

Discrimination and Sequencing of Polymers with Biological Nanopores

Mordjane Boukhet

▶ To cite this version:

Mordjane Boukhet. Discrimination and Sequencing of Polymers with Biological Nanopores. Materials Science [cond-mat.mtrl-sci]. Université de Cergy Pontoise; Albert-Ludwigs-Universität (Freiburg im Breisgau, Allemagne), 2018. English. NNT: 2018CERG0984. tel-02311628

HAL Id: tel-02311628 https://theses.hal.science/tel-02311628

Submitted on 11 Oct 2019 $\,$

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





Discrimination and Sequencing of Polymers with Biological Nanopores

Dissertation to obtain a joint doctoral degree in Physics of the Albert-Ludwig-Universität Freiburg Germany and the Université Cergy-Pontoise France

> Mordjane Boukhet September 2018

Defended on the 19th of November 2018

Supervisors:	Prof. Günter Reiter	
	Prof. Juan Pelta	
First Reviewer:	Prof. Günter Reiter	
Second Reviewer:	Prof. Juan Pelta	
Third Reviewer:	Prof. Mathias Winterhalter	
Examiners:	Prof. Joachim Dzubiella	
	Dr. Laurent Bacri	
	Prof. Jan C. Behrends	
	Dr. Abdelghani Oukhaled	

Abstract

The technique of detection with nanopores at the single molecule level, is one of the most powerful method for the analysis of various molecules, of which biological and synthetic polymers, proteins and peptides, sugar molecules or metal nanoparticles. These pores can also serve as a platform for the study of fundamental physical and biological phenomenons. In the context of molecule analysis, this work, which is experimented using the technique of planar lipid bilayer painting, focuses mainly on the detection of polymers and their utility to portray fundamental processes of the α -hemolysin and aerolysin biological nanopores.

The first results chapter described the probing of flows through α -hemolysin and aerolysin using polyethylene glycols (PEGs) and α -cyclodextrines, and the effects of KCl and LiCl salts on the interaction of PEGs with these pores. One main finding was that there exists a stronger electoosmotic flow in aerolysin, responsible for the transport of the neutral molecules α cyclodextrines. The second finding was that the dynamics of PEGs with the nanopores are strongly dependent on the salt, showing drastic differences of frequency and dwell times *vs.* voltage for the two salts, although, the results of detection of mass of PEGs pointed to the fact that the nature of the interaction with the pore is similar in both salts.

The aim of the work presented in the second results chapter, was to detect precision polymers, and find the best conditions, which can lead to their sequencing with nanopores. The homo- an copolymers of poly(phosphodiester)s were probed using α -hemolysin, aerolysin and MspA. The first type of polymers investigated which contained a 3-polythymidine primer and a sequence of comonomers of type (0) showed a strong interaction with the pores that was interpreted as the promotion of ssDNA-primer to the binding with the pore, combined to a high flexibility of the first type of polymers. The polymers which contained alkyne and triazole side chains, were found to have more complex interactions, but interacted for shorter durations with the pore indicating them to be stiffer. The second type of polymers seemed to be clustering in solution due the interaction between side chains, which proved the importance of performing characterization of these molecules in solution using wave scattering in the context of detection and ultimately sequencing.

The study of the third result chapter, focused on the dynamics of small oligonucleotides with the aerolysin pore. The interaction of polyadenines (A3, A4, A5) showed complex dynamics and

kinetics with pore, which was investigated via analysis of the events pattern. The whole process was found to be governed by two binding sites and energy barriers inside the pore that the molecules have to overcome. These results were combined to a developed kinetic model which allowed a complete description of the binding and translocation (or failure of it) of these polyadenines.

The last results chapter described the interaction of bigger polyadenines (A6-A7-A8-A9-A10) with the aerolysin nanopore. The analysis of amplitude of currents of the adenine-induced blocks inside this pore showed an orientation dependent interaction of the molecules with the pore. This orientation dependent interaction started to be apparent for the A7 molecule and became the dominant effect for A9 and A10. Due to the flexibility of ssDNA, this effect is not observed for smaller sized molecules (A6 and below) because of their possibility of reorientation while inside the pore.

Resumé

La technique de détection à l'aide de nanopores au niveau de la molécule unique est l'une des plus puissantes pour l'analyse de diverses molécules, dont les polymères biologiques et synthétiques, les protéines et les peptides, les molécules de sucre ou les nanoparticules métalliques. Ces pores peuvent également servir de plate-forme pour l'étude de phénomènes physiques et biologiques fondamentaux. Dans le cadre de l'analyse de molécules, ce travail, expérimenté en utilisant la technique de la peinture de bicouche lipidique, porte principalement sur la détection des polymères et leur utilité pour sonder les processus fondamentaux des de l' α -hémolysine et de l'aérolysine.

Le premier chapitre de résultats décrit l'analyse des flux à travers l'hémolysine et l'aérolysine à l'aide des polyéthylèneglycols (PEG) et des α -cyclodextrines, ainsi que les effets des sels de KCl et de LiCl sur l'interaction des PEG avec ces pores. L'une des principales conclusions est qu'il existe un flux électoosmotique plus fort dans l'aérolysine, responsable du transport des molécules neutres, les α -cyclodextrines. La seconde constatation concerne la dynamique des PEG avec les nanopores qui semblait être fortement dépendante du sel, montrant des différences drastiques de fréquence et de durée d'interaction en fonction de la tension pour les deux sels, bien que la détection de la masse de PEG dans les deux conditions indique que la nature de l'interaction avec le pore est similaire dans les deux types de sels.

Le but des travaux présentés dans le deuxième chapitre de résultats était de détecter les polymères de précision et à trouver les meilleures conditions pouvant conduire à leur séquençage avec des nanopores. Des homopolymères et copolymères de poly(phosphodiester)s ont été sondés en utilisant l'hémolysine, l'aérolysine et MspA. Le premier type de polymères étudiés contenant une amorce 3-polythymidine et une suite de comonomères de type (0) a montré une forte interaction avec les pores qui a été interprétée comme la promotion de la liaison avec les pores, due à l'amorce d'ADN simple brin, combinée à une grande flexibilité du premier type de polymères. Les polymères qui contenaient des chaînes latérales alcyne et triazole se sont révélés avoir des interactions plus complexes, mais ont interagi pendant des durées plus courtes avec les pores indiquant qu'ils étaient plus rigides. Le second type de polymères semble s'agréger en solution du fait de l'interaction entre les chaînes latérales, ce qui prouve l'importance de la caractérisation de ces molécules en solution par diffusion de rayons, dans le cadre de la détection et finalement de leur séquençage.

L'étude du troisième chapitre de résultats, a porté sur la dynamique de petits oligonucléotides avec le pore d'aerolysine. Les polyadénines (A3, A4, A5) ont montré une dynamique complexe d'interaction avec le pore, qui a été étudiée par l'analyse et la quantification des différentes propriétés des événements. L'ensemble du processus s'est avéré être régi par deux sites de liaison et des barrières énergétiques à l'intérieur du pore que les molécules doivent surmonter. Ces résultats ont été combinés à un modèle cinétique qui a permis une description complète de la liaison et de la translocation (ou son non succès) des polyadénines.

Le dernier chapitre des résultats décrit l'interaction de plus grandes polyadénines (A6-A7-A8-A9-A10) avec l'aérolysine. L'analyse de l'amplitude des courants des blocs induits par l'adénine à l'intérieur de ce pore montre une interaction dépendante de l'orientation des molécules avec le pore. Cette interaction dépendante de l'orientation a commencé à apparaître pour la molécule A7 et est devenue l'effet dominant pour A9 et A10. En raison de la flexibilité de l'ANDsb, cet effet n'est pas observé pour les molécules de plus petite taille (A6 et inférieures) en raison de leur possibilité de réorientation à l'intérieur du pore.

Zusammenfassung

Der Nachweis von Einzelmolekülen mit Nanoporen ist eine der besten Methoden zur Analyse verschiedener Moleküle, darunter befinden sich biologische und synthetische Polymere, Proteine, Peptide, Kohlenhydrate oder Metallnanopartikel. Diese Poren können auch als Plattform für das Studium grundlegender physikalischer und biologischer Phänomene dienen. Auf Basis der Molekülanalyse konzentriert sich diese Arbeit, die von der planaren Lipiddoppelschichten durchgeführt wurde, hauptsächlich auf die Analyse von Polymeren und ihren Nutzen, um grundlegende Funktionsweisen der biologischen Nanoporen α -Hämolysin und Aerolysin zu untersuchen.

Das erste Kapitel beschreibt die Interaktion von PEGs und α -Cyclodextrinen sowohl mit α -Hämolysin als auch mit Aerolysin, sowie den Einfluß von KCl- und LiCl-Salzen. Ein Ergebnis dabei war, dass es in Aerolysin einen stärkeren elektoosmotischen Fluss gibt, der für den Transport der neutralen α -Cyclodextrin Moleküle verantwortlich gemacht wird.

