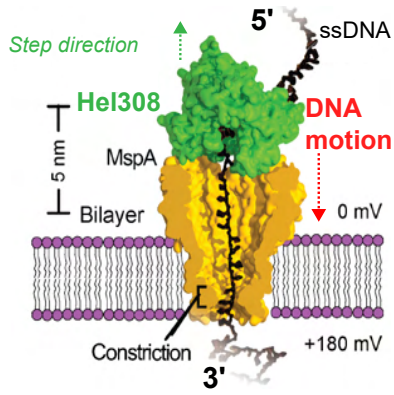


Figure 8.3: Scheme of Hel308 helicase in a force-assisting SPRNT experiment. The Hel308 sits on top of the MspA nanopore and controls the translocation through the nanopore by gradually stepping over the DNA. Image from [150].

The experiments were performed using a short 96 bases ssDNA template in a force-assisting configuration under a 180 mV voltage bias. In this configuration, the electric field pulls on the DNA assisting the Hel308 movement over the ssDNA (as exemplified in Fig. 8.3). Hel308 was measured in a 0.5 M KCl buffer, with 10 mM Tris at $pH = 7.5$. The raw data acquired during a Hel308 experiment are shown in Fig. 8.4a. It can be observed that during the experiments, short current deviations from the baseline (I_0) happen very often. These current blockades are due to ssDNA translocating without any Hel308 helicase attached. These events happen fast in the μs timescale. When performing SPRNT experiments, we are looking for current blockade events, normally of a few seconds, that have a staircase structure, indicating that the helicase is slowly stepping over the ssDNA molecule. However, often, long current blockade events with a flat signal are observed, like the one at the end of the trace in Fig. 8.4a. These events correspond to partial clogging of the nanopore or nanopore gating. When these events occur, we need to flip the voltage to unclog or get the nanopore out of a gating state. Therefore, the current must be monitored in real-time

during the experiment to flip the voltage when needed.

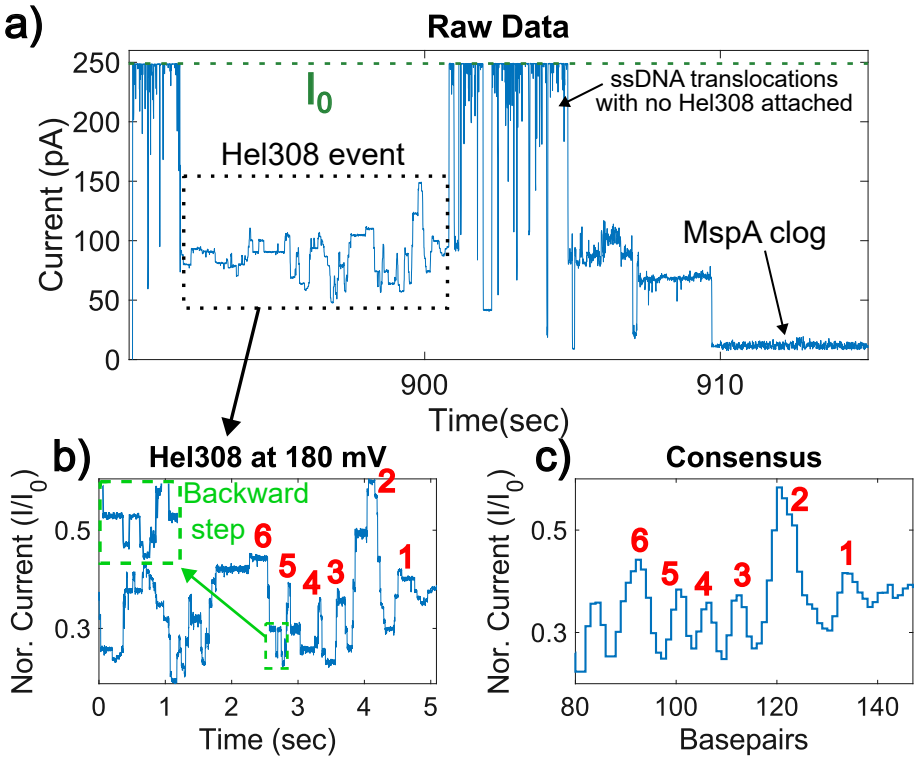


Figure 8.4: **SPRNT experiment with Hel308.** a) Shows the electric current signal during a SPRNT experiment, showing the different events that cause current drops during the experiment. b) Is a zoom-in of a translocation event controlled by Hel308. The normalized current I/I_0 is used for the y-axis. On the top left, a zoom-in of a backward step is shown. c) Consensus signal of the 96 bases ssDNA template used for the experiment. The red numbers indicate the peaks that have been visually identified in b).

For the analysis, long events with multiple levels are inspected by eye and assigned as potential Hel308 events or discarded. Figure 8.4b shows a selected event. The signal has been normalized by the baseline current (I_0). To confirm the signal is a Hel308 moving on DNA, we need to compare the signal to a consensus signal. The consensus is generated using a quadromer map from the known sequence of the ssDNA. The quadromer map maps the 256 possible current drops that can occur due to all the possible sequences of four nucleotides that can block MspA. As the ssDNA sequence is known, we can do a prediction levels sequence that we will observe during a real event. However, it is important to note that it is only a prediction of the levels that we will observe and not the residence time of the different events. The consensus signal is the signal we would get if the enzyme moved one

step at a time, with no back steps or skipping steps, and it took the same time for all the steps. The consensus of the 96 bases molecule is shown in Fig. 8.4c. Notice that the x-axis is the base number and not time. Comparing the event (Fig. 8.4b) to the consensus (Fig. 8.4c), it can be observed that the dwell times of the levels in b are very different and that the enzyme even backsteps sometimes. This illustrates the complex behavior of the Hel308 movement on ssDNA, which has been extensively studied in different works [20]. Moreover, we can correlate peaks in the real data to the ones observed in the consensus by comparing the relative height between them. By identifying some of the peaks from the consensus in our signal, we can ensure that the event corresponds to Hel308 translocation over the ssDNA 96 bases template.

8.3 Measuring gp41 with SPRNT

Measuring new enzymes using SPRNT is not straightforward. For an enzyme to be reliably measured with SPRNT, several conditions need to be fulfilled. First, the enzyme must be able to properly sit on top of the MspA nanopore while the DNA is translocating through it. This is not always achieved with new enzymes, as interactions between MspA and the enzyme may prevent it. Moreover, these interactions can also provoke the enzyme to detach from the DNA, making the enzyme's processivity very low. Some enzymes require specific sequence or structure conditions to bind to the DNA template. Therefore, an appropriate DNA template for studying the enzyme must be found. Proper buffer conditions and incubation conditions may also need to be determined. Finally, if possible, it is very useful to stall the enzyme in a specific position on DNA so that once the DNA-enzyme construction falls into the nanopore, the assisting force of the pore initiates the enzyme movement over DNA at a specific position.

Here, preliminary experiments for measuring gp41 using SPRNT are presented. gp41 is a bacteriophage T4 gp41 replicative helicase, that is a hexameric helicase, conformed by 6 equal monomers [106] that unwinds DNA with 5' to 3' polarity [151]. The six subunits conforming gp41 assemble and form a stable complex when ssDNA or forked DNA is present [152].

Three different DNA templates were measured to select a good one for gp41. Templates were measured in a 0.5 M KCl buffer, with 10 mM Tris at $pH = 7.5$. These conditions had previously been used to study gp41 [153] experiments in magnetic tweezers. A schematic representation of the three templates is shown in Fig. 8.5a. All templates have a cholesterol tag at one end for insertion into the bilayer Fig. 8.5a, to increase the translocation rate. The first template consists of an ssDNA with a short oligo hybridized at its 3' end, which contains the cholesterol tag at the 5' end of the oligo

(Fig. 8.5a top). The second template is similar but with the oligo at the 5' end of the ssDNA. The third template is a DNA hairpin with the cholesterol tag at the 3' end.

Helicase translocation events were only observed using the top template in Fig. 8.5a, which is 201 nt long. This could be due to the short ssDNA fragment at the 5' end of the two other templates, which is just $\sim$ 20 nucleotides long. gp41 needs a long ssDNA sequence at the 5' end to bind to the ssDNA, a condition not met by the middle and bottom templates. The experimental configuration using this template is shown in Fig.8.5b, left. When the nanopore captures a DNA with a gp41 attached, it translocates until gp41 stops the translocation, Fig.8.5b right. As gp41 moves in a 5' to 3' direction, the same direction the DNA translocates, the experiment is force assisting. The DNA template is a 201 nt ssDNA with a sequence designed for high resolution in SPRNT experiments. These sequences alternate sections of higher current blockades with sections of low current blockades. At the 3' end, the oligo has a tail of 26 thymines. This thymine tail is complementary hybridized to an 11 nt adenine segment of a short 26 nt oligo with a cholesterol tag at its 5' end (Fig. 8.5a, top).

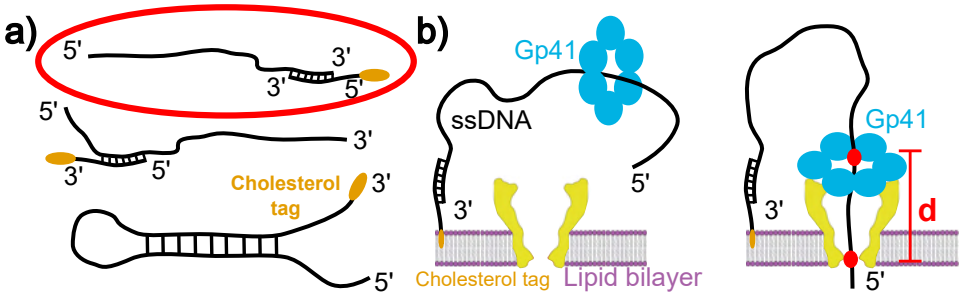


Figure 8.5: **SPRNT experiments with gp41.** a) The three templates that were tested with gp41. The highlighted top template was the only one for which gp41 events were observed. b) Illustration of gp41 translocation dynamics

In the case of the template with a long 5' end, events were only observed when the template was incubated for 5 min with gp41 before flowing it into the cis-chamber in a buffer containing:

1. 2 μL 10 mM ATP.
2. 1 μL 10 mM DTT.
3. 1 μL 0.5 M KCl.
4. 1 μL 10 mM DNA template.
5. 5 μL gp41 monomers.

gp41 events

In Fig.8.6 and b, different events obtained for gp41 at 90 mV and 120 mV are shown. No events were observed at 180 mV, indicating that the high force on the gp41 may unbind it from DNA. It can be observed that although the events do not show the expected levels of the consensus, as in the case of Hel308, the traces have a common pattern. They all start with some higher levels and then end with a long tail where no levels are appreciated. Comparing the traces to the consensus for the sequence of the DNA template, it can be observed that they match the final part of the consensus (inset in Fig.8.6b). The last part of the consensus contains peaks followed by a long flat level corresponding to the thymine tail.

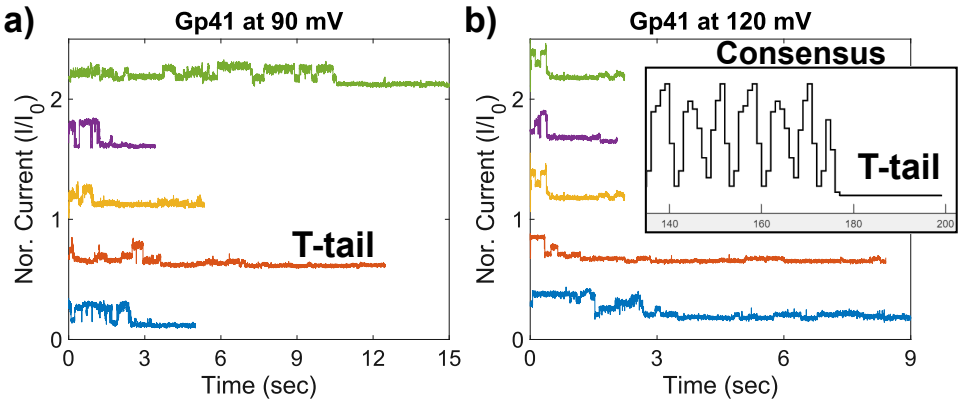


Figure 8.6: **SPRNT traces of gp41.** Events observed using the highlighted template in Fig. 8.5a. Events with an applied voltage of 90 mv a) and 120 mV b).

It is important to point out that the nucleotides measured by the nanopore are $\sim$ 20 nt upstrem (before) the position of the Gp41[150]. This is because of the spatial distance (see red *d* Fig.8.5b right) between the nanopore constriction and the sequence at the gp41.

We hypothesize that only the last nucleotides of the template are being measured because when the template is captured by the nanopore gp41 at the duplex between the 201 nt template and the short oligo, as shown in Fig. 8.7a. Hence, only some peaks followed by a long level are observed for the observed traces. A possible explanation for this would be that the duplex of the DNA template stalls the movement of gp41 over the DNA while in bulk, and gp41 will only start to unwind the dsDNA with the assisted force of the nanopore Fig. 8.7a. The exact position at which the template and the oligo hybridize can vary between the different molecules as the thymine tail is 26 nt long, and the oligo has an 11 nt adenine segment. This would explain why the starting position seems slightly different from different events in Fig.8.6. However, looking in detail into the last peak before the T-tail some

common levels can be observed between the events (red numbers in Fig. 8.7b), indicating that, indeed, gp41 is stepping over the ssDNA

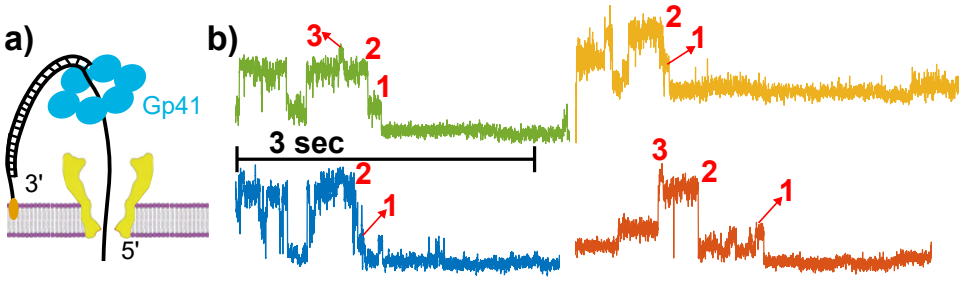


Figure 8.7

8.3.1 Conclusions

This chapter presented the SPRNT technique and how it can be used to study the mechanisms of DNA helicases at single-base pair resolution. SPRNT was used to measure the motion of Hel308 over DNA with astonishing resolution [20]. Thanks to its high resolution, SPRNT can provide the exact location of the enzyme along the DNA sequence. This means that SPRNT can be used to answer important questions about motor enzyme kinetics, such as sequence-dependent enzyme motion, backstepping, and pausing [150, 105]. This information is often difficult to obtain with other techniques due to the lack of resolution in determining the exact position of the enzyme along the DNA.

Our results show that gp41 can be measured using SPRNT. We found that the incubation of gp41 with the DNA template is necessary to observe events. Additionally, a long 5' end is required for gp41 to attach to DNA, as no events were observed for templates with short 5' ends. From the observed events, the dsDNA at the 5' end of the template seems to stall the movement of gp41 over DNA while in bulk. These two facts are important for designing a new template for measuring longer DNA sequences with gp41. Hence, increasing the statistics would allow for a more in-depth study of gp41's motion over DNA.