Als zweiter Befund ergab sich, dass die Dynamik der PEG Moleküle mit den Nanoporen stark vom Typ des verwendeten Salzes abhängig ist. Dies hat drastische Unterschiede in Frequenz und Verweilzeit gegenüber der elektrischen Klemmspannung zur Folge. Jedoch zeigen die Ergebnisse der Massendetektion von PEGs darauf hin, dass die Wechselwirkung mit der Pore bei beiden Salztypen ähnlich ist. Des Weiteren wurde die Identifikation optimaler Präzisionspolymeren Sequenzierungsbedingungen von mit Nanoporen verfolgt. Die Homocopolymere von Poly-(phosphodiestern) wurden mit Hilfe von a-Hämolysin, Aerolysin und MspA analysiert. Der erste Typ der untersuchten Polymere, welcher aus einem 3-Polythymidin-Primer und Comonomeren des Typs 0 besteht, zeigte eine starke Wechselwirkung mit den biologischen Nanoporen. Dies wird als eine Verbesserung der Bindung mit der Pore durch den ssDNS-Primer angesehen.

Die Polymere, die Alkin- oder Triazolseitenketten enthielten, zeigten komplexere Wechselwirkungen. Sie interagierten für kürzere Zeitspannen mit der Pore, was auf eine steifere Polymerstruktur hindeutet. Der zweite Typ von Polymeren schien sich aufgrund der Wechselwirkungen von einzelnen Seitenketten untereinander in Lösung zusammen zu lagern. Dies betont die Bedeutung der, der Sequenzierung vorangehenden, Charakterisierung des Verhaltens individueller Moleküle in Lösung. Das dritte Ergebniskapitel konzentriert sich auf die Dynamik von kleinen Oligonukleotiden mit der Aerolysin-Pore. Die Wechselwirkung von Polyadeninen (A3, A4, A5) zeigte eine komplexe Wechselwirkung und Dynamik mit der Pore, die durch Analyse des Ereignismusters untersucht wurde. Der gesamte Prozess wurde durch zwei Bindungsstellen und eine Energiebarriere in der Pore bestimmt, die die Moleküle überwinden müssen. Aus diesen Ergebnissen wurde ein kinetisches Modell entwickelt, das eine vollständige Beschreibung der Bindung und Translokation (oder des Versagens davon) von Polyadeninen ermöglicht*e*.

Das letzte Ergebniskapitel beschreibt die Wechselwirkung von größeren Polyadeninen (A6-A7-A8-A9-A10) mit der Aerolysin-Nanopore. Die Stromamplitude der Adenin-induzierten Blockierung innerhalb dieser *Pore* zeigte eine orientierungsabhängige Wechselwirkung der Moleküle mit der *Pore*. Diese orientierungsabhängige Wechselwirkung zeigte sich erstmals mit dem A7-Molekül und wurde zum dominierenden Effekt bei den A9 und A10 Molekülen. Aufgrund der Flexibilität der ssDNA wird dieser Effekt für kleinere Moleküle (A6 und darunter) wegen ihrer Fähigkeit zur Reorientierung innerhalb der Pore nicht beobachtet.

To the memory of Loïc Auvray Who passed away too soon For all his kindness and help For everything, thank you

Content

CH	APTER 1: INTRODUCTION, THEORY AND PREVIOUS STUDIES	1
I.	Physics of polymers an confinement effects	
1.	Conformation of polymers in solution	2
2.	Polymers in Confinement and Dynamics	5
II.	Previous Experiments	16
1.	Experiences with neutral polymers	16
2.	Experiments with DNA	
III.	Summary	
CH	APTER 2: BIOLOGICAL PORES AND ION CHANNELS	27
I.	Comparison of the different existing nanopores for molecule sensing	29
II.	Pore forming toxins	31
1.	Alpha-Hemolysin (α-HL)	32
2.	Aerolysin	
III.	Mycobacterial porins	40
1.	MspA	41
IV.	Summary	43
CH	APTER 3: EXPERIMENTAL METHODS	45
I.	Lipid types and self-assembly	45
II.	Planar lipid bilayer painting and methods of current measurement	49
1.	Classical measurements	51
2.	Recent method of lipid painting: measurement with MECA-chips	57
III.	Nanopore insertion into membranes, analyte interaction	63
IV.	Data analysis	64
1.	Igor-Pro algorithm	65
2.	Labview-based software: DetectEvent	66
3.	Current histograms and characteristic dwell time analysis	67
CH	APTER 4: INTERACTION OF PEGS AND CYCLODEXTRINS WITH TWO	
PRO	OTEIN CHANNELS	71
I.	Analytes : Polyethylene Glycol and Cyclodextrines	
1.	Polyethylene Glycol	
2.	Cyclodextrins	73
п.	Results	75
1.	Interaction of cyclodextrin and aerolysin (AeL)	75

2	. Measurement of ion selectivity	79
3	. Interaction of PEGs and aerolysin	84
4	. Interaction of PEGs and α-hemolysin	87
5	. Effect of salt on the discrimination of PEGs in aerolysin	92
Ш.	Summary	95
СН	IAPTER 5: POLY(PHOSPHODIESTER)S DETECTION AND CHARACTERIZA	TION 97
I.	Sequence controlled polymers	98
II.	Scattering Techniques for characterization	100
1	. Dynamic Light Scattering (DLS)	100
2	X-Ray Scattering	102
III.	Results	104
1	. Poly(phosphodiester)s with comonomer (0) and the primer sequence of 3-thymidines (P1a and P1b)	105
2	Poly(phosphodiester)s with comonomer (0) and the primer sequence 1-thymidine (P1c)	120
3	. Poly(phosphodiester)s with comonomer (1) and the primer sequence 1-thymidines (P3 and B1)	122
4	Poly(phosphodiester)s with comonomer (2) and the primer sequence 1-thymidine (P4)	131
5.	. Interaction with MspA	135
IV.	Summary	138
I.	The ssDNA molecule	142
П.	Results	142
1	Pattern of events in 4M KCl	142
2	. Analysis of the A _n -induced events	145
3	. Model for the description of the dynamics of A ₃₋₅ with aerolysin	153
4	Extraction of fit values from experimental results	160
III.	Summary	162
CU		
AE	ROLYSIN: A6-10	163
I.	Results	163
1.	. Dynamics of poly-A6	163
2	. Dynamics of poly-A7	166
3	. Dynamics of poly-A8	170
4	Dynamics of poly-A9	173
5	. Dynamics of poly-A10	176
II.	Discussion	180
Ш.	Summary	181

CONCLUSION	
REFERENCES	
LIST OF PUBLICATIONS	
ACKNOWLEDGMENTS	

Chapter 1: Introduction, theory and previous studies

The study of polymers and their interaction with membrane channels is a relatively recent field that started to be explored in the late 80s and early 90s and allows detection of nanometric particles. This technique is a modernized miniaturized Coulter counter [1], which was used for sensing and sizing of micrometric particle like blood cells. Nanopore based detection has been developed a little after the birth of the Patch Clamp technique of Neher and Sakman [2] in the late 70s and 80s which allowed for the first time to measure currents through single ion channels in biological membrane. Since the first pioneering experiment in DNA detection with the α hemolysin nanopore [3], various experiments have since been explored in regards to various molecules using biological and solid state nanopores in experiments of which DNA detection and sequencing [4], [5] to synthetic polymer mass spectrometry [6]. More recently, it has been of wide interest to use this method to investigate fundamental physical and biological phenomena, such as the activity of specific enzymes and the unfolding or dynamics of proteins [7]. The detection of molecules with the nanopores relies on various effects, mainly it is possible to detect molecules because of the nature of interaction with the pore namely binding reactions, as opposed to diffusion through the pore (~100 ns) that would not be possible to measure using current-day electronics.

The main focus of this thesis is to study the interaction with the pores through electrical measurements, from which is extracted through statistical analysis either the polymer's or pore's properties. We focused in this work mainly in the use of biological nanopores, which possess many interesting properties, to study various types of biological and synthetic polymers, for fundamental understandings as well as biotechnological applications in the sensing of polymers.

In this first chapter we introduce concepts of the theory of polymers in relation to our experiments, namely the behavior of polymers in solution and their physical behavior, dynamics in confinement, the fluidic flows through nanopores and the influence of the electrical field in such systems. We also review the state of the art experiments in this field on which a main part of the thesis is based on, with an emphasis on DNA and neutral polymers of PolyEthylene Glycol (PEG).

I. Physics of polymers an confinement effects

It is important to understand the physics of polymers in solution in the frame of these experiments. The conformation of polymers in the solvent is a crucial part of the interaction because the experiments are conducted in an aqueous solution, which will directly influence the interaction of the polymer with the pore and some of the main factors which affect the interaction are the conformation and the concentration of polymers in solution. Therefore, it is necessary to consider all the physical aspects in order to carefully and precisely conduct the experiment. We first introduce the main concepts of polymers in solution namely the behavior of polymer in good and bad solvents and explain the influence of the concentration of polymers on solution, from dilute regimes to melts.

1. Conformation of polymers in solution

1.1. Conformation of a neutral chain

The simplest description of a neutral and flexible polymer chain is through the model of a random walk, where the chain is under no external perturbation. The movement of the chain is described by a network dimension d with N step of length a. Here, each monomer moves independently from the other, each step is independent from the previous walking step and the probability of movement is the same in each direction. For example, in d=3 in a cubic network, the probability of a step in a given direction is 1/6.

The mean square distance value between the two extremities of the chain (end-to-end distance) here is derived from the quadratic sum of all steps over the total number N of monomers $\langle R^2 \rangle = \sum_{i,j}^N \langle a_i, a_j \rangle$. As each step is independent of *j*, this equation becomes $\langle R^2 \rangle = \sum_{i}^N \langle a_i^2 \rangle$, which gives the following distance.

$$R = a\sqrt{N} \tag{1}$$



Figure 1: Model of the random walk of an ideal chain. Extracted from [8]

1.2. Conformation of polymer chains in a solvent and excluded volume effects

In reality, the chain is not totally flexible and there are a number of interactions between the monomers and the solvent that need to be taken into consideration. A polymer chain is constituted of monomers that will interact with each other and with the solvent, the chain is, in this case, semi-flexible. The method used to describe the polymer conformation is the Flory method.

a. Good solvant

Isolated chain – diluted regime

If the concentration of polymers in the solvent is much lower than a critical concentration c^* ($c^* \approx N/R^3$), the chains don't interact with each other (spatially separated). Instead, the interaction between the monomers and the solvent are very strong and the chains are swollen. The radius of one chain is described by the Flory relation:

$$R = a N^{\nu} \tag{2}$$

Where the factor v = 3/(d + 2) and d is the spatial dimension. If d= 3 then v = 3/5.

When the polymer chain is diluted in a solvent, the interaction between the monomers will be more or less strong, depending on the strength of interaction with the solvent.

Melted Regime

In the case where the concentration of polymers in the solvent is very high (c >> c *), the main constituent of the solution is the polymer, thus the designation of polymer melt. In this case we use the distance of the ideal chain ($R = a \sqrt{N}$) because all the excluded volume effects are screened.

Semi diluted Regime

The polymer is in an intermediate regime, $(c \sim c^*)$, excluded volume effects are in effect, but the polymer chains are entangled. Here we introduce a network size ξ , below which the polymers are in a diluted regime and above which they are in a melt configuration.

$$\xi = a(ca)^{3/4} \tag{3}$$

a is the monomer size and c is the concentration of the polymer.

b. Bad solvent

In the case of a bad solvent, there is no affinity of interaction. The polymer chains minimize the contacts between the monomers and the molecules of the solvent, and the chains are collapsed or in a globular conformation. This is for example the case for proteins that fold into globular conformations, where α helixes and β sheets interact via hydrogen bonds or S-S bridges.



Figure 2: Chain conformation in different solvents a) bad solvent b) good solvent

2. Polymers in Confinement and Dynamics

When the polymer is in a confined environment it will take upon specific physical properties; its conformation and behavior will change because of additional imposed constraints. In nanopores many of the polymers studied in this thesis are much bigger than the pore's diameter thus will be a confined state, it is therefore important to take into consideration this added property in our experiments. In addition to this fact it is possible for the pore to move in this pore under the influence of external forces, for example an electrical or a hydrodynamic force, which also need to be taken into account for the studies that are discussed in the thesis.

In the next paragraphs will be first reviewed the theoretical considerations of a static polymer in confinement and more specifically, in the two examples, where its dimensions are of the same order or much bigger than the pore's in models developed respectively by Daoud, DeGennes and Brochard and Muthukumar. The dynamics of the polymer under the influence of external forces of electrical or hydrodynamic nature will be discussed in the second part of this section.

2.1. Confinement of a static polymer

a. Polymer and pore with similar dimensions

When a polymer is confined, the external forces acting on it differ from when free in solution. If the pore has dimensions similar to or larger than the polymer dimensions, one can use some theoretical descriptions proposed by Daoud, DeGennes and Brochard.

It is possible to describe a confined static polymer in a pore of diameter D by a "blob" model [9]. If the polymer in a diluted regime with excluded volume effects the Flory Radius is $R = a N^{3/5}$ as explained in equation (2). However in the case below when the polymer is confined, in a cylinder of diameter D, its conformation changes because of the decrease of its degrees of freedom and it can be described using the following reasoning:

If $D \gg R$ where R is the flory radius of the polymer in the bulk (outside the pore) then there are no changes in the chain conformation, but if R > D, then the chain doesn't have the same degrees of freedom as in the bulk. The chain can be pictured as a sequence of "blobs", with a number g of monomers per blob. These blobs have the same diameter than the pore. Because the scale is inferior to the pore's diameter, inside each blob the monomers cannot sense the constraint due to the confinement and they can move freely and take the same conformation than the polymer in bulk. Following this we can write: $D \approx ag^{3/5}$, which describes the diameter of the blob in relationship to the number of monomers g inside the blob



Figure 3: Cartoon of "Blob" model for a confined polymer

The total number of blobs is $\frac{N}{g}$ and the total length of the confined chain is $L = \left(\frac{N}{g}\right) D = aN\left(\frac{a}{D}\right)^{\frac{2}{3}}$. A free energy of k_BT per blob is needed to confine the chain and we can write:

$$F \approx k_B T \left(\frac{L}{D}\right) \approx k_B T N \left(\frac{a}{D}\right)^{5/3} \approx k_B T \left(\frac{R}{D}\right)^{5/3}$$
 (4)

By comparing this free energy with that of an unconfined diluted free chain which is of order k_BT leads to calculating the partition coefficient p. The partition coefficient is defined as a ratio of polymer concentration inside and outside the pore at thermodynamic equilibrium and using eq. (4) we write:

$$p \approx e^{\left(-F/k_BT\right)} \approx e^{-\left(\frac{R}{D}\right)^{\frac{5}{3}}} \approx e^{-N\left(\frac{a}{D}\right)^{\frac{5}{3}}}$$
(5)

We clearly see that p decreases exponentially with the degree of polymerization N (polymer sizes) suggesting that larger chains ($R \gg D$) do not spontaneously penetrate into the nanopore. As a consequence, in order for the chain to enter the pore one would need to use an external force

for injection. This is the case for the many of the polymers (neutral polymers in chapter 3 and polyelectrolytes in chapter 5) presented in this thesis which need to move under the influence of a force in order to penetrate the pore.

b. Polymer through a small hole

A second description of polymers in confinement proposed by Muthukumar [10], [11] is in the case where the polymer is confined in a small hole. This describes the fact that the radius of gyration of the polymer is much bigger that the membrane thickness and the pore's diameter [11]– [13]. The pore here is embedded within an impermeable membrane, assumed to be an infinite wall. The pore is small enough to accommodate only one monomer at a time and the energy of confinement of the chain is negligible. On one side of the membrane there is *m* monomers and thus *N-m* on the other side. The partition function is the product of the two partition functions, as explained by Muthukumar et al. [11]

$$Z = Z_1 (N - m) Z_2(m)$$
(6)

And the partition coefficient for a tail of large enough N segments in the half space is given by:

$$Z_{half}(N) = N^{\gamma - 1} e^{-\frac{\mu N}{k_B T}}$$
⁽⁷⁾

 μ is the chemical potential and γ is a critical exponent with different values (depending on the solvent conditions). The critical exponent is $\gamma=1/2$ for an ideal solution or $\gamma \approx 0,69$ for a self-avoiding chain (excluded volume effect).

The free energy in this case is

$$F_{half}(N) = -k_B T ln Z_{half}(N) \tag{8}$$

Replacing the expression of Z_{half} from equation (7) into equation (8), we obtain:

$$\frac{F_{half}(N)}{k_B T} = (\gamma - 1) \ln N - \left(\frac{\mu N}{k_B T}\right)$$
⁽⁹⁾

The free energy of the whole chain is the sum of energies of two segments (m and N-m monomers), neglecting an unnecessary constant and assuming the critical exponent is the same on both sides, we obtain:

$$\frac{F(m)}{k_B T} = (1 - \gamma) \ln m + (1 - \gamma) \ln(N - m) - m \frac{\Delta \mu}{k_B T}$$
(10)

 $\Delta \mu = \mu_1 - \mu_2$, is the difference of the chemical potential of monomers between the two compartments.

This type of description could for example be applied to the translocation of some big polymers in comparatively small pores, in biological processes like through of RNA through the nuclear pore, or in the filtration of voluminous polymers through thin membranes (few nm thick).