A design of this new template could be the one presented in Fig. 8.8. This template has a long 5' end to increase the probability of gp41 being attached to the template. In addition, the dsDNA section in the middle of the template would stall gp41 so that when it is in the nanopore, the unwinding of the ~ 40 oligo can be measured, followed by the translocation along the ssDNA. Therefore, this template could be used to compare gp41 unwinding (unwinding the green ~ 40 nt oligo) with its translocation mechanism (blue ~ 200 nt oligo).

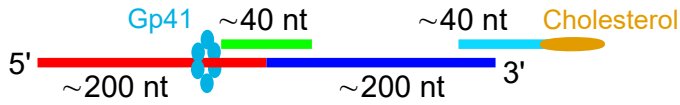


Figure 8.8: **Template for gp41 SPRNT experiments.** A possible template for gp41 that contains a long 5' end and a dsDNA section in the middle to stall the molecule, followed by a long $\sim$ 200 nt to measure gp41 translocation.

To conclude, it is important to point out that although SPRNT has proven to be an excellent technique for studying enzyme movement over DNA due to its unmatched resolution, the results obtained with SPRNT still need to be thoroughly compared to those obtained with other techniques. During SPRNT experiments, the enzymes are subjected to a strong ion flow around them that could alter their activity compared to their bulk activity. Similar to how, for example, phototoxicity can damage cells in optical experiments.

Part V

FINAL CONCLUSIONS

Chapter 9

Final Conclusions and future perspectives

In the last decades, single-molecule techniques have emerged as powerful tools to study the behavior of biomolecules with unprecedented resolution. The ability of these techniques to manipulate and monitor molecules has made them very attractive for biological applications. The study of individual molecules has proven valuable for investigating the inherent fluctuations in biological systems that are hidden from bulk techniques, as they give average information over samples containing a large number of molecules. Moreover, the capacity to apply forces in the picoNewton range allows us to explore the energetics of fundamental processes such as DNA hybridization. During this thesis, three single-molecule experimental techniques, nanopipette translocation, optical tweezers, and SPRINT, have been used to study DNA's behavior in the regime of high salt concentration, as well as the motion of enzyme motor proteins over ssDNA.

The thesis was divided into three parts, each containing the results of one single-molecule technique. Part II contained the results of electrical measurements using nanopipettes, divided into three chapters: 4, 5, and 6. Part IV included experiments with optical tweezers presented in chapter 7. Finally, part V showed some preliminary experiments with the novel SPRNT technique for studying motor enzymes in chapter 8.

Part II of this thesis starts with chapter 4, where the nanopipette conductance was characterized over a broad range of concentrations. The important contribution of surface conductance at low concentrations was shown, an effect inherently linked to the nanoscale. Furthermore, we used two conductivity models, a linear and a non-linear, to model the conductivity of salt solutions with concentration. The nonlinear model reproduced the nanopipette's conductance better over the broad range of salt concen-

trations studied. This indicated the importance of using nonlinear conductivity when estimating the diameter of the nanopipette from electric conductance measurements. Additionally, the flicker noise of nanopipettes was investigated, showing that it increases with voltage and salt concentration. Moreover, nanopipettes exhibited an asymmetric behavior of flicker noise between positive and negative voltages. This chapter's results provided the foundations necessary for interpreting and performing the DNA translocation experiments presented in the next chapters.

In chapter 5, the results of λ-DNA translocation experiments through quartz nanopipettes were presented. We studied how different translocation parameters, such as dwell time (t_D), current blockade (ΔI), and charge blockade (Q_B), varied under different experimental conditions, including voltage, salt concentration, monovalent salt type, and nanopipette tip diameter. Translocation events at a given voltage have a constant charge blockade value; however, charge blockade increases with voltage. The increase was more significant for the lower 1 M salt concentration, in contrast to 2 and 4 M. Moreover, it was found that smaller cations such as Li^+ cause larger current blockades and longer dwell times than the bigger cations such as Na^+ and K^+. The results correlated with previous simulation results that show the different interaction mechanisms of the smaller Li^+ cation with DNA. These ion-specific results are important because traditional theoretical approaches do not account for them. However, they are especially relevant in biophysics as different cations are present in biological organisms. Ions, in particular, cations, are important mediators in the interaction of DNA with proteins and screening charge effects in RNA folding; therefore, understanding the binding mechanisms to DNA is crucial for comprehending the biomolecular interactions.

Further work could be done by studying λ-DNA translocation in the presence of divalent ions such as $MgCl_2$ because divalent ions are much more effective at screening DNA's charge. It would be interesting to study if the current blockade is bigger for divalent ions, such as Mg^{2+}, as they interact more specifically with DNA. Presumably, divalent ions would increase the signal-to-noise ratio of translocation experiments. Moreover, divalent ions should strongly increase the dwell times, producing slower translocations that would allow for better discriminating DNA features.

Chapter 6 provided a study of the confined configurations of DNA during translocation. These configurations depend on the DNA molecules' 3D initial configuration at the time translocation starts and, therefore, are not influenced by voltage. Nonetheless, salt concentration increases DNA screening, reducing the DNA persistence length and causing more densely packed initial configurations. Hence, more compact confined configurations with higher knotting probability are observed for translocations at high salt

concentrations. In this chapter, we also showed how λ-DNA molecules are captured more often at their ends than expected if all parts of the molecule were equally likely to be captured. This indicates that the ends are more likely to be found outside the DNA coil and have higher mobility than the inner parts. The last section investigated the dwell time dispersion of λ-DNA translocations depending on the compactness of the confined configuration. It was shown how dwell time dispersion is mainly caused by fluctuations in DNA translocation velocity, which are a consequence of the different initial configurations of λ-DNA in the cis chamber prior to translocation. Initial configurations with a greater average distance from the nanopipette produce longer dwell times. The velocity fluctuations are smaller for translocations with compact confined configurations. By simulating multiple initial configurations of λ-DNA, we found that for molecules captured closer to the middle position (more compact confined configurations), the average distance distribution to the nanopipette had a lower mean value and a smaller dispersion. We concluded that the lower dwell time dispersion of more compact confined configurations is caused by the initial configuration being, on average, closer to the nanopipette and having less dispersion. These results are significant because reducing dwell time dispersion is essential for characterizing DNA molecules of different sizes, similar to gel electrophoresis. High dwell time dispersion will not make it possible to distinguish between the different molecules with similar sizes. Therefore, finding methods to reduce the initial configurations could, in principle, reduce the dwell time dispersion. Additional experiments regarding these chapter results could involve translocation experiments with a neutral crowder agent, such as polyethylene glycol (PEG). The crowder would cause the DNA molecules to have a more densely packed initial configuration, which would help reduce the dwell time dispersion.

Part III contained experiments carried out with optical tweezers. In Chapter 7, the effect of high salt concentrations and various salt types on the mean unzipping force of a 3.6 kbp DNA hairpin was investigated. It was found that the smaller Li^+ cation stabilizes DNA more than bigger cations (Na^+ and K^+), therefore a higher force ($\sim$ 1 pN) is required to unzip the hairpin at 1 M concentration in LiCl. This result relates to the longer dwell times observed for LiCl in Chapter 5. The mean unzipping force (f_m) had been shown to increase with increasing ionic strength in the 10 mM to 1 M range. However, we found that the mean unzipping force decreases for concentrations over 1 M in KCl and NaCl and over 2 M in LiCl. For the case of NaCl, for which elastic parameters of DNA are known, the mean free energy per base pair was computed and showed a minimum at 1 M. This indicated that 1 M is the concentration at which DNA is most stable, DNA becomes

destabilized above 1 M. The decrease in stability was attributed to a structural change of DNA, transitioning from its common B-DNA structure to a Z-DNA structure. Z-DNA has been reported to form at high concentrations of ~ 4 M NaCl because of the high screening of DNA's negatively charged backbone. We found indicators that this transition occurs in KCl and LiCl salts as well. Z-DNA is involved in reducing tension accumulated in DNA; this fact, together with its smaller crystallographic diameter, indicates that Z-DNA is probably more flexible than B-DNA. In the λ-DNA translocation experiments, we observed a significant increase in the current blockades over the 3^{rd} and 4^{th} levels at 4 M, indicating that the higher flexibility of Z-DNA is causing more densely packed initial configuration, increasing the occurrence of higher current blockades. These results are relevant for nanopore translocation experiments, where high salt concentrations are routinely used to increase translocation times and to improve the signal-to-noise ratio.

Future work should characterize the flexibility of Z-DNA by performing stretching experiments on DNA molecules and studying how high salt concentrations influence the elastic properties of DNA. Such information would be highly valuable for the nanopore field. Additionally, translocation experiments could be performed to investigate whether the percentage of translocation with higher current blockades further increases at higher salt concentrations (> 4 M). Finally, experiments in divalent salts such as $MgCl_2$ would be interesting, as the transition to Z-DNA will probably occur at lower concentrations due to the higher DNA screening by of divalent ions.

Finally, part IV presented the results of an international stay to learn about the novel SPRNT technique to study helicase motors. In Chapter 8, the basic concepts of the novel SPRNT technique were introduced. The technique's unprecedented resolution was shown from experiments performed with Hel308, which translocates at sub-angstrom steps over ssDNA. Moreover, preliminary experiments with the gp41 helicase showed that SPRNT could be used to analyze helicase dynamics by identifying the template restrictions required to study this helicase. Future experiments with an appropriate DNA template will provide longer reads and sufficient statistics to investigate gp41 with a precision not achieved to date. These experiments will help to further understand the mechanisms involved in gp41 movement on DNA.

Part VI

APPENDIX

Appendix A

A Single-Molecule Translocation Beginners Guide

This section contains all the necessary information required to start performing single-molecule translocation experiments. We describe in detail how to design and assemble the microfluidic chips for performing DNA translocation experiments. This section describes how the experimental setup works and what considerations must be taken into account to perform low-noise single-molecule translocation experiments. Special care is taken to explain crucial steps in designing and fabricating the microfluidic chip. Finally, some simple tests that may be useful to test whether different parts of the setup are working properly will be presented.

A.1 Design and Fabrication of microfluidic chips for single-molecule translocation

This section describes all the relevant steps for fabricating a microfluidic chip to perform DNA translocation experiments through nanopipette.

A.1.1 Material Selection for the microfluidic chip

Polydimethylsiloxane (PDMS), specifically Sylgard 184, is employed to produce the microfluidic chips. PDMS is a silicone-based polymer extensively used in the fabrication and prototyping of microfluidic devices and microimprint stamps within scientific research disciplines. We selected PDMS as our working material based on several advantageous properties, including flexibility, transparency, biocompatibility, and chemical inertness, all of

which contribute to its suitability for microfluidic applications. The inertness of PDMS is particularly significant as it prevents undesired chemical reactions within the microfluidic chip environment.

Beyond its intrinsic properties, PDMS is easy to handle, as it can be molded into almost any microstructure. Additionally, replica molding facilitates a fast and cost-effective way to produce microfluidic devices. Finally, PDMS can be bonded to various surfaces, such as glass or PDMS itself, through plasma treatment. This provides a straightforward means for achieving an effective seal within the microfluidic chip, thereby ensuring the integrity and functionality of the device.

A.1.2 PDMS mold design and fabrication

The molds are designed using SolidWorks, a software for 3D CAD design. The 3D design of the mold is shown in Fig. A.1a. We use a 3D printer to fabricate the molds. 3D printing is a fast, easy, and cheap way to fabricate molds that do not require precisions higher than $\sim 100\ \mu m$, which is the typical resolution limit of a resin 3D printer.

However, one must take care when using 3D-printed molds for PDMS chips that want to be sealed with glass by plasma treatment. An effective PDMS-glass seal requires both surfaces to be extremely flat and smooth, as surface roughness reduces the surface's contact area, allowing fewer bonds between the glass and the PDMS and making the sealing less strong.

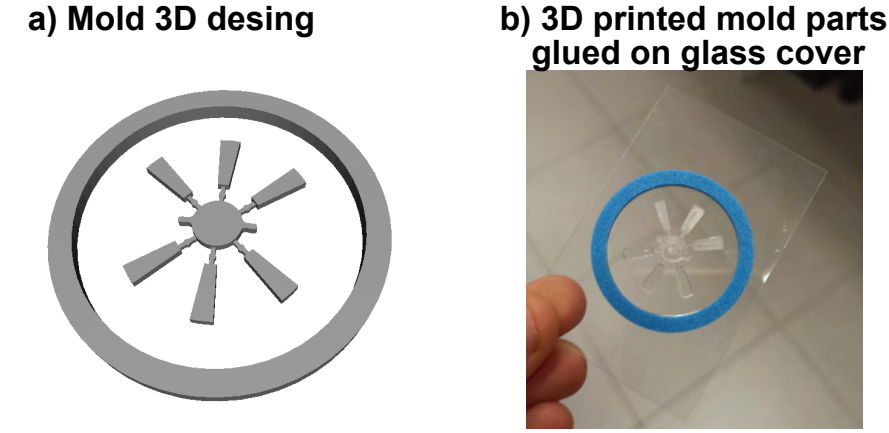


Figure A.1: a) 3D design of the mold created using SolidWorks. b) Final mold after gluing the 3D printed parts to the glass coverslip.

Because 3D printers work by continuously depositing thin layers of melted plastic, the surfaces they produce are far from smooth; they look more like a plowed field at a microscopic scale. This makes the PDMS surfaces in contact with the 3D-printed parts very rough and unsuitable

for plasma bonding. Different approaches were tested to smoothen the 3D-printed surfaces by sanding them or chemically melting the PLA (Plant-Based Resin) plastic's top layers by putting them into an acetone vapor bath to flatten them. However, the best solution was to use a glass slide as a base and glue the 3D-printed structures to it, as glass has an extremely smooth surface. This means we print a piece for the microfluidic channels and a piece that would serve as the rim of the mold (see Fig. A.1a). As we do not need much precision for the rim piece, we use a filament 3D printer with PLA. We need more precision for the channel piece, so we use a resin printer to print it.

The 3D-printed structures for the microfluidic channels are then glued onto a $75x50$ mm glass slide (1 mm thick) using a strong glue (Pattex Nural23). The final mold after gluing the parts together is shown in Fig. A.1b. When pouring PDMS into the mold, the surface of the PDMS in contact with the mold's glass slide (see Fig. A.1b) will be plasma bonded to a slide to make the microfluidic chip. As the glass surface of the mold is very flat, the PDMS bottom surface will be very flat, too, providing a much more effective seal when plasma bonding the PDMS to the glass. This is crucial, as any micro or nano connection appearing between the reservoirs in which we have the electrodes will have a lower resistance than the one of the nanopipette. As this connection is in parallel with the nanopipette, it will suppress the nanopipette signal.

After testing different microfluidic chip designs, the final design found a compromise between having as many nanopipettes as possible on the chip for higher throughput and ensuring a good seal between all the different reservoirs to avoid undesired connections. The final chip is the one shown in Fig. 2.3a.

PDMS sticks firmly to the mold surfaces to make the demolding easier and avoid damaging the PDMS chips during the demolding process; the molds are covered with silane (Trichlorosilane, in our case, from Sigma Aldrich). The silane layer is deposited onto the mold surface by vapor deposition. During this process, the mold is placed for a few hours in a vacuum chamber with a few hundred μl of liquid silane. The silane then evaporates and is deposited onto the surface.

In conclusion, our molds can be easily and cheaply designed and fabricated using 3D printing to make the structures for the microfluidic channels. By gluing them onto a glass slide, we achieve a much stronger seal of the chip using plasma bonding. A silanization treatment is used to ensure proper demolding. The only drawback of the molds is that, as they contain 3D-printed plastic parts, they cannot resist temperatures over 50° without melting. So, the PDMS needs to be cured at this temperature.

A.2 Fabrication of nanopipettes using a laser-based pipette puller

A P-2000 laser-based pipette puller from Sutter Instruments is used for the fabrication of the nanopipettes (see Fig. A.2a). Laser-based pipette pullers use a powerful laser to melt glass capillaries while pulling on them to produce pipettes. In this case, the laser is a 20 W class IV CO2 laser. CO2 lasers produce infrared light with a principal wavelength of around 9.6-10.6 μm.

a) P-2000 Pipette Puller **b) P-2000 screen showing the pulling protocol**

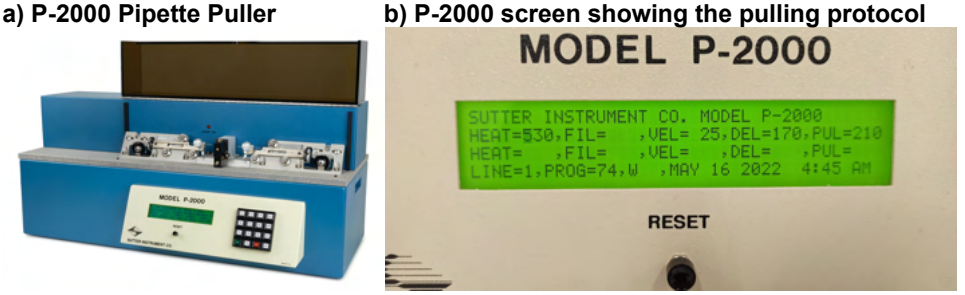


Figure A.2: **P-2000 pipette puller.**a) Image of the P-2000 pipette puller from Sutter Instruments. b) Screen display showing the parameters of the protocol used to pull nanopipettes with a tip diameter of 10-50 nm.

The pipette puller has five parameters that can be adjusted to produce pipettes of different sizes and geometries. In Fig. A.2b, the P-2000 screen with the different parameters is shown. These parameters are:

- **HEAT:** Range from 0 to 999. HEAT specifies the output power of the laser and, consequently, the amount of energy supplied to the glass.
- **FILAMENT:** Range from 0 to 15. FILAMENT (FIL) specifies the scanning pattern of the laser beam that is used to supply HEAT to the glass.
- **VELOCITY:** Range from 0 to 255. The VELOCITY (VEL) parameter specifies the velocity at which the puller bar must move before executing the hard pull.
- **DELAY:** Range from 0 to 255. The DELAY (DEL) parameter controls the timing of the start of the hard pull relative to the deactivation of the laser.
- **PULL:** Range from 0 to 255. The PULL parameter controls the force of the hard pull.

A more detailed explanation of all the parameters can be found in the P-2000 Operation Manual [154]. A brief summary of how a protocol in the P-2000 produces pipettes is presented here to get a feeling of how the parameters influence the shape and size of the final pipette:

1. The laser heats the capillary using a specific power given by HEAT. This heating needs to be evenly around the capillary. For this, two parabolic gold mirrors are used to focus the laser light onto the capillary from both sides. Higher heat values produce faster protocols as the capillary heats more quickly. This generally results in pipettes with a shorter taper, as heat has less time to transfer to adjacent regions of the capillary that are not directly heated by the laser.

2. The filament parameter (FIL) controls how much the laser is scanned along the capillary, determining how big is the glass region that is being heated. Higher FIL values produce pipettes with a longer taper as more glass is being melted and then pulled.

3. When the glass starts to melt, the two ends of the capillary start to move apart because the capillary is under a constant load in the puller. The puller monitors the velocity at which they move apart. This velocity is related to the viscosity of the glass, and so it is an indirect way to monitor the glass temperature. When the velocity reaches the desired value (determined by the VEL parameter), the puller stops the heating laser. Higher velocities indicate that the glass is being heated and pulled by the constant load until the glass is less viscous (more melted). Higher velocities produce longer tapers.

4. After stopping the laser, the DEL parameter controls how long of a delay time the puller waits until performing the final hard pull. This is the time the capillary is cooling off before the final pull.

5. The PULL parameter controls the force of the final hard pull that separates the capillary into two pipettes.

To produce pipettes with tip diameters in the nm range, capillaries with a thick wall are needed as a general rule of thumb for producing nanopipettes. To make nanopipettes, glass needs to be pulled until the nm range, and if the glass walls are too thin, they will break before the tip is of nm size. This also needs to be considered in the puller protocols, as protocols that pull the glass very thin before the hard pull will not produce pipettes in the nm range.

For the fabrication of our nanopipettes, we used quartz capillaries that are 5 cm long, have an inner diameter of 0.2 mm, and have an outer diameter

of 0.5 mm from Sutter. The protocol used to produce nanopipettes with tip sizes of 10-50 nm is shown in the following table:

Table A.1: **P-2000 protocol.** Protocol to generate nanopipette.

Parameters	**Heat**	**Filament**	**Velocity**	**Delay**	**Pull**
Value	530	0	25	170	210

Disclaimer: This protocol is specific to our lab's pipette puller and may not produce nanopipettes of the same size when used with other pipette pullers. Protocols, especially those for producing such small nanopipettes, are puller-dependent.

A.2.1 Controlling environmental parameters when fabricating nanopipettes

Apart from the different parameters in the puller, variations of environmental parameters also influence the size and geometry of the pipettes we fabricate. So, these parameters need to be controlled to minimize the dispersion in tip sizes as much as possible. The main environmental parameters are:

1. Properly centering the capillary into the groove. The capillaries need to be precisely placed into the groove of the pull bars of the puller to ensure the laser hits the capillary in the middle.

2. Pressure at which the capillaries are fixed to the pull bar of the pipette puller. It is important to fix the capillaries, applying always pressure on them. For this, some marks can be drawn onto the screws used to fix them. And then they can be screwed each time until the same mark.

3. The gold mirror and alignment of the laser need to be checked regularly. See pipette-puller maintenance section for more details[154].

4. The temperature and humidity of the room where the puller is used must be as stable as possible.

A.2.2 Pipette puller maintenance

The pipette puller needs regular maintenance to produce similar nanopipettes using the same protocol over long periods of time. The maintenance explained in this section should be done once every $2-3$ months.

1. Checking the gold mirror. Due to the high temperature at which the quartz is heated, dirt on the capillaries' surface evaporates and then deposits onto the mirror's gold surfaces. Using a regular mirror, one can check if the gold mirror is clean, and use a Thorlabs lens cleaning paper with some ethanol or acetone to gently clean its surface if needed.

2. Checking the laser alignment. The puller laser may become slightly misaligned over time. To check the alignment, we place a capillary into the puller and some laser paper behind it. We then use protocol P78 of the puller, which turns on the laser very briefly. We then check the pattern on the paper. The pattern is symmetric when the laser is properly aligned (as seen in Fig. A.3a). If it is not symmetric, as observed in Fig. A.3b, the micrometer screw on the back of the puller is used to move the laser position up or down to make the pattern symmetric as in Fig. A.3a.

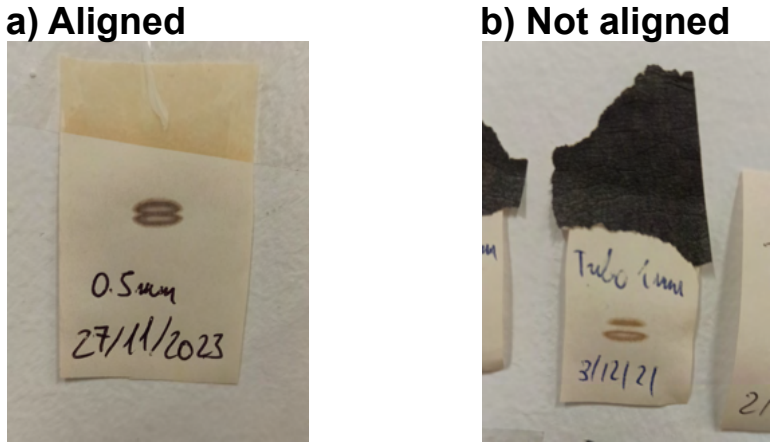


Figure A.3: a) The pattern on the laser paper when the pipette puller laser is properly aligned and centered on the glass capillary. b) Pattern when the laser is misaligned.

A.3 Characterizing the nanopipettes with electron microscope

After fabricating the nanopipettes, one must characterize them to determine their size and geometry. For this, we use an electron microscope, as the size of the nanopipette tips is well under the diffraction limit of an optical microscope. Two electron microscopes were used to characterize the

nanopipettes: an SEM (scanning electron microscope) and a TEM (transmission electron microscope). The SEM microscope is easier to use, the sample preparation is simple, and one can measure many samples relatively quickly. However, in the SEM, the nanopipettes are observed from above, so we only measure the size at the tip, not their geometry. In principle, one can also observe them from the side, but the SEM cannot provide information about the internal structure of the pipette, as it only gathers electrons reflected from the sample. Furthermore, the tip size of our nanopipettes is close to its resolution limit, so it is difficult to tell their exact size for small tips. A TEM microscope is used to characterize the nanopipettes to overcome these limitations. TEM microscopes provide higher resolution and allow a more detailed analysis of the nanopipettes' tip size and geometry. In addition, TEMs allow for imaging the nanopipettes in cross-section, providing valuable information about the internal structure of the pipettes, for example, the angle of the cone and if this angle is constant. However, the sample preparation to study the nanopipettes using a TEM microscope is more complex, requiring the nanopipettes to be carefully cut 2-3 mm away from the tip to be glued on a 5 mm big TEM grid. This has to be done with extreme care, as the nanopipettes are very delicate, and it is easy to break them when they are cut so close to the tip. Apart from the preparation, TEM measurements are much more time-consuming than SEM measurements. As TEM needs a very high vacuum, this vacuum has to be performed every time we change the grid, and one can only fit two nanopipettes per grid.

A.4 Assembling the microfluidic chamber

Several steps are undertaken to perform the final assembly of the microfluidic chip. First, one must punch some holes into the PDMS mold that will be used for placing the electrodes and to flow the liquid PDMS that then seal the microfluidic channels. We then cleaned the glass coverslip and the PDMS. For this, we use water and ethanol and then let them dry. We also use some tape to clean off bigger dust particles or other contaminants from the surfaces, as those prevent a good plasma bond. After cleaning them, we have to place the fabricated nanopipettes into their corresponding positions in the PDMS mold. The pipettes need to be cut to an appropriate size with a cutter.

Once all the pipettes are placed, we put the PMDS mold and the glass coverslip into the plasma cleaner to perform the plasma bonding. We then perform a partial vacuum inside the cleaner and flow some oxygen to replace any other gas inside it. We then use a 45-50 seconds oxygen plasma treatment to activate the surfaces and ensure proper bonding between the

PDMS and glass. To ensure that we have proper oxygen plasma, we can look through the small window of the plasma cleaner, and we should see white plasma. If the color is more pink/purple, it means that we have not properly replaced the nitrogen gas present in the air with oxygen. Once the plasma treatment is finished, one has to quickly press the glass and PDMS together to initiate the bonding process. To ensure that it is properly bonding, one can place the chip against the light, as the non-bonded areas where there is still air between the glass and the PDMS appear different. Areas with good bonding between the coverslip and the PDMS can be identified by looking backlight at the chip, as areas where there is still air between the PDMS and the coverslip will look different. We can gently press on those areas with a finger to facilitate the bonding, but avoid pressing on already bonded areas, as it can disrupt the bonding. Exposing the chip to high temperatures for a few minutes after the bonding process has been completed will reinforce the bonding. So we place it for 10 min onto a hotplate at 100^oC. Finally, we have to flow the liquid PDMS into the microfluidic channels where we placed the nanopipettes to seal them so that the only connection between the reservoirs is the nanopipette. For this, a syringe is used to inject the liquid PDMS into the microfluidic channels carefully, and then the chip is left on the hotplate for 20 more minutes. The assembled chips can then be stored and ready to use.

A.5 Filling and using the chip for an experiment

Due to the small size of the nanopipettes, they need a special procedure to be properly filled with the desired buffer. Before the filling process, the chip is placed for 5-6 min into the plasma cleaner. This helps not only to clean all the surfaces inside the chip from contaminants but also renders all these surfaces hydrophilic, facilitating a proper filling of the nanopipettes when we follow the buffer into the chip. The filling needs to be done as soon as possible after the plasma cleaning process, as the surfaces lose their hydrophilic property over time. A syringe is used to follow the desired buffer into the chip. The buffer can be followed into the chip chambers by punching through the PDMS while the air flows out through the holes we punch into the chip. The chip can be wrapped with a plastic film and placed in the fridge overnight to enhance the filling, especially for small nanopipettes. This prevents buffer evaporation while we wait for the nanopipettes to fill.

After filling the chip, we need to wait at least 30 min before performing measurements on it. As all the chip surfaces are very hydrophilic, the outer surfaces of the chip can get covered by a thin layer of liquid condensing onto the surface due to its high hydrophilicity or by any spill that may have occurred during the filling process. This liquid can connect the different

chambers of the tip so that the nanopipettes will not be the only connection between them anymore. As these connections will be much larger (and will therefore have a much lower resistance), they will overcome any measurements we try to make on the nanopipettes.

Finally, to perform the experiments, we place one Ag/AgCl electrode in the central chamber and one in the chamber at the end of the nanopipette we want to measure. We used the Axopatch 200B to apply a voltage across the nanopipette and measure its current.

Appendix B

Synthesis of DNA hairpin for Optical Tweezers experiments

In this section, the synthesis of a 3594 bp DNA hairpin with a tetra loop (ACTA) and two 29-bp dsDNA handles at the ends is described. At the ends of the handles we have biotin and digoxigenins (DIG) so that they attach to the beads that are used in the experiment. The beads are covered with streptavidin and anti-digoxigenin. The preparation of the beads is not described here. The synthesis is based on a previously described method[73]. The DNA sequences used for the synthesis are shown in Table. The λ-DNA for the synthesis and all the necessary enzymes were ordered from New England Biolabs. The custom-made oligos were ordered from Merk. In summary, the synthesis consisted of a digestion of λ-DNA with EcoRI to obtain the 3530 bp segment. Then, this segment was phosphorylated. In parallel, the ordered oligos need to be labeled with biotin or digoxigenins. Finally, the oligos and the λ-fragment are annealed and then ligated. The final DNA construct with all the required oligos is shown in Fig.B.1

B.1 Required DNA fragments and oligos for the synthesis

To synthesize the DNA hairpin, 5 different oligos are needed. These oligos are ligated to a 3530 λ-DNA fragment. As the hairpin handles are dsDNA, two oligos are required for each handle (see Fig.B.1). These oligos are then labeled with biotin and digoxigenins at the ends. Both oligos are labeled to make the interaction stronger and more reliable. In addition, an oligo to make the loop of the hairpin is also needed. In Table B.1 the different oligos

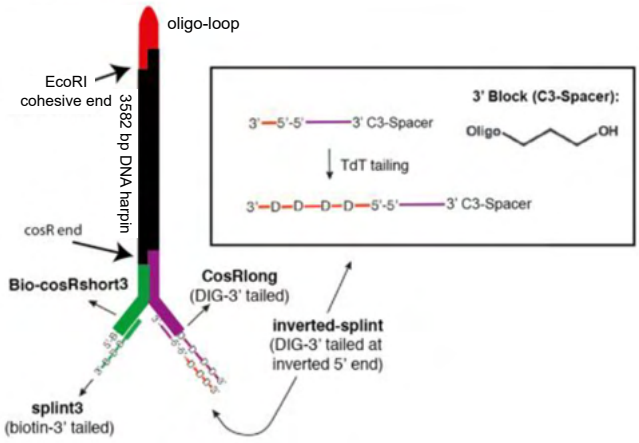


Figure B.1: Schematic of the 3582 bp DNA hairpin construct used in optical tweezers experiments.

are shown. In bold blue color, the 4 bases that form the loop are shown. In bold green, the labeled ends of the hairpin are shown. Finally, in different colors, the complementary regions between the different oligos and between the oligos and the DNA fragment are shown.