Figure 4: Polymer through a small hole, with N-m and m monomers on respective sides, from [11]

2.2. Dynamics of polymers

In general polymers in solution are subjected to two different kinds of forces: a random force that gives rise to the Brownian motion of the chain and a viscous force of the fluid that counter acts the movement of the chain. The viscous force has the form $f = 6\pi r\eta v$, where r is the diameter of the molecule and η the viscosity of the polymer solution. Therefore in general in experiments carried in fluids this forces have to be considered in theoretical descriptions. The nanopore experiments is carried in an aqueous environment, where these types of forces need to be taken into account. In addition to these aspects it is again necessary to consider the influence of the confinement on the movement of the polymer under the influence of those forces.

a. Model for "affine" polymer injection at pore's entrance

The model proposed by Daoudi and Brochard as explained in refs. [14], [15] describes the deformation of a polymer chain under the influence of competitive effects, namely the flux imposed by the fluid on the polymer and the chain relaxation (from entropic origin). At the pore entrance, the hydrodynamic flux is convergent and the flexible molecules are stretched out. The polymer, under certain conditions, is then highly stretched near the pore entrance.

The Zimm relaxation time is $\tau \propto \frac{R_F^2}{D}$ and D is the diffusion coefficient described by $D = \frac{k_B T}{6\pi R_F \eta}$ where R_F is the gyration radius and η is the fluid viscosity. Combining these two expressions gives $\tau \propto \frac{\eta R_F^3}{k_B T}$ (Kirkwood approximation).

The conservation of flux through a sphere from a distance r of the pore can be written as $J = 2\pi r^2 v$.

We can write the elongational shear related to the fluid velocity as below.

$$S = -\frac{dv}{dr} \approx \frac{J}{r^3} \tag{11}$$

If the shear rate is low, an isolated object in the fluid is only slightly deformed. If this shear is higher than a critical value $S \ge S_c$, the deformation is called affine. This means that the solvent imposes the deformation on the macromolecule and the polymer is then deformed analogously to an element of the fluid (high elongational shear). Because the viscous forces are very strong in that region, each monomer follows the fluid movement independently from the others. This critical shear S_c is of the order of the Zimm relaxation frequency $S_c \approx \frac{1}{\tau_{Zimm}} \Rightarrow S_c \approx \frac{k_B T}{\eta R_F^3}$.

This critical shear S_c corresponds to a critical radius r_c from the pore, and the polymer is stretched at a distance $r \leq r_c$. The deformation of the polymer is affine as long as it is not

completely stretched. In that case the stretching out of the chain is proportional to the elongation of a volume of the fluid. The transverse radius of the chain R_{\perp} is then equal to:

$$R_{\perp}(r) = \frac{r}{r_c} R_F \tag{12}$$

The polymer is sucked into the pore if $R_{\perp}(r_p) = r_p \Rightarrow R_F = r_c$ (r_p is the pore diameter). Using the critical value $S_c \approx \frac{k_B T}{\eta R_F^3}$ and injecting it into equation (11) gives a critical flux for polymer injection J_c with $J_c \approx k_B T/\eta$. In this model no polymer is injected under this critical flux and the entire polymer is sucked into the pore when $J > J_c$. One might notice that J_c is independent of all the polymer's properties and pore's dimension.

b. Injection with a hydrodynamic flow

Experiments by Long and Anderson [16] showed that the injection of the polymer happens gradually between the polymer's withholding and its complete passage inside the pore. Using the reasoning by De Gennes and Brochard [17]:

$$E_{tot} = E_{conf} + E_{hydro} \tag{13}$$

As showed previously in equation (4) for a chain where the length l is in the confined region, we have:

$$E_{conf} \approx k_B T\left(\frac{l}{D}\right)$$
 (14)

The energy E_{hydro} is the contribution of the viscous forces on the confined blobs, where each one of the blobs is subject to a force ηvD , where η is the fluid viscosity. The total force on all the confined blobs is $F_{hydro} = \eta vD\left(\frac{l}{D}\right) = \eta vl$.

We can relate the velocity to the hydrodynamic flux with the relation $v = J/D^2$.

The work of the viscous force gives $W_{hydro} = \eta v l^2/2$, thus

$$E_{hydro} = -\eta v l^2 / 2 \tag{15}$$

Using equation (14) and (15) we can calculate the maximum $dE_{tot}/dl = 0$ which yields:

$$l^* = \frac{k_B T D}{\eta J} \tag{16}$$

this gives

$$E_{tot}^{*} = \frac{(k_B T)^2}{\eta I} = k_B T \frac{J_c}{I}$$
(17)

Where $J_c = \frac{k_B T}{\eta}$ is the critical flux for aspiration of the polymer into the pore.

The same result for the critical flux of injection is found here again, similarly to the previous deformation model.

This critical flow value has been proved experimentally in ref. [18] for DNA injection under a pressure gradient into gold coated polycarbonate membranes.

We have explained the influence of hydrodynamic fluxes in the injection of polymers in the nanopores. However, in the experiments carried in the nanopores, one has to consider the influence of the electrical field on the polymers and their dynamics under these conditions which will be presented in the next section

2.3. Flows dragging polymers and influence of the electrical field

We can distinguish different influences on the movement of polymers in pores under the influence of the electrical field. Two main influences which can drag polymers to and inside the pore are the electroosmotic flow and electrophoresis, we will explicit both of them and their influence on the velocity of the molecules in such a context in the two following paragraphs.

a. Electroosmotic flow in charged channels

The electroosmotic flow is the phenomenon of a fluid movement relative to a charged surface under the influence of an electrical field. Near the charged surface there are more counter ions, when an electric field is applied, the counter ions start moving collectively, dragging molecules from the solvent and creating a flow near the charged surface.

The electroosmotic flow is important to take into consideration for experiments with charged channels or pores. In the case of cylindrical channels with uniformly charged walls, the fluid

moves in respect to these walls, creating a plug flow. This flow, if strong enough, is then able to drag analytes along with the solvent.

We can calculate this flow by balancing the forces acting on the fluid, including the electrical driving force which gives the Navier-Stokes equation below:

$$-\eta \nabla^2 \mathbf{v}(\mathbf{r}) = \rho_e(r) \mathbf{E} \tag{18}$$

Where η is the fluid viscosity, **v** is the velocity vector, ρ_e the charge density of the surface, and **E** the electrical field vector. We omit, in this case, the term of pressure force as we consider that we have a constant pressure in the system (incompressible system).

For a charged planar surface that we assume infinite in the x and z direction, we consider an electrical field in the direction of the plane in the x direction, where ψ is the surface potential. For infinitely long distances along the x and z axis we consider the charge and velocity uniform, so their derivatives are zero in respect to these variables.

Using the Poison equation relating the surface potential and the charge density,

$$\nabla \cdot (\epsilon_{\omega} \epsilon_0 \nabla \psi) + \rho_e = 0 \tag{19}$$

Where ϵ_0 and ϵ_{ω} are respectively the dielectric permittivity of vacuum and of the fluid.

Combining equation (18) and (19) we obtain

$$\eta \frac{d^2 v_x}{dy^2} - \epsilon_\omega \epsilon_0 \frac{d^2 \psi}{dy^2} E_x = 0 \tag{20}$$

Integrating the equation (20) two times, considering that at $y \to \infty$, $\frac{dv_x}{dy} = 0$, $\frac{d\psi}{dy} = 0$ and the no slip boundary condition at y = 0 thus $v_x(0) = 0$, we obtain this final equation for the velocity:

$$v_x(y) = \frac{\epsilon_\omega \epsilon_0}{\eta} [\psi(y) - \zeta] E_x \tag{21}$$

Where ζ is the surface potential at y = 0, also called the zeta potential.

In the Debye-Hückel approximation, $\psi(y)$ falls exponentially with y. This yields a plug flow with velocity increasing in the double layer and reaching a maximum outside the double layer, where the potential vanishes. This gives a final velocity of:

$$v_{\chi}(y) = -\frac{\epsilon_{\omega}\epsilon_0}{\eta}\zeta E_{\chi}$$
⁽²²⁾

This final relation is known as the Helmholtz-Smoluchowski equation.

The value of this flux can be quite important in the case of solid state nanopores with charged walls, and has to be taken into consideration. In the case of biological nanopores, one has to take into account the non-uniformity of charge distribution and calculate this flow as a dependence of the channel ion selectivity. Then, depending on the distribution and density of charge of the channel, this flux might be of more or less importance. It has been shown to play an important role in the transport of molecules in biological nanopores in refs. [19]–[24].

b. Electrophoresis

It is known that a charged particle moving in an electrical field will move towards the electrode with a charge opposite to its own, following a Coulomb force. With that, in the case of a charged particle in a fluid, the situation gets more complex.

In an electrolyte solution a charged particle will move in regards to the electric field, but there would also be additional viscous forces due to the fluid acting on the particle. Because the particle in solution is surrounded by a cloud of counter ions, the velocity of the particle will be derived at the steady state from a balance between an electrical force applied on the cloud of counter ions (the particle's charge and the counter ions cloud charge) and the viscous force due to the fluid.