For the main part of the DNA hairpin, a fragment of λ-DNA that is obtained by digesting it with EcoRI is used. The fragment with its corresponding single-stranded ends is shown in Table B.2. The single-stranded ends that are complementary to some part of the oligos are shown in different colors.

B.2 Synthesis steps

The different steps of the synthesis process are explained in this section.

B.2.1 Digestion of λ-DNA with EcoRI

To digest a λ-DNA molecule with EcoRI the following reaction is prepared (Table B.3). This preparation is **incubated for 3 h at** 37^oC**, and then EcoRI is thermally inactivated at** 65^oC **for 20 min**. From the digestion of the λ-DNA molecule with EcoRI, we should obtain the fragments shown in Fig.B.2 a).

If we want the sample to be very clean, so that only 3530 bp DNA fragments are present in the sample, we can perform a purification by gel electrophoresis. This step is not strictly necessary, as the custom-order oligos are designed to anneal only with the 3530 bp fragment. Only the

Oligo-Name	Oligonucleotide sequence
oligo-loop	5'-Pho- AAT T GC CAG TTC GCG TTC GCC AGC ATC CG **A CTA** CGG ATG CTG GCG AAC GCG AAC TGG C -3'
Bio-cosRshort3	5'-**Bio**- GAC TTC ACT AAT ACG ACT CAC TAT AGG GA A ATA GAG ACA CAT ATA TAA TAG ATC TT -3'
cosRlong	5'-Pho- GGG CGG CGA CCT AAG ATC TAT TAT ATA TGT GTC TCT ATT AGT TAG TGG TGG AAA CAC AGT GCC AGC GC -**Dig**-3'
splint3	5'- TCC CTA TAG TGA GTC GTA TTA GTG AAG TC -**Bio**-3'
inverted-splint	3'-**Dig**-AAA AA-5'-5'- GCG CTG GCA CTG TGT TTC CAC CAC TAA C (SpC3)-3'

Table B.1: Oligos used for the hairpin synthesis of the 3594 bp DNA hairpin.

λ-DNA 3530 fragment	Sequence
3530 bp fragment	5-Pho'- AGG TCG CCG CCC - λ-3530 bp Seq. -3'
	3'- λ-3530 bp Seq. - TTAA -<u>Pho</u>-5'

Table B.2: λ-fragment used for the synthesis of a 3594 bp DNA hairpin

loop will anneal to other fragments, as the EcoRI always leaves the same cohesive end when it cleaves DNA. However, if we don't purify the hole sample, it is always good to run a gel with a small portion of it to ensure that the digestion was performed properly, and that we have fragments of the desired size. For this, we prepare a 0.8% agarose gel with TBE and run the digested λ-DNA for 1-1,5 h at 100-120 V. After this time, we should see a clear separation of the 3530 bp fragment (as seen in Fig.B.2 b)).

The digestion of λ-DNA can be avoided by purchasing commercially available λ-DNA pre-digested with EcoRI, which is sold by some companies as a marker for gel electrophoresis.

B.2.2 Phosphorylation of λ-DNA/EcoRI to put 5'-Pho at cosL end

This step is performed to add a phosphate to the 5' end of the 3530 bp fragment (underlined phosphate in Table B.2). This is needed because when EcoRI cleaves, the fragments do not conserve this phosphate. This phos-

Reactant	Volume
λ-DNA ($500ng/ml$)	$20\mu L$
10x Buffer NE EcoRI	$10\mu L$
BSA	$1\mu L$
EcoRI	$5\mu L$
MiliQ Water	$64\mu L$
Total Volume	$100\mu L$

Table B.3: λ-DNA digestion with EcoRI

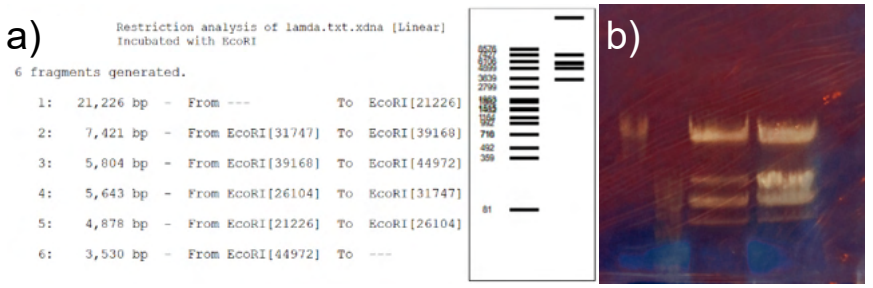


Figure B.2: a) Fragments obtain from λ-DNA digestion with EcoRI. On the right, an image of how the bands should look when performing a gel with a specific DNA marker is shown. b) A photo of a gel performed with a λ-DNA digested with EcoRI.

phate is important as it is afterward needed for the ligation. The preparation for this reaction is shown in Table B.3, and is then **incubated 30 min at** 37^oC **and then polynucleotide kinase is thermally inactivated at** 65^oC **for 20 min.**

Reactant	Volume
λ-DNA/EcoRI digested	$50\mu L$
10x PNk Buffer	$10\mu L$
ATP 10 mM	$10\mu L$
polynucleotide kinase	$1\mu L$
MiliQ Water	$19\mu L$
Total Volume	$100\mu L$

Table B.4: 3530 bp fragment phosphorylation.

B.2.3 Purification of λ-DNA/EcoRI with QIA quick purification Kit

This step is to remove the residues of EcoRI and polynucleotide kinase. And it is performed by following the steps explained in the QIA quick purification Kit.

B.2.4 Oligonucleotides labeling

In this step, cosRLong, inverted splint, and splint oligos are labeled with biotin or digoxigenins at their ends (bold green Bio and Dig in the oligo TableB.1). The cosLshort3 oligo is already ordered and labeled with biotin (bold green Bio in TableB.1).

Dissolved the new oligos

If the oligos are new, we first need to dissolve them so that we have them all at a $100\mu M$ concentration. We centrifuged the oligos, added the necessary MiliQ water, and then incubated them for 15 min at 50^oC to better dissolve them.

Dig Labeling of oligos

For labeling the oligos with digoxigenins, the reaction in Table B.5 is prepared. Then it is **incubated 1 h at 37^oC and then thermally inactivated at 75^oC for 20 min.**

Reactant	**Volume**
cosRlong ($100\mu M$) or inverted splint ($100\mu M$)	$1\mu L$
reaction buffer	$4\mu L$
CoCl2 solution	$4\mu L$
Dig-dUTP	$1\mu L$
dATP	$1\mu L$
MiliQ water	$8\mu L$
terminal transferase	$1\mu L$
Total Volume	$20\mu L$

Table B.5: Biotin labeling of cosRlong and inverted splint oligo.

Bio Labeling of oligos

For labeling the oligos with biotin, the reaction in Table B.6 is prepared. Then **was incubated for 1 h at 37^oC and then thermally inactivated at 75^oC for 20 min.**

Reactant	**Volume**
splint3 $(100\mu)M$	$1\mu L$
reaction buffer	$4\mu L$
CoCl2 solution	$4\mu L$
Bio-dUTP	$1\mu L$
dATP	$1\mu L$
MiliQ water	$8\mu L$
terminal transferase	$1\mu L$
Total Volume	$20\mu L$

Table B.6: Digoxigenin labeling of splint3 oligo.

Add $2\mu L$ of EDTA 0.2 M, pH 8.0

EDTA is added to stop the terminal transferase reaction, as it sequesters the Co+2 cations needed by the terminal transferase to work.

Purification of oligonucleotides with QIAquick Nucleotide removal kit

Resuspension is done in $50\mu L$ of EB buffer. If we assume 100% efficiency, the concentration would be $2\mu M$.

Dilution of cosRshort3 and loop oligos for annealing

The Bio-cosRshort3 at $100\mu M$ is diluted with miliQ water 1/20 to get $5\mu M$. The loop at $100\mu M$ is diluted with miliQ water 1/5 to get $20\mu M$. The other oligos are at a concentration of about $2\mu M$ after the resuspension after the purification kit.

B.2.5 Annealing reaction

To perform the annealing of all the oligos and the 3530 bp fragments, they are incubated in the preparation shown in TableB.7. The preparation is then **incubated at 70^oC for 10 min, and then the bath is turned off until the sample cools down to room temperature.**

Reactant	Volume
λ-DNA/EcoRI/PNK	$50\mu L$
cosRlong ($\sim 2\mu M$)	$4\mu L$
inverted splint ($\sim 2\mu M$)	$4\mu L$
splint3 ($\sim 2\mu M$)	$4\mu L$
loop ($20\mu M$))	$4\mu L$
Bio-cosRshort3 ($\sim 5\mu M$)	$1\mu L$
Tris 1M pH 7.5	$2\mu L$
NaCl 5M	$1.5\mu L$
MiliQ water	$4.5\mu L$
Total Volume	$75\mu L$

Table B.7: Annealing reaction.

B.2.6 Ligation reaction

Finally, a ligation of the oligos with the DNA fragment needs to be done to have the final DNA hairpin construct. For this, the reaction in Table B.8 is prepared, and **incubated overnight at** 16^oC **(about 15 Hours) and then thermally inactivated for 15 min at** 65^oC**.**

Reactant	Volume
Annealed sample	$75\mu L$
T4 ligase buffer	$10\mu L$
ATP 10mM	$10\mu L$
T4 ligase	$5\mu L$
Total Volume	$100\mu L$

Table B.8: Ligation reaction with T4 liagse.

Appendix C

Appendix Chapter 5

This appendix contains additional figures from Chapter 5.

Conductance drop for different concentrations

In Chapter 5, it was shown that the relative conductance drop ($\left|\frac{\Delta G_1}{G}\right| \cdot 100$) is dependent on the cation type. In Fig. C.1, we show that the conductance drop shows no dependence on salt concentration.

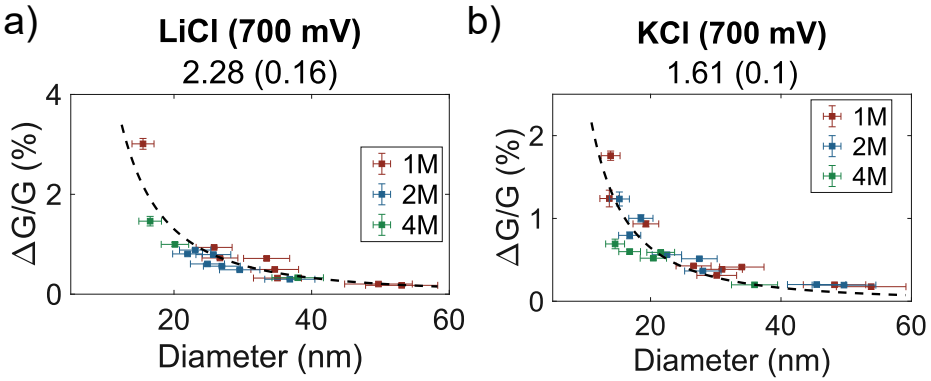


Figure C.1: Conductance drop percentage of the first blockade level $\left|\frac{\Delta G_1}{G}\right| \cdot 100$ for 1, 2 and 4 M of a) LiCl and b) KCl. The data shown are from translocations at 700 mV.

Charge blockades with concentrations

In Chapter 5, it was shown that charge blockade increases with concentration for KCl and that there is no clear dependence of the charge blockade on diameter. In Fig. C.2, we show the case for NaCl and LiCl, where the same results can be observed. It can be appreciated that the charge blockade of LiCl is also larger than for NaCl at a given concentration, as expected from the results in Chapter 5.

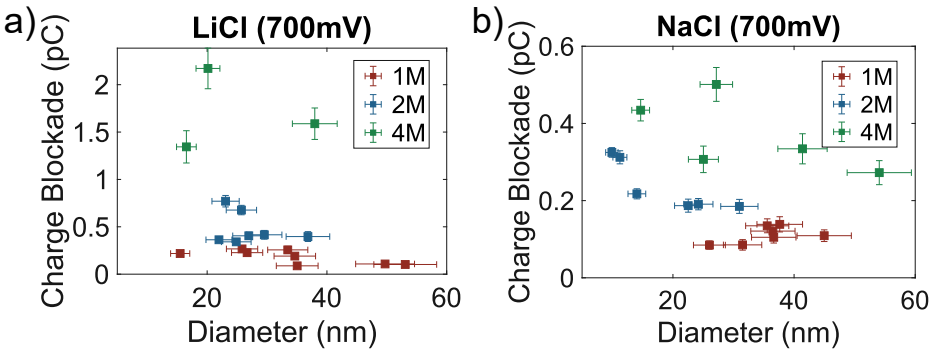


Figure C.2: Charge blockade for 1, 2, and 4 M of a) LiCl and b) NaCl. The data shown are from translocations at 700 mV.

Investigation of charge blockade

As shown in Chapter 5, the mean charge blockade $\langle Q_B \rangle$ of translocations events at different voltages for the same nanopipette is not always a constant value. In Fig.C.3 it can be observed that while at 1 M $\langle Q_B \rangle$ increases with voltage, this voltage dependence decreases for higher 2 and 4 M concentrations, and at 4 M, there is almost no voltage dependence of $\langle Q_B \rangle$.

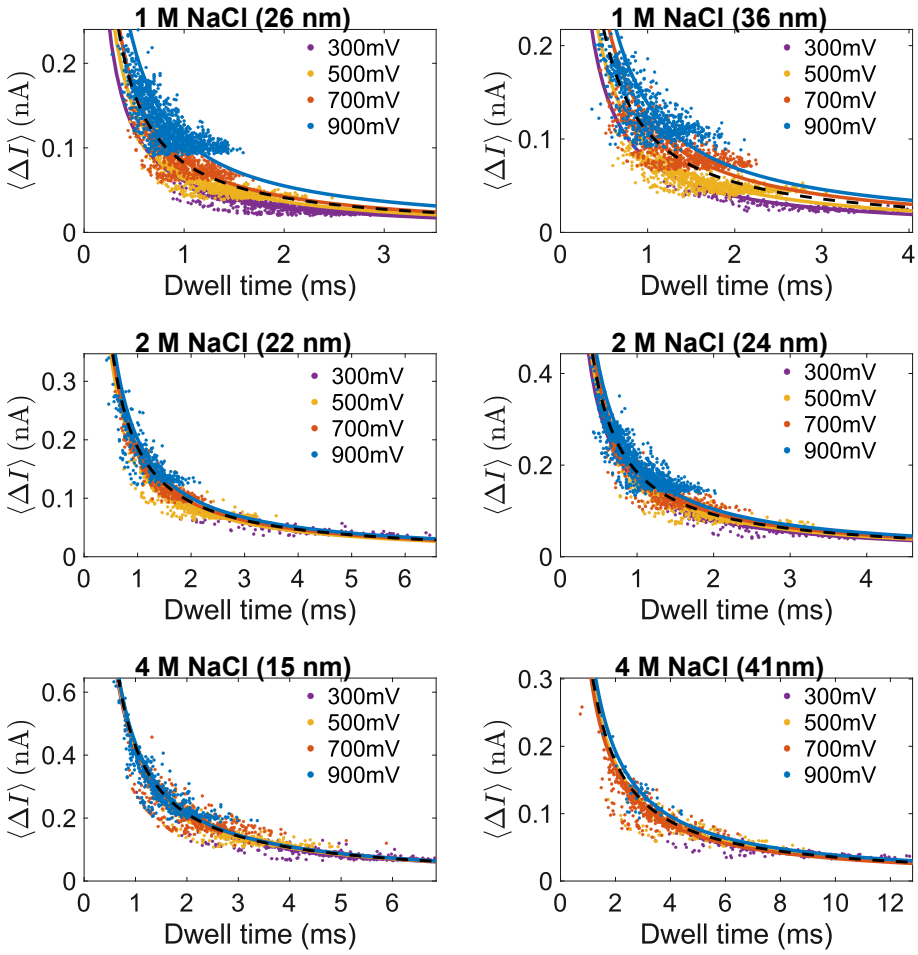


Figure C.3: Scatter plots of dwell time vs. mean current drop $\langle \Delta I \rangle$ for 1, 2, and 4 M of NaCl from different voltages. The concentration and tip diameter of the nanopipettes are indicated on the top. The mean charge drop for each voltage is indicated by a solid line in the corresponding color.