For simplification we will talk here about the case of a spherical charged particle in solution (it is possible to use this model for simple enough systems), with radius a, we can use the same reasoning for derivation of the velocity and express the amount of charge in a layer of the double layer shell as described in ref. [25]:

$$dq = 4\pi r^2 \rho_E \, dr \tag{23}$$

Using the Poisson equation in spherical symmetry, we can write

$$q = -\int_{a}^{\infty} 4\pi\epsilon \frac{\partial}{\partial r} \left(r^{2} \frac{\partial \psi}{\partial r} \right)$$
(24)

Assuming there is no potential at $r \rightarrow \infty$

$$q = 4\pi a^2 \left(\frac{\partial \psi}{\partial r}\right)_{r=a} \tag{25}$$

The surface charge distribution can is written as:

$$q = -\epsilon \left(\frac{\partial \psi}{\partial r}\right)_{r=a} \tag{26}$$

We can write the surface potential ψ as :

$$\psi(r) = \frac{A}{r} e^{-r/\lambda_D} \tag{27}$$

Where A a constant and λ_D debye's length. It is possible to write then $\left(\frac{\partial \psi}{\partial r}\right)_{r=a}$ using the fact that at the surface the potential is ζ with $\zeta = \frac{A}{a} e^{-a/\lambda_D}$:

$$\left(\frac{\partial\psi}{\partial r}\right)_{r=a} = -\zeta \left(\frac{1}{a} + \frac{1}{\lambda_D}\right) \tag{28}$$

The derivation gives for surface charge distribution

$$q_s = \epsilon \zeta \left(\frac{1}{a} + \frac{1}{\lambda_D} \right) \tag{29}$$

From the definition of distribution of a surface charge at r = a, $q_s = \frac{q}{4\pi a^2}$, we can derive the zeta potential as below:

$$\zeta = \frac{q}{4\pi\epsilon a \left(1 + \frac{a}{\lambda_D}\right)} \tag{30}$$

or

$$\zeta = \frac{q}{4\pi\epsilon a} - \frac{q}{4\pi\epsilon(a+\lambda_D)} \tag{31}$$

This relation means that the ζ potential is the sum of two potentials, one generated from a charge q by the surface charge at r = a and the second generated from a charge -q sphere of radius $a + \lambda_D$ which corresponds to the cloud of counter ions distributed in the concentrical layer with Debye length.

Now we can derive the velocity balancing the forces acting on the particle. In the steady state, the electrical force (acting on the net charge of the particle q) and the stokes drag force are balanced which gives

$$qE = 6\pi\eta a \mathsf{v} \tag{32}$$

The exact calculation of the velocity is rather complicated and often requires numerical solutions. In the limit of a small or big Debye length (in comparison to the particle size) it is possible to derive a simple form of the velocity, and it can be used in approximation.

Combining equation (32) with equation (30). For a big Debye length limit we obtain:

$$\mathbf{v} = \frac{1}{2} \frac{\epsilon \zeta}{\eta} E \tag{33}$$

In the case of small Debye length, in other words an infinitely thin double layer, we can neglect curvature effects and we can apply the same reasoning that we applied for the derivation of the electroosmotic velocity (for planar pore walls),

$$\epsilon \psi E = \eta V + \text{constant} \tag{34}$$

The two velocities only differ in the final integrated form. Here v is the velocity of the fluid parallel to the surface. The boundary conditions here are $\psi = 0$ and v = 0 at the edge of the diffuse layer and $\psi = \zeta$ and V = v (the velocity of the particle) at the surface, which gives the surface velocity of

$$\mathbf{v} = \frac{\epsilon \zeta}{\eta} E \tag{35}$$

And we find again here the Helmholtz-Smoluchowski equation.

II. Previous Experiments

We will discuss in this section, the state of the art experiments which are closely related to the studies carried out in this thesis, especially studies of DNA and PolyEthylene Glycol (PEG) detection, which were one of the firsts to be carried out in this field. We will review the pioneering experiments, which showed the most potential in this field and led to the increasing number of nanopore experiments in the late decades.

1. Experiences with neutral polymers

Krasilnikov and collaborators started to work with synthetic polymers and observe the interactions of these polymers with biological pores in the late 80s. This proved to have great potential for molecular probing and detection, allowing probing of the pore's or polymer's properties. In the 1992 study by Krasilnikov et al. [26] it was shown that the geometry of the α -Hemolysin pore could be probed using polyethylene glycols. This was the first experiment using PEGs with a biological nanopore. Later in 1994, Bezrukov and collaborators [27] used PEGs with another biological nanopore, the alamethicin, to count molecules passing through the pore. These were the first counting experiments done with biological nanopores.

Bezrukov, Kasianowicz *et al.* [28] later conducted additional PEGs experiments and in 1996 they published a study about the dynamics and free energy of PEGs in the α H pore. It was demonstated that the diffusion is not able to describe the interaction of the PEG with the pore and that a polymer had an attractive interaction with the pore, by being reversibly adsorbed to the pore wall. They showed that the predictions using the scaling theory (De Gennes scaling concepts on polymer) and the simple diffusion model of a hard sphere into a cylinder were in disagreement with their results, as shown in Figure 5 (with the two dashed line representing the disagreeing model). The solid line in Figure 5 represents the fitting of the relative average conductances, with the model they proposed in their study, in presence and absence of the polymer in the solution. They were able to extract the partition coefficient $p(w) = e^{-\left(\frac{w}{w_0}\right)^{\alpha}}$ representing the average density of monomers inside the pore to those in the bulk, where *w* is the molecular weight of the PEG and w_0 and α are extracted from the experimental fitting.



Figure 5: Hemolysin conductance ratio & Hydrodynamic Radius vs PEG molecular Weight. Extracted from [28]

They accounted for the polymer pore interaction by using an "Ad-Hoc" description of the diffusion coefficient that took into account the polymer's weight (i.e size) dependence of interaction with the pore (polymer size dependent activation process) with a diffusion coefficient $D(w) \propto e^{-\frac{\gamma w}{w_0}}/R_D$, where R_D is the hydrodynamic radius of the polymer. In this paper they suggested that the interaction of the PEG with the pore walls might be of hydrophobic nature, given the usual hydrophobic nature of the PEG-protein interactions. This finding was surprising because of the aqueous nature of this type of channels. This study suggested a complex set of interactions besides diffusion, between the analyte and the pore.

In another study by Merzlyak et al. [29], they showed that using an asymmetrical addition of PEG (one side only of the pore) is able to better probe the pore geometry than in the symmetrical case, where the PEG is present on both sides of the nanopore (the symmetrical case was described in a previous study [26]). Using a filling parameter which describes the ratio of pore length (accessible to the polymer) relative to the whole length of it, $F(w) = \frac{[g_0 - g(w)]\chi(w)}{[\chi_0 - \chi(w)]g(w)}$, they plotted it against the polymers hydrodynamic radii and they performed the experiment by placing the PEG asymmetrically in the *cis* or the *trans* side (to probe the differences in the pore's geometry from both sides). As shown in Figure 6, they observed that the polymers with hydrodynamic radii over 1.2-1.3nm don't partition into the pore anymore. Based on this result, they concluded an almost similar opening from both sides of the pore. Furthermore, they were able to deduce from this method the channel's constriction of 0.6-0.7nm and 0.9nm. These results were in close agreement

with the light crystallographic results by Song et al. [30], showing the powerful tool that nanopores are for such measurements.



Figure 6: Dependence of Filling coefficient (cis and trans) on the PEG hydrodynamic radius. Ref [29]

Another study by Movileanu *et al.* [31], showed polymer partitioning into the α -Hemolysin pore for PEG ranging from 940 to 6000 Da. They used the partition coefficient and scaling laws of Daoud and De Gennes and they deduce from their results that molecules with molecular weights higher than 5kDa seem to not partition into the pore.

Salt also showed to play an effect on the interaction of the polymers with the nanopores. Krasilnikov, Kasianowicz and collaborators revealed in different studies that close to the limit of solubility of the salt they observed an enhanced interaction between the PEG and the α -hemolysin pore. In 2006 Krasilnikov, Rodrigues and Bezrukov [32] demonstrated the dependency of interaction for different molecular weights of PEG with α -hemolysin in a 4M KCl solution. Increasing salt concentration changed polymer partitioning and allowed them to observe time-resolved PEG-induced blockades. They found that dwell times increased with molecular weight until a certain size (~PEG 3400), where the polymer seems to not be completely accommodated in the pore anymore and decreased for higher molecular weights. They explained this phenomenon by the polymer's tail, which was not accommodated inside the pore and acts like an entropic spring pulling on the trapped part of the polymer.