Appendix D

MATLAB codes developed for DNA translocation analysis

A significant part of the work done during this thesis was developing and creating codes to analyze nanopore translocation data, as our group did not have a tool for this type of analysis. This appendix presents the most relevant and most used codes for this purpose. The codes presented here have been developed using MATLAB App Designer, which is a versatile tool for creating interactive applications. This makes it very useful for analyzing nanopore translocation data. When analyzing nanopore translocation data, it is important to have a visual check of the analysis to confirm the quality of the analysis. Thanks to its graphical user interface (GUI) capabilities, a Matlab app can combine complex data processing workflows with a user-friendly app that allows for real-time visualization of the results from the analysis. These attributes make App Designer an ideal platform for designing codes for translocation analysis.

D.1 MATLAB app for the pretreatment of DNA translocation data

The primary objective of this first app is to convert '.TDMS' files generated by the LabVIEW program, which records the Axopatch signal, into a more user-friendly format that is compatible with MATLAB for further analysis. During nanopore experiments, transient issues such as pore clogging or excessive noise can degrade the signal. To address this, the app enables pre-treatment of the signal, allowing users to identify and remove noisy segments before proceeding with detailed analysis. Additionally, the recording

system does not automatically save critical experimental data, such as salt type, salt concentration, pH, DNA concentration, and applied voltage. The app provides functionality to manually input and associate this metadata with the experimental data, ensuring that both the signal and contextual information are preserved for subsequent analyses.

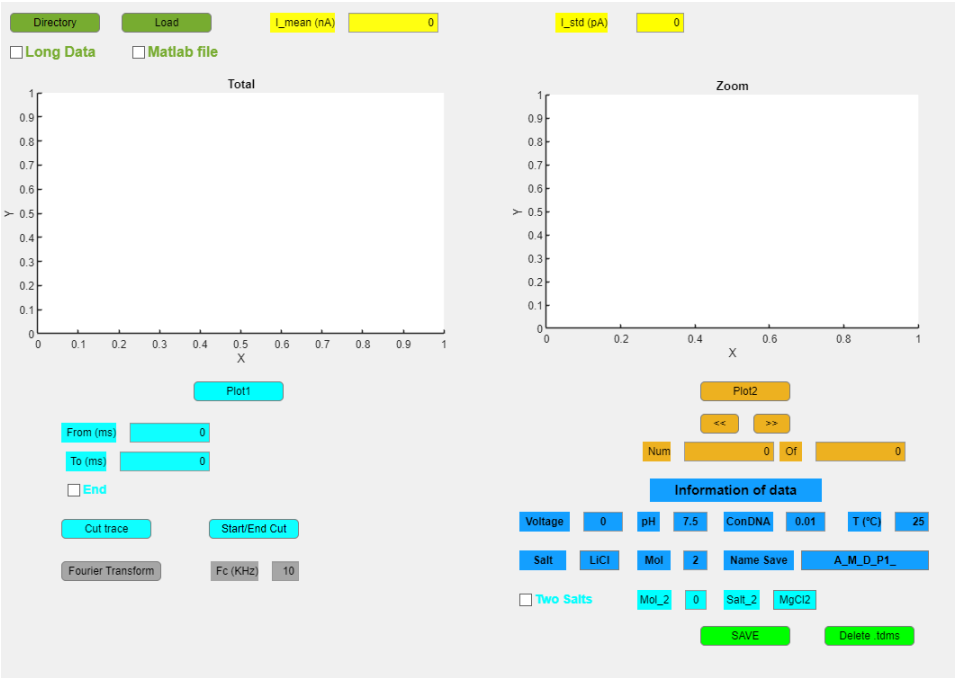


Figure D.1: MATLAB App GUI allows pre-treating the translocation data and saving them together with important parameters about the experiment.

D.2 MATLAB app for the analysis of DNA translocation data

This app is specifically designed to analyze DNA translocation data, offering a systematic and user-friendly workflow for processing the current traces acquired during translocation experiments. The app detects the DNA translocation events and calculates important parameters for each event, such as the dwell time, the mean current blocked, or the charge blockade. The app follows these key steps:

1. Data Filtering: Raw signals are pre-processed using a low-pass filter to remove high-frequency noise, improving the signal clarity for subsequent analysis.

2. Threshold Selection: A threshold is used to detect translocation events. The threshold depends on experimental factors such as nanopipette tip diameter, salt concentration, and noise. It is commonly chosen to be half the current blockade observed for the first level of translocation events. More information in Sec. 5.2.

3. Noise Reduction and Event Identification: The signal undergoes box filtering to suppress noise and reduce data points, which enhances computational efficiency. This processed signal is then analyzed to detect where DNA translocations occur. More information in Sec. 5.2.

4. Event Extraction: Identified translocation events, along with a few milliseconds of data before and after each event, are extracted from the signal. These events are saved for further analysis.

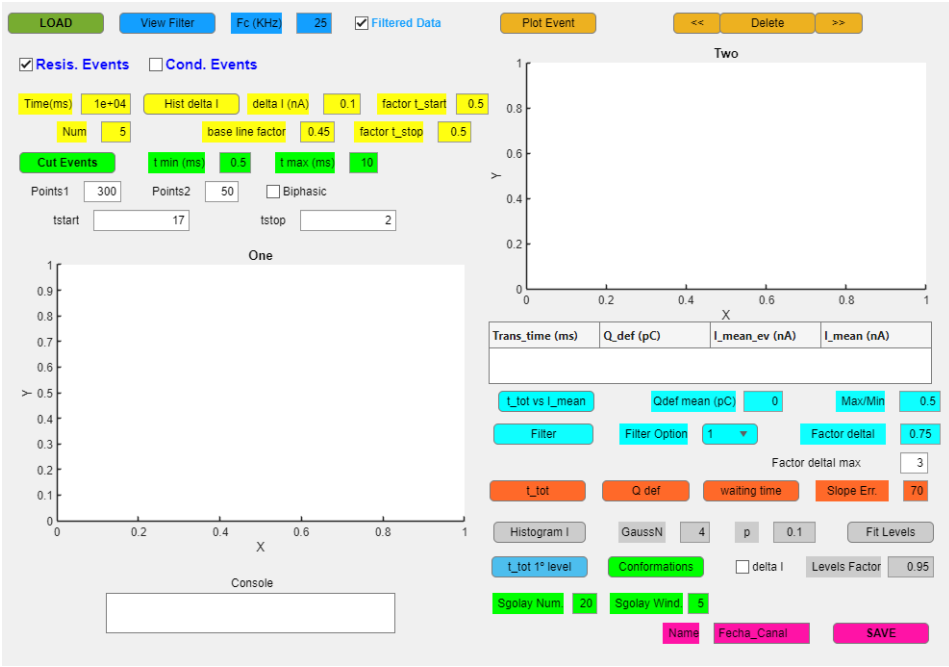


Figure D.2: MATLAB App GUI for analyzing DNA translocation data. More details about the app's analysis can be found in the text.

5. Event Verification: The app allows for visual inspection of individual translocation events to verify their correct identification, ensuring the analysis is accurate.

6. Parameter Calculation: For each translocation event, parameters such

as dwell time, mean current drop, and charge blockade are calculated from the current trace of each event.

7. Event Filtering: Based on these parameters, the app filters events to identify those corresponding to specific molecule sizes and distinguishes them from fragments or multi-molecule translocations.

8. Mean Statistics: The app calculates mean statistics from the selected events, including mean dwell time, mean charge blockade, the mean current drop for different levels, and translocation rate.

9. MATLAB Structure Creation:

    - A MATLAB structure is created to store all translocation events, including the raw current traces, baseline-subtracted traces, absolute time of occurrence, dwell times, mean current drops, charge blockade, and elapsed time since the previous event.
    - A second structure is generated to store key mean parameters, including dwell times, charge blockades, current drops for different levels, and translocation rates. The different parameters used in the analysis are also saved in this structure.

10. Data Saving: Both structures are saved for further analysis.

Appendix E

Appendix Chapter 4

This Appendix contains extra figures about the results in Chapter 6. Additionally, tables summarizing the analyzed data are provided.

E.1 Simulations of λ-DNA using the Worm like chain model

In Chapter 6, simulations of λ-DNA conformations were performed. Here, information on how the simulations are performed is provided. The Worm-like chain model was described in Sec. 3.1.2, where it was explained that this model considers DNA as a thin, inextensible, and continuously flexible rod. For the simulations, λ-DNA is divided into segments of $L_s = 1$ nm. All the simulations start at the origin, and the position of the segments is calculated iteratively using the equations:

$$\begin{aligned} x(i) &= x(i-1) + L_s \cdot sin(\phi)cos(\theta) \\ y(i) &= y(i-1) + L_s \cdot sin(\phi)sin(\theta) \\ z(i) &= z(i-1) + L_s \cdot cos(\phi) \end{aligned} \tag{E.1}$$

where values of the angles θ and ϕ is updated at each step by:

$$\begin{aligned} \theta &= \theta + d\theta \\ \phi &= \phi + d\phi \end{aligned} \tag{E.2}$$

where $d\phi$ and $d\theta$ are randomly picked each step from a normalized Gaussian distribution center at zero ($\mu = 0$) and with a standard deviation $\sigma = \kappa \cdot \sqrt{L_s/P}$, where P is the persistence length. $\kappa = 1.2$ is a parameter that is adjusted so that the correlation between tangent vectors satisfies $\left\langle \hat{t}(x) \cdot \hat{t}(0) \right\rangle = e^{-x/P}$. The tangent vectors are the vector going from $\{x(i), y(i), z(i)\}$ to $\{x(i+1), y(i+1), z(i+1)\}$

In Fig. E.1, the correlation of tangent vectors ($\langle \hat{t}(x) \cdot \hat{t}(0) \rangle$) averaged over 25 simulations is shown for two different persistence length values, $P = 50\ nm$ in a, and $P = 150\ nm$ in b. With a red continuous line, the function $e^{-x/P}$ is shown, and it can be seen that it perfectly reproduces the correlation between tangent vectors.

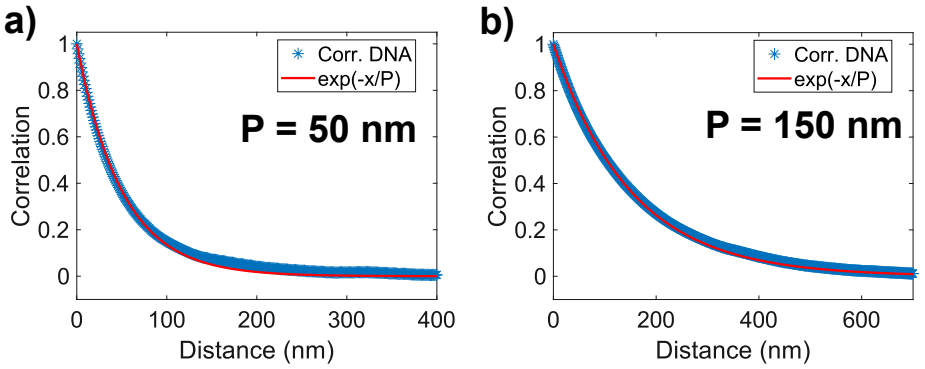


Figure E.1: Average over 25 simulations of the correlation between tangent vectors. In a) the simulation is performed with a persistence length of $P = 50\ nm$, in b) $P = 150\ nm$.

The simulated λ-DNA conformations are shown in Fig. E.2, projected onto the x/y plane. From the figure, it can be observed how the conformations corresponding to a larger persistence length ($P = 150\ nm$) are more extended, and therefore, DNA is less densely packed than for the case of $P = 50\ nm$. At the bottom left of the plots, a 1000 nm-long nanopipette with a tip diameter of 30 nm is shown for scale comparison with DNA conformations.

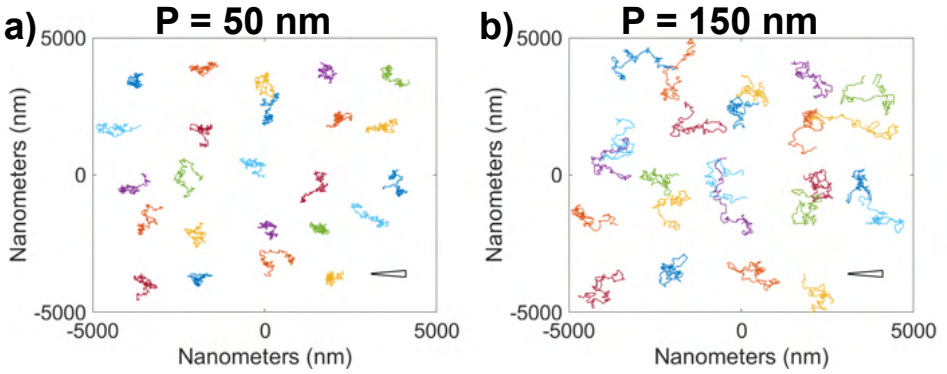


Figure E.2: Projection on the x/y plane of 24 λ-DNA conformations, simulated with a persistence length of $P = 50\ nm$ in a and $P = 150\ nm$ in b.

E.2 Effect of voltage on levels residence times and level transitions

In Sec. 6.3, the results of the levels analysis of a 22 nm diameter nanopipette in 2 M NaCl for 300 and 900 mV were presented. For each figure shown in Sec. 6.3, we provide the results for 500 and 700 mV to confirm voltage independence.

- 2D-histogram of current blockade vs. dwell time

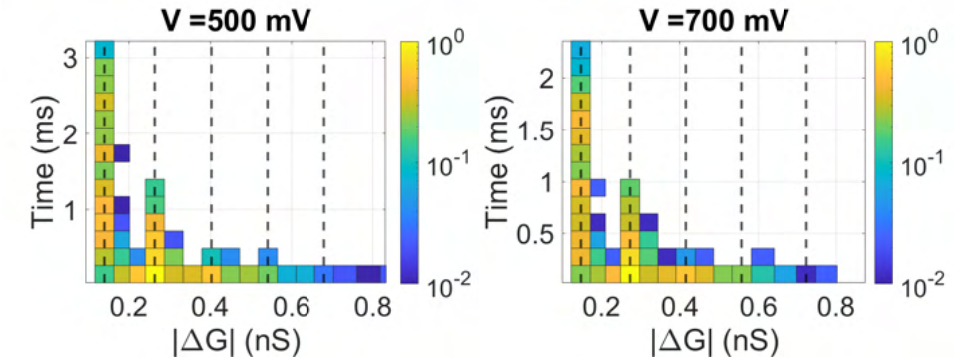


Figure E.3: 2D-histogram of current blockade vs. dwell time of all calculated levels for a 500 mV (left) and 700 mV (right) voltage. Data are from a nanopipette with a 22 nm diameter and 2 M NaCl. The histograms have been normalized so that the scale is from $0-1$, and the color is in a log scale. The mean dwell times are 1.98 ms (500 mV) and 1.44 (700 mV).

- Levels residence time histograms.

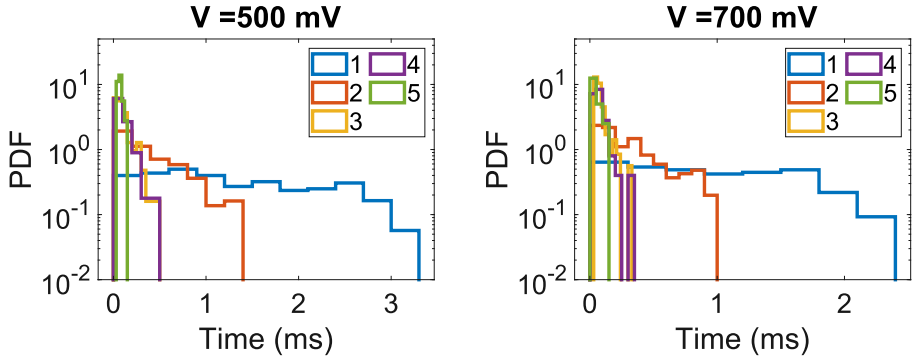


Figure E.4: **Residence time distributions.** Residence time distributions of different levels for 500 mV a), and 700 mV b). Data are from a nanopipette with a 22 nm diameter and 2 M NaCl.

- Transition Matrices.

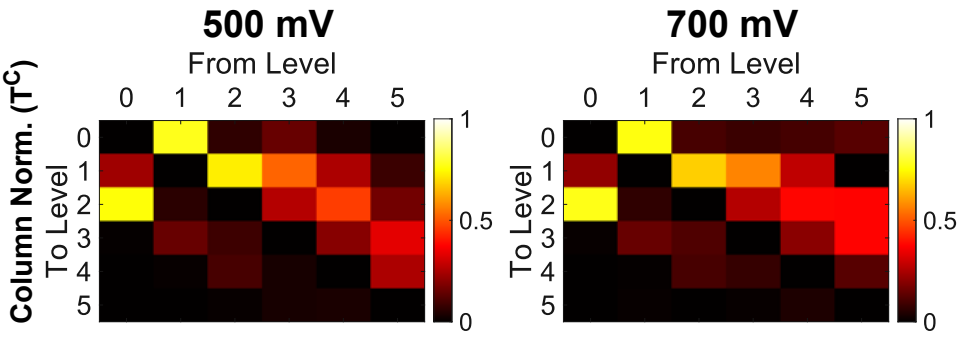


Figure E.5: Transition matrix ($\boldsymbol{T^C}$) for 500 and 700 mV. Data are from a nanopipette with a 22 nm diameter and 2 M NaCl.

E.3 Summary of analyzed data

The levels or configuration analysis described in Sec. 6.2, in Chapter 6, yields a lot of information that can be presented in various manners, as we have seen in the previous section. However, when comparing the results of different nanopipettes, it is easier to condense the analysis into data that can be visualized in a table. The following tables summarize the configuration analysis for λ-DNA translocation. The tables contain relevant information such as salt type, concentration, nanopipette tip diameter, voltage, number of translocation events, and mean dwell time. Additionally, the parameters that resume the configuration analysis are:

- $\langle$Level$\rangle$**.** This number is the mean level of all the translocation events. For a 0120 translocation, the mean level is 1.5. Hence, it indicates if the translocations are more linear or more compact

- **Yes.** This number indicates the percentage of events that satisfy the statement 'the event code does not contain an even number between two uneven numbers of the code'. Hence, it is an estimation of level misassignment. Closer to 100% indicates fewer levels of misassigned.

- **0210 and 0120.** These percentages indicate how many elements with configurations 0210 and 0120 have been detected. This shows the high asymmetry between transitions from levels 2 to 1 and 1 to 2, also observed from the transition matrices.

- **Over** 3° **and** 4°**.** These numbers indicate the percentage of translocation events that have current blockades higher than the mean current

blockade estimated for the 3° and 4° levels. The errors are calculated assuming an error of ±3% in the detrition of the mean current blockade of the 3° and 4° levels.

- **Knot probability.** This number indicates the probability of a translocation event containing a knot. Knots detected as events have codes that satisfy: n_1kn_2 or $n_1k_1k_2n_2$, with $k \geq (n_1 + 2)\&(n_2 + 2)$ or $k_1\&k_2 \geq (n_1 + 2)\&(n_2 + 2)$.

Salt	Conc.	Dia. (nm)	V (mV)	Events	SNR	Dwell (ms)	⟨Level⟩	Yes (%)	0210 (%)	0120 (%)	Over 3º (%)	Over 4º (%)	Knots prob.
NaCl	1	26	1000	771	13.6	0.77	1.73	93.0	42.0	0.2	23±2	6±1	0.20
			900	1135	11.9	0.85	1.74	90.7	44.9	1.1	22±2	5±1	0.20
			700	522	9.0	0.97	1.80	91.8	40.5	0.2	24±2	10±2	0.21
NaCl	1	37	1000	644	12.8	0.96	1.73	91.9	30.2	1.0	21±2	5±1	0.16
			900	742	11.7	1.04	1.74	88.8	26.7	1.8	21±3	4±1	0.17
			700	650	9.3	1.31	1.76	88.6	28.2	1.6	23±2	8±2	0.20
NaCl	2	23	900	325	15.2	1.18	1.83	89.2	50.0	0.7	24±3	7±2	0.23
			700	354	13.6	1.44	1.83	86.2	50.5	0.7	26±3	9±2	0.26
			500	411	9.2	1.98	1.84	85.6	50.9	0.6	27±2	9±2	0.26
			300	208	5.5	3.75	1.75	73.1	55.9	0	28±2	11±2	0.22
NaCl	2	14	900	309	23.8	0.77	1.78	91.9	43.8	1.1	24±2.	8.7±2	0.22
			700	439	19.2	1.00	1.81	88.2	49.6	0.3	26±3	8.0±2	0.26
			500	415	15.4	1.47	1.74	90.6	56.4	0	25±2	7.4±2	0.24
			300	427	9.0	2.47	1.75	93.2	65.1	0.3	20±2	5.9±2	0.19
NaCl	4	15	900	402	23.4	1.56	1.96	86.6	40.8	0.3	40±4	16±3	0.44
			700	409	14.4	2.06	1.93	88.3	45.2	0.3	35±3	13±3	0.40
			500	362	13.9	2.71	1.95	85.6	49.7	0	36±2	14±3	0.42
			300	415	8.9	4.52	1.88	91.8	61.2	0	38±2	12±2	0.30
NaCl	4	27	900	302	13.9	2.34	2.04	83.8	39.9	0	39±3	17±3	0.40
			700	246	12.3	3.02	1.96	78.5	43.0	0	33±3	9±2	0.25
			500	254	9.7	4.50	1.97	87.0	33.5	1.8	33±2	9±2	0.28

Salt	Conc.	Dia. (nm)	V (mV)	Events	SNR	Dwell (ms)	⟨Level⟩	Yes (%)	0210 (%)	0120 (%)	Over 3º (%)	Over 4º (%)	Knots prob.
KCl	1	19	900	591	14.6	0.63	1.68	92.4	44.1	0.7	17±3	3±1	0.13
KCl	1	13	900	209	14.3	0.59	1.53	99.5	63.5	0.5	7±1	1.0±0.5	0.05
KCl	2	26	900	630	0.16	0.89	1.68	94.9	54.2	0.7	14±2	5±1	0.10
KCl	2	18	900	896	21.6	0.70	1.74	93.1	40.4	0.5	16±2	4±1	0.11
KCl	4	20	900	487	19.8	1.08	1.86	92.4	46.7	0.4	26±2	8±2	0.23
KCl	4	36	900	897	12.6	1.48	1.89	87.7	44.2	0.6	29±2	8±2	0.23
LiCl	1	14	900	330	18.2	0.67	1.65	89.4	61.9	0.3	13±2	3±1	0.12
LiCl	1	26	900	426	12.2	1.59	1.68	91.8	57.9	0.3	19±2	4±1	0.16
LiCl	2	37	900	317	9.6	3.32	1.78	89.6	50.4	0.7	23±3	8±2	0.22
LiCl	2	22	900	515	22.9	1.66	1.90	87.4	48.2	0.7	29±3	9±2	0.24
LiCl	4	20	900	123	20.8	5.49	2.42	37.7	15.0	0	79±3	41±2	0.87

Appendix F

A SPRNT (Single-molecule picometer resolution nanopore tweezers) beginners Guide

This section aims to guide anyone wanting to perform their first SPRNT experiments. It will explain experimental protocols, setup/build design, and data analysis work.

F.1 Performing SPRNT experiments

In summary, to do a SPRINT experiment, one must establish a lipid bilayer and then get a single MspA nanopore to insert into it (see Fig.F.1c). Once this is achieved, one can flow the desired enzyme and DNA template for the experiment and apply a voltage to measure the enzyme movement/kinetics on the DNA. Although these sound simple, several experimental protocols need to be followed to achieve this. The protocols will be described in the following sections.

F.1.1 Formation of the lipid membrane

A schematic of the u-tube microfluidic (from now on called plug) where the lipid bilayer is formed is shown in Fig.F.1a and b. The bilayer is formed across the aperture of approx. 20 μm, in the cis-chamber. To do this, one has to paint the lipids onto the aperture and then form the bilayer by bubbling over the aperture to establish a lipid bilayer. The protocol to follow to form a bilayer is explained here:

1. All experiments start with taking a clean plug (plug cleaning is explained in its corresponding section). The plug is then fixed onto the plug holder, and the electrodes are placed into their corresponding chambers.

2. Once the plug is fixed, the u-tube is backfilled with the desired buffer for the experiment. Care must be taken to avoid bubbles in the u-tube, as those may block the aperture during the experiment. After filling the u-tube, both the cis and trans chambers are filled.

3. Next, one needs to prepare the lipids for painting. For these, the lipids are mixed with hexadecane on a cover slide (the preparation and handling of the lipids are explained in the corresponding section). We scrape some lipids off the cover slide and then mix them with hexadecane on a clean spot of the coverslip; a very fine brush with 3–5 bristles is used. The lipid/hexadecane mixture needs to have a consistency similar to toothpaste to be painted onto the aperture.

4. Once one has the lipids at the desired consistency, the brush is used to paint them as uniformly as possible around and over the aperture. A voltage is applied with the Axopatch during the painting process, and the current across the aperture is measured. This allows us to see when the lipids clog the aperture, as the current will drop to zero. By clog, we refer to a chunk of lipids that blocks the aperture, but where the lipids are not forming a bilayer, we cannot get a nanopore insertion into it. Once the lipids/hexadecane mixture has been spread around and over the aperture, the aperture will be clogged at some point, and then one can stop painting.

5. To unclog the aperture, some buffer is back-flowed from the trans chamber through the u-tube. This unclogs the aperture. These can be seen as the Axopatch will measure a current again. To form a bilayer, one needs to form a bubble over the aperture. While doing this, a lipid bilayer is formed on the air-water interface and then deposited on the aperture when the bubble blocks the aperture. A micropipette of $\sim 20\ \mu L$ is used to form the bubbles. After creating the bubble, it is gently sucked into the pipette again. If this is done gently, then the bilayer will remain over the aperture without braking. These can be again observed in the current measurement, as it will droop to zero.

6. To ensure that we have a bilayer and not a clog, we can use the Zap bottom of the Axopatch to apply a quick ± 1 V pulse. As bilayers are more delicate, if we have a bilayer, it will break, and we will measure a current across the aperture again. If we have a clog instead, a ± 1 V

pulse will not break it, and we have to break it by back flowing through the u-tube. To ensure that the bilayer is forming properly, we can establish it by bubbling and then breaking it a few times. If, on the contrary, we get a clog when we make the bubble, it probably means that the lipids have not been properly spread around the aperture or that we have too many lipids. So we need to use the brush to paint again and remove some lipids or spread them around more evenly.

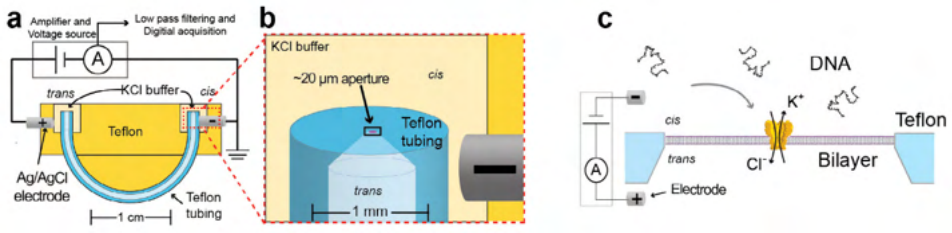


Figure F.1: Schematic of the u-tube bilayer setup at various scales (originally described in [155]. Image from [20].

F.1.2 Flowing the MspA nanopore to get an insertion

Once we have a bilayer, we have to flow the MspA nanopore into the cis chamber and wait until we get an insertion of an MspA into the membrane. The protocol to achieve this is as follows:

1. We start by pipetting 1 μL of MspA nanopore into the cis chamber.

2. We can monitor for the insertion by looking at the current that the Axopatch is measuring. The current measure should be zero when we just have the bilayer. When a nanopore is inserted, the current will increase to 180 pA or 220 pA (if we are working in a buffer of 0.5 M KCl). The current will be 180pA for a forward nanopore (vestibule oriented to the cis chamber, like in Fig.F.1c) or 220pA for a backward one (vestibule to the trans chamber)

3. To facilitate the pore insertion, it is better to keep breaking and bubbling the bilayer. In this manner, one can get a nanopore insertion faster. If everything is working, one should be nanopore insertions in just a few minutes of rebubbling.

4. Once an insertion has happened, one needs to quickly and carefully exchange the buffer in the cis chamber, as it contains more nanopores that could insert into the membrane. For this, we flow $2-3$ pipettes

of 1 mL of buffer into the cis chamber while using the waste syringe to remove the excess buffer. To avoid breaking the lipid membrane while flowing, we should place the syringe tips under the height of the lipid membrane, as it will break if we flow directly on it.

F.1.3 Running the experiment

Now that we have an Mspa nanopore inserted into the lipid membrane, we can add the DNA template and the molecular motor we want to study. The procedure may differ depending on the molecular motor, as some need to be incubated with the DNA template. Here, I will explain how to perform an experiment with Hel308 that does not require incubation.

1. We first carefully pipette 1 μL of the DNA template over the lipid membrane. The DNA will then fuse into the bilayer thanks to the cholesterol tag at the end of the DNA template. The cholesterol tags on the DNA can be used to increase the number of events, however they are not strictly needed.

2. Once we start to see DNA translocations through the nanopore, we can proceed to flow 1.5 μL of Hel308 into the nanopore.

3. The current needs to be monitored during the whole time during the experiment, as the small Mspa nanopore is prone to get clogged or enter into a gating state. In these cases, reversing the voltage bias for a brief moment usually helps to unclog the pore or to bring it out from the gated state.

4. We should avoid applying voltages over 250 mV during the experiments, as these voltages may break the bilayer.

5. After running the experiment for $\sim$ 30 minutes, one must flow some new ATP and DNA.

Bibliography

[1] Thorsten Ritz, Ana Damjanović, and Klaus Schulten. The quantum physics of photosynthesis. *ChemPhysChem*, 3(3):243–248, 2002.