Figure 7: PEG residence time *vs* PEG molecular weight. PEG is fully accommodated until ~3400, then the entopic spring pulls on the polymer for higher molecular weights. Ref [32]

The study of Kasianowicz, Krasilnikov *et al.* of 2008 [33], showed the mass-dependent interaction of the α -Hemolysin and a polydisperse solution of PEG. Using the amplitude of a polymer blockade, they obtained a mass spectrum in agreement with the MALDI-tof spectrum (Figure 8) and they demonstrated that characteristic dwell time values depend on the size of the polymer in solution with differences of one monomer. Indeed, it seemed that the pore was highly sensitive to the size of the polymer and that the interactions between the pore and different molecular weights of the polymer showed clear conductance differences.


Figure 8: Conductance and PEG blockades in hemolysin nanopore and corresponding histograms (left) Histogram of nanopore conductance in the presence of polydisperse PEG1500 and corresponding MALDI TOF. Ref [33]

Later, in a more detailed study about salt concentration effects on the pore-PEGs interaction [34] the authors clearly showed that increased 4M KCl electrolyte concentration increased the onrates, which became hundreds of times larger than in 1M, while off rates become several hundreds of times smaller. They deduced that the neutral polymers, in this case, can form complexes with ions and that this changed interaction with the α -hemolysin pore. In this case, they deduced the ability of PEG to be charged in a high salt condition is because of a non-linear voltage dependency of constant rates. They also observe a high increase in the on-rate constant (and amplitudes) allowing the recognition of single molecules, due to salting out effects, change in polymer conformation, and Coulombic interactions.

These kinds of experiments with increased salt solution were enlarged for different molecules (neutral polymers and polyelectrolytes) and also pore types, as they proved to be powerful tools for the sizing of various analytes [33], [35], [36]. In many of our experiments, we also use a high salt solution to observed enhanced effects of interaction between the pores and our polymers of interest.

Other recent experiments working with neutral polymers include probing the effect of ions on PEG partitioning into the α -hemolysin pore [37], showing the importance of alkali cation type in such experiments and another one on the electroosmotic flow in the same channel and PEG [38]. These last ones are closely related to our studies and will be further discussed in the next chapters.

2. Experiments with DNA

The first pioneer experiment in the study of DNA transport through a protein channel was carried out by Kasianowicz et al in 1996. This experiment focused of the study of polynucleotides. In the experiments the authors used DNA and RNA: Poly[A]₅₀ and polyUracil (Poly[U]₅₀), highly negatively charged molecules that were driven by an electrical field towards the pore to achieve translocation [3]. By showing that the lifetime of blockades were proportional to the mean length of the RNA molecules, that the increase of voltage decreases the lifetime of the blockades in general and finally, by using PCR, they demonstrated that ssDNA translocates through the pore in contrast to double stranded DNA, which doesn't, it was suggested that the molecules completely thread through the pore as extended rods.

Later on, based on their results, they explained the voltage dependence of the capture rate using a Van't Hoff Arrhenius transition state equation [39].

$$R = \kappa v e^{\frac{\Delta U - U^+}{kT}}$$
(36)

Where κ is the probability factor $\kappa \sim 1$, ν is the frequency factor, ΔU the energy driving the polynucleotide, and U^+ the activation energy, i.e the barrier height. Assuming that the electrical energy of the polymer is $\Delta U = ze|V|$ one can rewrite the rate equation as:

$$R = Roe^{\frac{ze|V|}{kT}}$$
(37)

To get a value for the activation energy they estimated the value of Ro, where Ro is described by $\frac{CDA}{l}$. Here, C is the concentration of polymer, D the diffusion coefficient, and A cross section of the channel and I the channel length. The energy barrier was estimated to be ~ $8k_BT$

Assuming there is only one barrier for polymer entry, it is automatically deduced that an increase in the electrical field reduces the energy required to confine the polymer in the pore. By fitting the results in Figure 9, they obtained a valency of $z \approx 2$, explaining this result by an average of two nucleotides interacting with the pore entrance and initiating the translocation (as in these

conditions one phosphate group bares an average of one negative charge). They explain later that the polynucleotides are captured more frequently from one side than the other due to an asymmetry in the pore. They argue that the confinement barrier is higher on one side in comparison to the other due to a difference in geometry and that there is a higher density of negative charges on one side than on the other, which increases the electrostatic repulsion and thus the barrier for entry.



Figure 9: Rate of DNA blockades from *cis* and *trans* vs Voltage (absolute value). Ref [39]

Since the first DNA experiments, enhancing DNA sensing using biological nanopores has been one of the most challenging and important aspects of this field. The most widely used pore for this purpose up until today was the α -Hemolysin pore, which can easily form pores in lipid bilayers. With this pore, very stable recordings could be carried out for long durations of time. Through various experiments it was indeed shown with this particular pore that detection with various molecules could be improved by placing a molecular adapter at the pore entrance [40], [41], by genetically engineering the pore [42], [43]. Moreover, it has later been found by Bayley, Ghadiri, et al. [44] that ssDNA detection can be enhanced by trapping DNA hairpins, achieving measurements in current detection between different ssDNAs, up to one nucleobase. In the study by Clark, Bayley et al. [41] they demonstrated that a covalently attached molecular adapter could be used for continuous detection of indigenous mononucleotides. A β -cyclodextrin which narrows the constriction of the beta-barrel and enhances the chemical specificity of the pore allows detection of the single nucleotides of dAMP, dGMP, dCMP, dTMP (2'-deoxyadenosine 5'-monophosphate (dAMP), 2'-deoxycytidine 5'-monophosphate (dCMP), 2'-deoxyguanosine 5'-monophosphate (dGMP), and 2'-deoxythymidine 5'-monophosphate (dTMP)). They obtained a resolution up to a single nucleotide (every mononucleotide passing through the pore produced a different current signature) with a confidence interval of 99% in the optimal experiment condition. In a following study of Stoddart, Bayley et al. [43] it was shown that recognition of DNA molecules can be improved using engineered α -Hemolysin pores, and DNA was immobilized using a biotin-steptavidin linkage as (see Figure 10). This study showed that discrimination between types of nucleobases within one DNA molecule could be achieved by three different sites in the α -hemolysin pore, which was not the case for the WT pore.



Figure 10: A. Engineered α -Hemolysin pore ref [43]. a. Current trace with nucleotides blockades. b. Corresponding histograms ref [41]

Although the α -Hemolysin nanopore was, up until then, the most widely used pore for DNA sequencing, it presented a drawback. Indeed, the β -barrel is about 5 nm long and is able to accommodate up to 10 bases of ssDNA and the current is very much altered by all the nucleotides

at once. Even though engineering the pore makes single nucleotide differentiation in a homopolymer sequence containing a single different base possible, as shown in [43], the high fidelity of such measurements for much larger and complex DNA sequences was not proved. This was due to the fact that it was difficult to achieve high signal-noise ratios with such small differences in current detection.

Another pore, which has lately proved to be very powerful for the detection of ssDNA, is the MspA biological pore. This nanopore, which has a different geometrical shape, has a much narrower barrel constriction of 1.2 nm diameter and 0.5 nm long. With this barrel it is possible to accommodate around 4 base pairs, in contrast to 10 bp in the α -Hemolysin β -barrel. The experiments aiming to achieving this technology with MspA were started by Jens Gundlach and collaborators and have since been carried out mainly by the same group. The first experiment carried out was published in 2008 [45], and was the first study where DNA interaction with MspA was shown with a mutant of this pore. Negative charges at the narrowest constriction of the channel were eliminated, which allowed a higher capture of the DNA hairpins in contrast to the WT MspA (very low capture rate because of the repulsion between DNA and negative charges). They showed that the mutant with a positive distribution of charge is the most stable regarding the gating behaviour, and that the duration of interaction between the pore and the ssDNA becomes longer (100µs, in contrast to less than 10µs with the first mutant).

The latter method was further developed by the same group [46], showing that the three bases adjacent to the DNA hairpin are those responsible for the residual current values and that less than 4 bp are participating in the reduction of the pore's conductance. It was also demonstrated, as a proof of principle, that it is possible to achieve a single nucleotide read out by amplification of the single nucleotide and addition of duplexes of double stranded DNA in between the single stranded DNA nucleotides. This was called duplex interrupted nanopore sequencing: each nucleotide gives rise to a conductance value and then due to the high electrical field in the nanopore, the duplex dissociates, allowing the next nucleotide to move forward in the pore and so on.



Figure 11: Detection and translocation of DNA hairpins through the MspA nanopore. Ref [46]

Lately, a new biological pore less explored, the aerolysin nanopore, has been used to size short ssDNA oligonucleotides, detect temperature effects in the transport and detect pepitdes. These ssDNA experiments performed in the group of Yitao Long showed the power of discrimination of the aerolysin pore with short polynucleotides, (3 to 10 adenines) [36] and later they revealed the different sites responsible for the DNA binding through experiments and simulations [47]. This pore, which was not investigated with DNA or peptides before, seemed to have a great potential for such kind of experiments as it is also highly charged. We will discuss this point furthermore into the thesis as we have also used this pore into our studies with DNA as well as neutral polymers and synthetic charged polymers.

Mathé et al and Wanunu et al. [48]–[50] performed other experiments measuring escape times and orientation dependent interactions using DNA hairpins and biological nanopores. These experiments showed the dynamics of DNA hairpins in α -hemolysin with escape times and unzipping of the molecule under the applied voltage. They also observed different dynamics of interaction depending on the 3' and 5' orientation interaction of the molecule, with the pore validated by numerical simulations. We will discuss in more detail these last studies in our experimental section.

It is also possible to measure double stranded DNA using solid state nanopores. Different materials and techniques have been developed these past few years to optimize the measurements using these pores. The traditional material used was SiN because of its stable chemical nature and

low mechanical stress. The first group that used the SiN nanopore was Golovchenko and collaborators [51], reporting measurements of double stranded DNA in 5nm ion-beam sculpted pores. Other experiments measuring dsDNA have been reported using graphene films [52]–[54]. Using these films was attractive because of the comparable thickness of graphene films and spacing between DNA bases, which make them of great interest for single nucleobase detection. Lately, other techniques have been used coupling biological nanopores with solid state pores [55] or coating SiN surfaces and pores with lipid bilayers [56]. The major challenge with solid state nanopores is the control of the chemistry and charge reproducibility of the pores over the different experiments, which is of a major importance, especially in single molecule detection.

III. Summary

We have reviewed some of the most important experiments, which have led to the big advances in the field of nanopore detection and allowed sizing, sequencing and understanding of some fundamental concepts of the interactions of polymers with ion channels, and more specifically biological nanopores. These studies constitute the basis of most our experiments and will be recurrently referred to in the context of our investigations.

We will explain in the next chapters the different pores that were used in our experiments and their particular properties. We will also present the experimental setup used and technique of data analysis and finally discuss the results obtained during this thesis with the various biological and synthetic polymers studied.

Chapter 2: Biological pores and ion channels

Pores are of primary importance for biological studies. Indeed, they represent one of the main pathways of transport for prokaryote and eukaryote cells. The wide classes of pores that exist in nature play essential roles in the regulation and molecular transport through the cell membrane.

These pores can allow the transport of necessary nutrients, regulation of the cell's processes or signaling between cells and molecule transfer accomplished by various degrees of pore selectivity can critically vary within cell types of the same species and between different species. Internal transport (cytoplasm-nucleus transport) and external cell transport and signaling mechanisms (cell's external membrane-outer environment) are based on these pore functioning. Because of passive transport (the transport of molecules through diffusion) of small molecules, including ions, most of the pores are ion channels, however some pores only serve as a pathway for the selective transport of ions, and are called ion channels.

One of the main examples of regulatory internal signaling is transport through the nuclear pore [57], [58]. The nucleus which contains all the genetic information needs to have a strong protection from the outer environment. Its pore allows some passive transport (molecules <40kDa, ions, metabolites etc..) but is however highly selective to higher size molecules, through a complex pathway of selectivity governed mainly by nuclear pore proteins called FG-Nups (nucleoporins which contain repeat units of phenylalanine and glycine: FG) organized in a mesh-like structure, and permits transport of ribosomal complexes and external transport of RNA to the cytoplasm.

An instance of cell to cell signaling based on ion transport, is in the case of the nervous system. Neurons ion channels in the case of eukaryotic cells, which transport low amounts of ions, are highly specific to the ion type and are voltage gated (open only when the membrane potential changes). This specificity is crucial for the propagation of nerve impulses. During transmission of these signals one impulse is transferred through a very specific pathway of neurons and these ion channels play a crucial role for the polarization/depolarization of each neuron cell, giving rise to the mechanism of action potentials (threshold mechanism of ion transport necessary for transmission and directionality of signals in the nervous system). The signal which then travels through the neuron (due to membrane depolarization, the voltage gated channels open) to the end of axon is finally transferred through the synapses where a chemical substance is transferred to a

receptor opening ligand-gated sodium channels, causing in turn depolarization of the next neuron membrane, which in turn transfers the impulse.

One other main example of ion selective pore is the KcsA pore of the prokaryotic cell *Streptomyces lividans* (soil bacteria). It is a voltage-gated potassium channel, and it was the first K^+ channel to be characterized using X-Ray crystallography [59]. Research on this channel, has revealed important mechanisms in ion selectivity and gating function. In addition, part of its structure seemed to be highly conserved within the various K^+ channels of prokaryotes and eukaryotes [60], it became a model for the study of potassium channel.

These examples illustrate the complexity and selectivity of such systems that represent important instances necessary to carry processes in biology. However, other pores, which are mostly prokaryotic based are less selective than the previously cited examples, and can serve other purposes than regulation, as it is the case of the mycobacterium porins which serve nutrients transport into the mycobacterium or the pore forming toxins (PFTs), which are toxins that are secreted by the bacteria for host infection purposes. It is important to study these classes of pores to understand their functioning mechanism which is ultimately necessary for biological and medical research application, like drug delivery.

In the frame of our experiments these pores can also be used for different purposes, and they were proved to be powerful tools for sensing and biotechnological applications: because of their high stability in in vitro experiments and their lack of selectivity they can interact with various molecules and various types of polymers ranging from synthetic polymers like Polyethylene Glycol (PEG), or natural polymers like DNA and RNA, to various peptides and proteins, which gives them a high potential for characterization. In addition it is important to understand some processes like translocation of polymers like RNA through the nuclear pore [57], [58], [61] or DNA through viruses capsids [62], [63] for example which use some similar kinetic mechanisms for these processes, and these bacterial pores can be used as model systems to unravel these various biological and physical processes (like confinement, intermolecular forces, etc..).

I. Comparison of the different existing nanopores for molecule sensing

As briefly mentioned in the first chapter of this work, in the polymer-pore experiments, various types of pores can be used, ranging from biological to solid state nanopore, or other more "exotic" pores constructed from DNA (in the same principle than DNA origami). The various pores and their applications in addition to their advantages and drawbacks are described in refs.[64]–[66]. As explained, biological nanopores, show a high potential for the detection of various molecules: polymers (DNA, synthetic etc..), proteins and are one of the most stable structures for this purpose. Other pores which are solid state nanopores, or pore hybrids, are also used by other groups for detection of polymers for specific advantages and depending on the interest of the user. The following figure [65] and table [66] summarize the advantages and inconvenients of the various existing pores:



Figure 12: Comparison of the four main classes of pores and listing of their advantages and inconvenients. Extracted from [65]

Comparison between different nanopores.

Nanopore type	Biological	Synthetic in PL	Silicon materials	2D	Polymer	Glass nanopipettes
Fabrication process	Expression and purification	Synthesis	FIB TEM Di-electric breakdown asymmetric etching in KOH	Formed as intrinsic defects, FIB TEM Chemical routes	Track-etching technique	Pipette puller
Shape	Inherent to the protein	barrel	Hourglass, smooth and round	Round, atom thick triangle, atoms thick	Cylindrical, conical, hourglass, cigar, half-cigar, depends on the etching process	Conical
Cost for a single NP	Cheap	Cheap	Expensive excepted for the dielectric breakdown	Expensive	Expensive	Cheap
Cost for upscale membrane	Expensive	Expensive	Expensive	Depend the material	Cheap	Impossible
Reproducibility	High	High	does not allow for a reliable preparation of small, sub-20 nm pores		does not allow for a reliable preparation of small, sub-20 nm pores	does not allow for a reliable preparation of small, sub-20 nm pores
Sensitivity	High, engineered with atomic precision		Low	Atomic precision	Very low	Low
Life time	hour long recording	hour long recording	From several minute to hours	From several minute to hours	weeks	weeks
Operating voltage	Low ($\pm 200 \text{ mV}$)	Low (±200 mV)	Hundreds of mV	Hundreds of mV	Up to 10 V	Up to 10 V
Functionalization	Protein engineering	impossible	Possible, deposition, and/or chemical functionalization,	Depends the material	Easy, deposition, and/or chemical functionalization,	Easy, deposition, and/or chemical functionalization
Pore size	From 1.2 to 3.4 nm	Few nm	From nm to hundreds of nm	From nm to hundreds of nm	From nm to hundreds of nm	From nm to hundreds of nm
Noise Potential application	Low Single molecule detection	Low Single molecule detection	High Single molecule detection and biomimetic nanopore	High noise Single molecule detection Energy harvesting	Low noise Single molecule detection filtration, energy, harvesting, Ionic diode	Single molecule detection, and biomimetic nanopore

Table 1: Comparison between the different pores. Extracted from [66]

Based on the characteristics described in Figure 12 and Table 1, we have chosen to use biological pores, because of their high stability in bilayers their sensitivity in detection of polymers and the fact that they are cheap in comparison to other pores. They also have proved over the course of various studies, to be precise tools for the detection of H and D, divalent cations, polyethyelene glycols, sugar molecules, DNA and RNA, various proteins and peptides [6], and demonstrated great potential for the sensing, sizing, sequencing, and study of fundamental processes of the various molecules [6], [64].