[2] JD Watoson and FHC Crick. Molecular structure of nucleic acids. *Nature*, 171:737–738, 1953.

[3] Erwin Schrödinger. *What is life?: With mind and matter and autobiographical sketches*. Cambridge University Press, 1992.

[4] Craig Venter and Daniel Cohen. The century of biology. *New Persp. Q.*, 21:73, 2004.

[5] Chirlmin Joo, Hamza Balci, Yuji Ishitsuka, Chittanon Buranachai, and Taekjip Ha. Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.*, 77(1):51–76, 2008.

[6] Hyun-Kyu Choi, Hyun Gyu Kim, Min Ju Shon, and Tae-Young Yoon. High-resolution single-molecule magnetic tweezers. *Annual Review of Biochemistry*, 91(1):33–59, 2022.

[7] Ru-Jia Yu, Yi-Lun Ying, Rui Gao, and Yi-Tao Long. Confined nanopipette sensing: from single molecules, single nanoparticles, to single cells. *Angewandte Chemie International Edition*, 58(12):3706–3714, 2019.

[8] John Stanley and Nader Pourmand. Nanopipettes—the past and the present. *APL Materials*, 8(10), 2020.

[9] Jeffrey Mc Hugh, Stanislaw Makarchuk, Daria Mozheiko, Ana Fernandez-Villegas, Gabriele S Kaminski Schierle, Clemens F Kaminski, Ulrich F Keyser, David Holcman, and Nathalie Rouach. Diversity of dynamic voltage patterns in neuronal dendrites revealed by nanopipette electrophysiology. *Nanoscale*, 15(29):12245–12254, 2023.

[10] Krishna Jayant, Michael Wenzel, Yuki Bando, Jordan P Hamm, Nicola Mandriota, Jake H Rabinowitz, Ilan Jen-La Plante, Jonathan S

Owen, Ozgur Sahin, Kenneth L Shepard, et al. Flexible nanopipettes for minimally invasive intracellular electrophysiology in vivo. *Cell reports*, 26(1):266–278, 2019.

[11] Marshall Don. The coulter principle: foundation of an industry. *JALA: Journal of the Association for Laboratory Automation*, 8(6):72–81, 2003.

[12] Coulter. Means for counting particles suspended in a fluid, 1953.

[13] Ralph W DeBlois, Charles P Bean, and Roy KA Wesley. Electrokinetic measurements with submicron particles and pores by the resistive pulse technique. *Journal of colloid and interface science*, 61(2):323–335, 1977.

[14] Hagan Bayley and Lakmal Jayasinghe. Functional engineered channels and pores. *Molecular membrane biology*, 21(4):209–220, 2004.

[15] Hagan Bayley and Charles R Martin. Resistive-pulse sensing from microbes to molecules. *Chemical reviews*, 100(7):2575–2594, 2000.

[16] Nima Arjmandi, Willem Van Roy, Liesbet Lagae, and Gustaaf Borghs. Measuring the electric charge and zeta potential of nanometer-sized objects using pyramidal-shaped nanopores. *Analytical chemistry*, 84(20):8490–8496, 2012.

[17] Rajesh Kumar Sharma, Ishita Agrawal, Liang Dai, Patrick S Doyle, and Slaven Garaj. Complex dna knots detected with a nanopore sensor. *Nature communications*, 10(1):4473, 2019.

[18] Shaghayegh Agah, Ming Zheng, Matteo Pasquali, and Anatoly B Kolomeisky. Dna sequencing by nanopores: advances and challenges. *Journal of Physics D: Applied Physics*, 49(41):413001, 2016.

[19] Yue Wang, Qiuping Yang, and Zhimin Wang. The evolution of nanopore sequencing. *Frontiers in genetics*, 5:449, 2015.

[20] Ian M Derrington, Jonathan M Craig, Eric Stava, Andrew H Laszlo, Brian C Ross, Henry Brinkerhoff, Ian C Nova, Kenji Doering, Benjamin I Tickman, Mostafa Ronaghi, et al. Subangstrom single-molecule measurements of motor proteins using a nanopore. *Nature biotechnology*, 33(10):1073–1075, 2015.

[21] Yanfang Wu and J Justin Gooding. The application of single molecule nanopore sensing for quantitative analysis. *Chemical Society Reviews*, 51(10):3862–3885, 2022.

[22] Qi Chen and Zewen Liu. Fabrication and applications of solid-state nanopores. *Sensors*, 19(8):1886, 2019.

[23] Lorenz J Steinbock, Roman D Bulushev, Swati Krishnan, Camille Raillon, and Aleksandra Radenovic. Dna translocation through low-noise glass nanopores. *Acs Nano*, 7(12):11255–11262, 2013.

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

[25] Hai-Ling Liu, Kan Zhan, Kang Wang, and Xing-Hua Xia. Recent advances in nanotechnologies combining surface-enhanced raman scattering and nanopore. *TrAC Trends in Analytical Chemistry*, 159:116939, 2023.

[26] Lorenz J Steinbock, Oliver Otto, Catalin Chimerel, Joanne Gornall, and Ulrich F Keyser. Detecting dna folding with nanocapillaries. *Nano letters*, 10(7):2493–2497, 2010.

[27] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[28] Amit Meller, Lucas Nivon, and Daniel Branton. Voltage-driven dna translocations through a nanopore. *Physical review letters*, 86(15):3435, 2001.

[29] Lorenz J Steinbock, Alex Lucas, Oliver Otto, and Ulrich F Keyser. Voltage-driven transport of ions and dna through nanocapillaries. *Electrophoresis*, 33(23):3480–3487, 2012.

[30] Alessio Fragasso, Sonja Schmid, and Cees Dekker. Comparing current noise in biological and solid-state nanopores. *ACS nano*, 14(2):1338–1349, 2020.

[31] FN Hooge. 1/f noise. *Physica B+ c*, 83(1):14–23, 1976.

[32] Pulak Dutta and PM Horn. Low-frequency fluctuations in solids: 1 f noise. *Reviews of Modern physics*, 53(3):497, 1981.

[33] Ralph MM Smeets, UF Keyser, MY Wu, NH Dekker, and Cees Dekker. Nanobubbles in solid-state nanopores. *Physical Review Letters*, 97(8):088101, 2006.

[34] Chenyu Wen, Shuangshuang Zeng, Kai Arstila, Timo Sajavaara, Yu Zhu, Zhen Zhang, and Shi-Li Zhang. Generalized noise study of

solid-state nanopores at low frequencies. *ACS sensors*, 2(2):300–307, 2017.

[35] SJ Heerema, GF Schneider, M Rozemuller, L Vicarelli, HW Zandbergen, and C Dekker. 1/f noise in graphene nanopores. *Nanotechnology*, 26(7):074001, 2015.

[36] Friits N Hooge. 1/f noise is no surface effect. *Physics letters A*, 29(3):139–140, 1969.

[37] Harry Nyquist. Thermal agitation of electric charge in conductors. *Physical review*, 32(1):110, 1928.

[38] John Bertrand Johnson. Thermal agitation of electricity in conductors. *Physical review*, 32(1):97, 1928.

[39] Ya M Blanter and Markus Büttiker. Shot noise in mesoscopic conductors. *Physics reports*, 336(1-2):1–166, 2000.

[40] Jan Dolzer. Patch clamp technology in the twenty-first century. *Patch Clamp Electrophysiology: Methods and Protocols*, pages 21–49, 2021.

[41] E Sackmann, R Lipowsky, et al. Handbook of biological physics. *by Lipowsky R., Sackmann E., Elsevier*, 1:213–304, 1995.

[42] Robert John Hunter, Lee R White, and Derek YC Chan. Foundations of colloid science. *(No Title)*, 1987.

[43] Chenyu Wen, Shuangshuang Zeng, Shiyu Li, Zhen Zhang, and Shi-Li Zhang. On rectification of ionic current in nanopores. *Analytical chemistry*, 91(22):14597–14604, 2019.

[44] Z Siwy, Yuchun Gu, HA Spohr, D Baur, A Wolf-Reber, R Spohr, P Apel, and YE Korchev. Rectification and voltage gating of ion currents in a nanofabricated pore. *Europhysics Letters*, 60(3):349, 2002.

[45] Chang Wei, Allen J Bard, and Stephen W Feldberg. Current rectification at quartz nanopipet electrodes. *Analytical Chemistry*, 69(22):4627–4633, 1997.

[46] Javier Cervera, Birgitta Schiedt, Reinhard Neumann, Salvador Mafé, and Patricio Ramírez. Ionic conduction, rectification, and selectivity in single conical nanopores. *The Journal of chemical physics*, 124(10), 2006.

[47] David M Huang, Cécile Cottin-Bizonne, Christophe Ybert, and Lydéric Bocquet. Ion-specific anomalous electrokinetic effects in hydrophobic nanochannels. *Physical review letters*, 98(17):177801, 2007.

[48] Dominik Horinek and Roland R Netz. Specific ion adsorption at hydrophobic solid surfaces. *Physical review letters*, 99(22):226104, 2007.

[49] James E Hall. Access resistance of a small circular pore. *The Journal of general physiology*, 66(4):531–532, 1975.

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

[51] Jacob N Israelachvili. *Intermolecular and surface forces*. Academic press, 2011.

[52] Raymond M Fuoss. Review of the theory of electrolytic conductance. *Journal of Solution Chemistry*, 7:771–782, 1978.

[53] Kazuko Tanaka and Reita Tamamushi. A physico-chemical study of concentrated aqueous solutions of lithium chloride. *Zeitschrift für Naturforschung A*, 46(1-2):141–147, 1991.

[54] William M Haynes. *CRC handbook of chemistry and physics*. CRC press, 2016.

[55] Aaron Regberg, Kamini Singha, Ming Tien, Flynn Picardal, Quanxing Zheng, Jurgen Schieber, Eric Roden, and Susan L Brantley. Electrical conductivity as an indicator of iron reduction rates in abiotic and biotic systems. *Water resources research*, 47(4), 2011.

[56] Andrew H Laszlo, Ian M Derrington, Brian C Ross, Henry Brinkerhoff, Andrew Adey, Ian C Nova, Jonathan M Craig, Kyle W Langford, Jenny Mae Samson, Riza Daza, et al. Decoding long nanopore sequencing reads of natural dna. *Nature biotechnology*, 32(8):829–833, 2014.

[57] Jiali Li, Marc Gershow, Derek Stein, Eric Brandin, and Jene Andrew Golovchenko. Dna molecules and configurations in a solid-state nanopore microscope. *Nature materials*, 2(9):611–615, 2003.

[58] Roman D Bulushev, Sanjin Marion, Ekaterina Petrova, Sebastian J Davis, Sebastian J Maerkl, and Aleksandra Radenovic. Single molecule localization and discrimination of dna–protein complexes

by controlled translocation through nanocapillaries. *Nano letters*, 16(12):7882–7890, 2016.

[59] Stefan W Kowalczyk, Alexander Y Grosberg, Yitzhak Rabin, and Cees Dekker. Modeling the conductance and dna blockade of solid-state nanopores. *Nanotechnology*, 22(31):315101, 2011.

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

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

[62] Daniel Fologea, James Uplinger, Brian Thomas, David S McNabb, and Jiali Li. Slowing dna translocation in a solid-state nanopore. *Nano letters*, 5(9):1734–1737, 2005.

[63] Christopher M Edmonds, Yeny C Hudiono, Amir G Ahmadi, Peter J Hesketh, and Sankar Nair. Polymer translocation in solid-state nanopores: Dependence of scaling behavior on pore dimensions and applied voltage. *The Journal of chemical physics*, 136(6), 2012.

[64] Rena Akahori, Takanobu Haga, Toshiyuki Hatano, Itaru Yanagi, Takeshi Ohura, Hirotaka Hamamura, Tomio Iwasaki, Takahide Yokoi, and Takashi Anazawa. Slowing single-stranded dna translocation through a solid-state nanopore by decreasing the nanopore diameter. *Nanotechnology*, 25(27):275501, 2014.

[65] Amit Meller, Lucas Nivon, Eric Brandin, Jene Golovchenko, and Daniel Branton. Rapid nanopore discrimination between single polynucleotide molecules. *Proceedings of the National Academy of Sciences*, 97(3):1079–1084, 2000.

[66] Arnold J Storm, JH Chen, HW Zandbergen, and C Dekker. Translocation of double-strand dna through a silicon oxide nanopore. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 71(5):051903, 2005.

[67] John J Kasianowicz, Eric Brandin, Daniel Branton, and David W Deamer. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences*, 93(24):13770–13773, 1996.

[68] Jiali Li and David S Talaga. The distribution of dna translocation times in solid-state nanopores. *Journal of Physics: Condensed Matter*, 22(45):454129, 2010.

[69] Peng Chen, Jiajun Gu, Eric Brandin, Young-Rok Kim, Qiao Wang, and Daniel Branton. Probing single dna molecule transport using fabricated nanopores. *Nano letters*, 4(11):2293–2298, 2004.

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

[71] Alexander Y Grosberg and Yitzhak Rabin. Dna capture into a nanopore: interplay of diffusion and electrohydrodynamics. *The Journal of chemical physics*, 133(16), 2010.

[72] Daniel Fologea, Eric Brandin, James Uplinger, Daniel Branton, and Jiali Li. Dna conformation and base number simultaneously determined in a nanopore. *Electrophoresis*, 28(18):3186–3192, 2007.

[73] Xavier Viader Godoy. *Biophysical properties of single-stranded DNA studied with single-molecule force spectroscopy.* PhD thesis, Universitat de Barcelona, 2021.

[74] Gustav Mie. Beiträge zur optik trüber medien, speziell kolloidaler metallösungen. *Annalen der physik*, 330(3):377–445, 1908.

[75] JP Barton, DR Alexander, and SA Schaub. Theoretical determination of net radiation force and torque for a spherical particle illuminated by a focused laser beam. *Journal of Applied Physics*, 66(10):4594–4602, 1989.

[76] Carlos J Bustamante and Steven B Smith. Light-force sensor and method for measuring axial optical-trap forces from changes in light momentum along an optic axis, November 7 2006. US Patent 7,133,132.

[77] Steven B Smith, Yujia Cui, and Carlos Bustamante. [7] optical-trap force transducer that operates by direct measurement of light momentum. In *Methods in enzymology*, volume 361, pages 134–162. Elsevier, 2003.

[78] Carlos J Bustamante and Steven B Smith. Optical beam translation device and method utilizing a pivoting optical fiber, September 25 2007. US Patent 7,274,451.

[79] Josep M Huguet, Cristiano V Bizarro, Núria Forns, Steven B Smith, Carlos Bustamante, and Felix Ritort. Single-molecule derivation of salt dependent base-pair free energies in dna. *Proceedings of the National Academy of Sciences*, 107(35):15431–15436, 2010.

[80] Elizabeth A Shank, Ciro Cecconi, Jesse W Dill, Susan Marqusee, and Carlos Bustamante. The folding cooperativity of a protein is controlled by its chain topology. *Nature*, 465(7298):637–640, 2010.

[81] I Di Terlizzi, M Gironella, D Herráez-Aguilar, Timo Betz, F Monroy, M Baiesi, and F Ritort. Variance sum rule for entropy production. *Science*, 383(6686):971–976, 2024.

[82] Marta Gironella-Torrent, Giulia Bergamaschi, Raya Sorkin, Gijs JL Wuite, and Felix Ritort. Viscoelastic phenotyping of red blood cells. *Biophysical Journal*, 123(7):770–781, 2024.

[83] Maria Manosas, Delphine Collin, and Felix Ritort. Force-dependent fragility in rna hairpins. *Physical Review Letters*, 96(21):218301, 2006.

[84] Jan Liphardt, Bibiana Onoa, Steven B Smith, Ignacio Tinoco Jr, and Carlos Bustamante. Reversible unfolding of single rna molecules by mechanical force. *Science*, 292(5517):733–737, 2001.

[85] Douglas E Smith, Sander J Tans, Steven B Smith, Shelley Grimes, Dwight L Anderson, and Carlos Bustamante. The bacteriophage $\varphi 29$ portal motor can package dna against a large internal force. *Nature*, 413(6857):748–752, 2001.

[86] Sara De Lorenzo, Marco Ribezzi-Crivellari, J Ricardo Arias-Gonzalez, Steven B Smith, and Felix Ritort. A temperature-jump optical trap for single-molecule manipulation. *Biophysical journal*, 108(12):2854–2864, 2015.

[87] Álvaro Martínez Monge. *Free energy and information-content measurements in thermodynamic and molecular ensembles*. PhD thesis, Universitat de Barcelona, 2019.

[88] Jan Gieseler, Juan Ruben Gomez-Solano, Alessandro Magazzù, Isaac Pérez Castillo, Laura Pérez García, Marta Gironella-Torrent, Xavier Viader-Godoy, Felix Ritort, Giuseppe Pesce, Alejandro V Arzola, et al. Optical tweezers—from calibration to applications: a tutorial. *Advances in Optics and Photonics*, 13(1):74–241, 2021.

[89] Jay R Wenner, Mark C Williams, Ioulia Rouzina, and Victor A Bloomfield. Salt dependence of the elasticity and overstretching transition of single dna molecules. *Biophysical journal*, 82(6):3160–3169, 2002.

[90] Mark C Williams, Jay R Wenner, Ioulia Rouzina, and Victor A Bloomfield. Entropy and heat capacity of dna melting from temperature dependence of single molecule stretching. *Biophysical Journal*, 80(4):1932–1939, 2001.

[91] Josep Maria Huguet Casades. *Statistical and thermodynamical properties of DNA unzipping experiments with optical tweezers*. PhD thesis, Universitat de Barcelona, 2011.

[92] R.R. Sinden. *DNA Structure and Function*. Academic Press, 2012.

[93] Stephen Neidle and Mark Sanderson. *Principles of nucleic acid structure*. Academic Press, 2021.

[94] Pui S Ho, Michael J Ellison, Gary J Quigley, and Alexander Rich. A computer aided thermodynamic approach for predicting the formation of z-dna in naturally occurring sequences. *The EMBO journal*, 5(10):2737–2744, 1986.

[95] Michael J Ellison, Juli Feigon, Raymond J Kelleher III, Andrew HJ Wang, Joel F Habener, and Alexander Rich. An assessment of the z-dna forming potential of alternating da-dt stretches in supercoiled plasmids. *Biochemistry*, 25(12):3648–3655, 1986.

[96] Fritz M Pohl and Thomas M Jovin. Salt-induced co-operative conformational change of a synthetic dna: equilibrium and kinetic studies with poly (dg-dc). *Journal of molecular biology*, 67(3):375–396, 1972.

[97] Sung Chul Ha, Ky Lowenhaupt, Alexander Rich, Yang-Gyun Kim, and Kyeong Kyu Kim. Crystal structure of a junction between b-dna and z-dna reveals two extruded bases. *Nature*, 437(7062):1183–1186, 2005.

[98] Vinod Kumar Subramani, Subramaniyam Ravichandran, Varun Bansal, and Kyeong Kyu Kim. Chemical-induced formation of bz-junction with base extrusion. *Biochemical and biophysical research communications*, 508(4):1215–1220, 2019.

[99] Michael Rubinstein and Ralph H Colby. *Polymer physics*. Oxford university press, 2003.

[100] Otto Kratky and Günther Porod. Röntgenuntersuchung gelöster fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas*, 68(12):1106–1122, 1949.

[101] Steven B Smith, Yujia Cui, and Carlos Bustamante. Overstretching b-dna: the elastic response of individual double-stranded and single-stranded dna molecules. *Science*, 271(5250):795–799, 1996.

[102] John F Marko and Eric D Siggia. Stretching dna. *Macromolecules*, 28(26):8759–8770, 1995.

[103] Timothy M Lohman and Keith P Bjornson. Mechanisms of helicase-catalyzed dna unwinding. *Annual review of biochemistry*, 65(1):169–214, 1996.

[104] Martin R Singleton, Mark S Dillingham, and Dale B Wigley. Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, 76(1):23–50, 2007.

[105] Jonathan M Craig, Andrew H Laszlo, Ian C Nova, and Jens H Gundlach. Modelling single-molecule kinetics of helicase translocation using high-resolution nanopore tweezers (sprnt). *Essays in biochemistry*, 65(1):109–127, 2021.

[106] Feng Dong, Edward P Gogol, and Peter H von Hippel. The phage t4-coded dna replication helicase (gp41) forms a hexamer upon activation by nucleoside triphosphate. *Journal of Biological Chemistry*, 270(13):7462–7473, 1995.

[107] Sven H Behrens and David G Grier. The charge of glass and silica surfaces. *The Journal of chemical physics*, 115(14):6716–6721, 2001.

[108] Frank HJ van der Heyden, Derek Stein, and Cees Dekker. Streaming currents in a single nanofluidic channel. *Physical review letters*, 95(11):116104, 2005.

[109] Matthew R Powell, Craig Martens, and Zuzanna S Siwy. Asymmetric properties of ion current 1/f noise in conically shaped nanopores. *Chemical Physics*, 375(2-3):529–535, 2010.

[110] Shihao Su, Xun Guo, Yanjun Fu, Yanbo Xie, Xinwei Wang, and Jianming Xue. Origin of nonequilibrium 1/f noise in solid-state nanopores. *Nanoscale*, 12(16):8975–8981, 2020.

[111] FN Hooge, TGM Kleinpenning, and Lode KJ Vandamme. Experimental studies on 1/f noise. *Reports on progress in Physics*, 44(5):479, 1981.

[112] Ralph MM Smeets, Ulrich F Keyser, Nynke H Dekker, and Cees Dekker. Noise in solid-state nanopores. *Proceedings of the National Academy of Sciences*, 105(2):417–421, 2008.

[113] Ulrich F Keyser, Bernard N Koeleman, Stijn Van Dorp, Diego Krapf, Ralph MM Smeets, Serge G Lemay, Nynke H Dekker, and Cees Dekker. Direct force measurements on dna in a solid-state nanopore. *Nature Physics*, 2(7):473–477, 2006.

[114] ROBERT WILFRED Wilson, DONALD C Rau, and VICTOR A Bloomfield. Comparison of polyelectrolyte theories of the binding of cations to dna. *Biophysical journal*, 30(2):317–325, 1980.

[115] Gerald S Manning. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Quarterly reviews of biophysics*, 11(2):179–246, 1978.

[116] Cees Dekker. Solid-state nanopores. *Nature nanotechnology*, 2(4):209–215, 2007.

[117] Kidan Lee, Kyeong-Beom Park, Hyung-Jun Kim, Jae-Seok Yu, Hongsik Chae, Hyun-Mi Kim, and Ki-Bum Kim. Recent progress in solid-state nanopores. *Advanced materials*, 30(42):1704680, 2018.

[118] Liang Xue, Hirohito Yamazaki, Ren Ren, Meni Wanunu, Aleksandar P Ivanov, and Joshua B Edel. Solid-state nanopore sensors. *Nature Reviews Materials*, 5(12):931–951, 2020.

[119] Kaikai Chen, Nicholas AW Bell, Jinglin Kong, Yu Tian, and Ulrich F Keyser. Direction-and salt-dependent ionic current signatures for dna sensing with asymmetric nanopores. *Biophysical journal*, 112(4):674–682, 2017.

[120] Sergio Cruz-León, Willem Vanderlinden, Peter Müller, Tobias Forster, Georgina Staudt, Yi-Yun Lin, Jan Lipfert, and Nadine Schwierz. Twisting dna by salt. *Nucleic acids research*, 50(10):5726–5738, 2022.

[121] Stefan Kesselheim, Wojciech Müller, and Christian Holm. Origin of current blockades in nanopore translocation experiments. *Physical review letters*, 112(1):018101, 2014.

[122] Yong Woon Kim and Roland R Netz. Electro-osmosis at inhomogeneous charged surfaces: Hydrodynamic versus electric friction. *The Journal of chemical physics*, 124(11), 2006.

[123] Kaikai Chen, Ining Jou, Niklas Ermann, Murugappan Muthukumar, Ulrich F Keyser, and Nicholas AW Bell. Dynamics of driven polymer transport through a nanopore. *Nature Physics*, 17(9):1043–1049, 2021.

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

[125] VM M LOBO and JL Quaresma. Handbook of electrolyte solutions. b. *Physical sciences data*, 41, 1989.

[126] Sergio Cruz-León and Nadine Schwierz. Hofmeister series for metal-cation–rna interactions: the interplay of binding affinity and exchange kinetics. *Langmuir*, 36(21):5979–5989, 2020.

[127] Sébastien Guilbaud, Laurence Salomé, Nicolas Destainville, Manoel Manghi, and Catherine Tardin. Dependence of dna persistence length on ionic strength and ion type. *Physical review letters*, 122(2):028102, 2019.

[128] Sumanth Kumar Maheshwaram, Koushik Sreenivasa, and Gautam Vivek Soni. Fingerprinting branches on supercoiled plasmid dna using quartz nanocapillaries. *Nanoscale*, 13(1):320–331, 2021.

[129] Harshwardhan H Katkar and Murugappan Muthukumar. Role of non-equilibrium conformations on driven polymer translocation. *The Journal of chemical physics*, 148(2), 2018.

[130] David K Lubensky and David R Nelson. Driven polymer translocation through a narrow pore. *Biophysical journal*, 77(4):1824–1838, 1999.

[131] Daniel Y Ling and Xinsheng Sean Ling. On the distribution of dna translocation times in solid-state nanopores: an analysis using schrödinger's first-passage-time theory. *Journal of Physics: Condensed Matter*, 25(37):375102, 2013.

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

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

[134] Josep Maria Huguet, Marco Ribezzi-Crivellari, Cristiano Valim Bizarro, and Felix Ritort. Derivation of nearest-neighbor dna parameters in magnesium from single molecule experiments. *Nucleic acids research*, 45(22):12921–12931, 2017.

[135] Paolo Rissone. *Thermodynamics and Kinetics of Nucleic Acids Folding*. PhD thesis, Universitat de Barcelona, 2023.

[136] Aurelien Severino, Alvaro Martinez Monge, Paolo Rissone, and Felix Ritort. Efficient methods for determining folding free energies in single-molecule pulling experiments. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(12):124001, 2019.

[137] Peter Yakovchuk, Ekaterina Protozanova, and Maxim D Frank-Kamenetskii. Base-stacking and base-pairing contributions into thermal stability of the dna double helix. *Nucleic acids research*, 34(2):564–574, 2006.

[138] Mina Lee, Sook Ho Kim, and Seok-Cheol Hong. Minute negative superhelicity is sufficient to induce the bz transition in the presence of low tension. *Proceedings of the National Academy of Sciences*, 107(11):4985–4990, 2010.

[139] Alessandro Bosco, Joan Camunas-Soler, and Felix Ritort. Elastic properties and secondary structure formation of single-stranded dna at monovalent and divalent salt conditions. *Nucleic acids research*, 42(3):2064–2074, 2014.

[140] David B Haniford and David E Pulleyblank. The in-vivo occurrence of z dna. *Journal of Biomolecular Structure and Dynamics*, 1(3):593–609, 1983.

[141] Fei Zheng, Antonio Suma, Christopher Maffeo, Kaikai Chen, Mohammed Alawami, Jingjie Sha, Aleksei Aksimentiev, Cristian Micheletti, and Ulrich F Keyser. When knots are plectonemes. *arXiv preprint arXiv:2407.16290*, 2024.

[142] Christopher Maffeo, Lauren Quednau, James Wilson, and Aleksei Aksimentiev. Dna double helix, a tiny electromotor. *Nature Nanotechnology*, 18(3):238–242, 2023.

[143] Elizabeth A Manrao, Ian M Derrington, Andrew H Laszlo, Kyle W Langford, Matthew K Hopper, Nathaniel Gillgren, Mikhail Pavlenok, Michael Niederweis, and Jens H Gundlach. Reading dna at single-nucleotide resolution with a mutant mspa nanopore and phi29 dna polymerase. *Nature biotechnology*, 30(4):349–353, 2012.

[144] Matthew T Noakes, Henry Brinkerhoff, Andrew H Laszlo, Ian M Derrington, Kyle W Langford, Jonathan W Mount, Jasmine L Bowman, Katherine S Baker, Kenji M Doering, Benjamin I Tickman, et al. Increasing the accuracy of nanopore dna sequencing using a time-varying cross membrane voltage. *Nature biotechnology*, 37(6):651–656, 2019.

[145] Miten Jain, Ian T Fiddes, Karen H Miga, Hugh E Olsen, Benedict Paten, and Mark Akeson. Improved data analysis for the minion nanopore sequencer. *Nature methods*, 12(4):351–356, 2015.

[146] Tom Z Butler, Mikhail Pavlenok, Ian M Derrington, Michael Niederweis, and Jens H Gundlach. Single-molecule dna detection with an engineered mspa protein nanopore. *Proceedings of the National Academy of Sciences*, 105(52):20647–20652, 2008.

[147] Matthew H Larson, Rachel A Mooney, Jason M Peters, Tricia Windgassen, Dhananjaya Nayak, Carol A Gross, Steven M Block, William J Greenleaf, Robert Landick, and Jonathan S Weissman. A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science*, 344(6187):1042–1047, 2014.

[148] Irina O Vvedenskaya, Hanif Vahedian-Movahed, Jeremy G Bird, Jared G Knoblauch, Seth R Goldman, Yu Zhang, Richard H Ebright, and Bryce E Nickels. Interactions between rna polymerase and the "core recognition element" counteract pausing. *Science*, 344(6189):1285–1289, 2014.

[149] Andrew H Laszlo, Ian M Derrington, and Jens H Gundlach. Mspa nanopore as a single-molecule tool: From sequencing to sprnt. *Methods*, 105:75–89, 2016.

[150] Jonathan Michael Craig. *High Resolution Single-Molecule Enzyme Dynamics Using Nanopores*. PhD thesis, University of Washintong, 2017.

[151] M Venkatesan, LL Silver, and NG Nossal. Bacteriophage t4 gene 41 protein, required for the synthesis of rna primers, is also a dna helicase. *Journal of Biological Chemistry*, 257(20):12426–12434, 1982.

[152] Zhiquan Zhang, Michelle M Spiering, Michael A Trakselis, Faoud T Ishmael, Jun Xi, Stephen J Benkovic, and Gordon G Hammes. Assembly of the bacteriophage t4 primosome: single-molecule and ensemble studies. *Proceedings of the National Academy of Sciences*, 102(9):3254–3259, 2005.

[153] Maria Manosas, Xu Guang Xi, David Bensimon, and Vincent Croquette. Active and passive mechanisms of helicases. *Nucleic acids research*, 38(16):5518–5526, 2010.

[154] Laser Based Micropipette. *P2000 Operation Manual*. Sutter Instrument, 2010.

[155] Mark Akeson, Daniel Branton, John J Kasianowicz, Eric Brandin, and David W Deamer. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single rna molecules. *Biophysical journal*, 77(6):3227–3233, 1999.