We will present next the different pores that were used in the studies namely two PFTS and one mycobarterium porin. We will present their atomistic structures which are commonly determined via X-Ray crystallography (however for one of the presented PFTs -aerolysin- the situation is more complex and will be explicated further on) we will present in detail the various research that has been done to unravel their amino-acidic structure, their geometrical structure as well as their mechanism of insertion into the lipid membranes. It is important to know and understand the structure of these protein pores, in the frame of our molecule-detection experiments, because molecular and geometrical properties directly influence the interaction between the protein-pore and the polymers or molecules that we want to study or sense.

II. Pore forming toxins

Two of the biological nanopores used in our experiments are pore-forming toxins (PFTs). these toxins once formed on the outer layer of cells constitute highly stable structures in nature, this stability also confers it a property of use in *in vitro* bilayer systems and vast research showed that they have a high potential for sensing applications [6], [7], [66].

These toxins are secreted in nature by pathogenic bacterial organisms during infection of hosts. In general, they kill cells by forming unregulated pores in their outer membrane, causing lysis of targeted cells and ultimately cell death. Since the 1800s, these toxins, which are secreted in soluble forms in nature were studied in bacterial supernatants [67]. During the last decades, advances in technological instrumentation (development of microscopy techniques, X-Ray crystallography, patch clamp and *in vitro* bilayer experiments) led to an increasing number of experiments on such proteins

It was shown that PFTs were released as water soluble monomers in the organism and undergo various conformational changes in order to form water channels within the membrane [68], [69]. Upon binding to the membrane, their local concentration increases which enhances the possibility of oligomerization [69]. The oligomers, which are bound non-covalently, form pores through the cell membrane, with different sizes depending on the nature of the toxin. The protein has an amphipathic property, with both hydrophobic and hydrophilic regions, this facilitates insertion into the membrane and the formation of the water filled channel. This insertion is in agreement with the amphipatic property of the PFTs: hydrophobic region is in contact with the membrane and hydrophilic region in contact with water.

Even though these pores are not highly specific to the cells it has been shown that some degree of specificity can be observed which is dependent on the type of PFT, and that some cell receptors can be targeted by PFTs, helping the oligomerization and stabilization of their structure [68]–[70].

We will now detail the functional structure of the PFTs chosen for our experiments as well as their different properties which convey them their specificity in sensing experiments. In lab based experiments these proteins can be expressed in other model organisms that are less pathogenic (ex: *E.Coli*), and subsequently purified. The first presented pore α -hemolysin, has been one of the must widely used PFT in sensing applications.

1. Alpha-Hemolysin (α-HL)

1.1. Structure

Extensive research has been carried out with this toxin especially in *in vitro* lipid bilayer experiments [4]–[7], [66], [71] because of its high potential for sensing application, it has been shown to be able to sense and differentiate DNA and RNA molecules, or in sensing PEG polymers differentiating them up to single monomers in monodisperse solutions, as would be possible with mass spectrometry [33], [72], [73].

α-HL is a cytolytic protein released by the *Staphylococcus Aureus* which can infect the human species. During infection, it can bind to various types of cells in the organism; it is nonspecific and can form pores easily in various lipid membranes, however protein-cell receptors might help recognition and binding to the membrane [67], [69] in some cases It is secreted by the bacteria as water-soluble monomers of 33.2kDa that form heptameric structures weighting 232.4kDa upon interaction with the membrane of eukaryotic cells such as endothelial cells or human platelets [67].

The structure of this protein was determined through X-ray cristallography in 1996 by Gouaux *et al.*[30] with a resolution of 1.9Å. According to their study he protein reaches a total length of 100 Å and total diameter of 100Å too. They showed that it has a mushroom-like shape with its channel stem or β -barrel (a coiling β -sheet that forms a barrel) protruding through the membrane and its cap sitting on top of the membrane (see Figure 14). Within the solvent accessible regions, the stem has a diameter of 26Å part reaching a 15Å constriction on the top part of the stem region (Figure 14.c) .The diameter of the water filled channel in the cap part reaches a maximum diameter of 46 Å. The residues in the interior (shown in Figure 14.c) of the β -barrel are polar and uncharged except at the constriction site (stem neck) with paired rings of glutamic acid (GLU) and lysine (LYS) and at the channel's base with rings of lysine (LYS) and aspartic acid(ASP). In contrast, the cap part is more charged as displayed in Figure 14.b

As a consequence, the channel is slightly selective to anions and shows current rectification [30] as it conducts current asymmetrically under opposite voltage polarities: more current is conducted under the influence of positive than negative voltages.

1.2. Pore formation and stability

From the release into monomeric form to full oligomers, α -HL undergoes multiple stages before complete stable insertion into the membrane as explained in ref [74] (representation of PFTs formation is :



Figure 13: Representation schematics of the formation of PFTs from soluble form to oligomerization. The monomers are released by the bacteria in the inactive soluble form, followed by membrane binding and activation and subsequently pre-pore formation or membrane insertion. Finally after multiple stages, the final assembly leads to the stable pore embedded across the membrane. Extracted from [69].

According to various studies, circular dichroism shows that the β -sheet structure of the protein remains almost unchanged from α_1 to α_7 [74]. However, the sensitivity to cell proteolysis gradually increases through the different states and evolves towards a higher degree of resistance [74]:

- α_1 is digested within two regions: the amino latch and its glycine-rich region
- α_1^* and α_7^* are only digested in the amino latch
- α_7 is very resistant to protein digestion.

The final stability of the heptameric pore relies on the optimal state of higher proteolysis resistance of the α_7 structure.

Various experiments and simulations have been performed using this pore, it can also be purchased commercially which makes it very convenient to use. In our studies its high degree of stability upon formation into the bilayer and its high potential for sensing makes it a very good candidate for lipid bilayer experiments in the frame of this work on polymer sensing. As mentioned in the first chapter of this thesis, this pore is very resistant to different parameters (temperature, pH, chatropic agents...) in planar lipid experiments. Based on these characteristics, it is possible to obtain long and stable recordings on the order of hours with the same pore, in a wide range of experimental conditions: various salt concentrations [34], salt types [37], [75], different pHs [76], [77], or temperature [78]. All these properties makes it a very interesting pore to use for sensing applications.

The next pore that will be presented is also a pore forming toxin which possesses interesting properties useful for the same previous named purposes (sensing and investigation of polymers and their interaction mechanisms with the pore) however less research has been carried out so far, as it is only recently that more has been discovered on its structure, leading in parallel to increasing bilayer experiments.



Figure 14: Simulation of the atomistic α -hemolysin structure using Jmol (structure from the protein data bank PDB). (a) Cartoon structure of the secondary structure. Color code per monomer unit. From left to right: side view, top view, single monomer unit in the pore. (b) View of the charged state of the pore. Blue = positive amino acids, red = negative amino acids. From left to right: side and top view. (c) Sagittal section. From left to right: solvent accessible surface, charged state and corresponding charged amino acids.

2. Aerolysin

2.1. Rationale of research to unravel the structure

Aerolysin is secreted by *Aeromonas Hydrophila*, as a water-soluble monomer, in an inactive structure called proaerolysin [79]. The final pore is a heptameric structure and the oligomerization into a heptamer of the aerolysin monomers can only happen once the proaerolysin is activated via proteolysis. Despite multiple efforts, crystallization of the final pore of aerolysin has not been possible, because of the fact that the protein might be highly flexible in that form. In contrast, proaerolysin crystallization was possible and the atomic structure has been available since 1994 [79]. This proaerolysin was observed in dimeric forms in crystals and the dimer structures have also been observed at higher toxin concentrations in solution (above 1mg/mL) [80], [81], and because of these observations, this dimeric structure was thought to be the form secreted by the bacterial pathogen [82]. It is only recently that a more advanced description of aerolysin pore structure was provided [83], [84]. The authors used a combination of cryo-EM, X-Ray and modeling to describe the structure and the mechanisms of insertion of the pore into the cell membrane.

In the study by parker *et al.* [79], it was shown that proaerolysin contains two main domains: an elongated part with three non-continuous sub-domains and a smaller globular domain with a -NH₂ terminal end (Figure 16.a). Protein activation relies on cell proteolysis. Indeed, the C-terminal peptide at domain 4, necessary for the monomer to be water soluble, needs to be cleaved by cell proteases at the loop connecting the CTP-site (cytidine triphosphate) with the main body (*cf.* C-terminal in Figure 16.a). Removal of this part exposes a hydrophobic patch in that domain, as described in refs. [82], [83] which allows oligomerization into the final heptameric structure that can insert into the membrane. Detailed mechanisms will be further described in this part.

2.2. Structure, heptamerization and insertion mechanism

The structure studies in refs [83], [84], which reached a resolution of 7.4 Å, showed that compared to hemolysin, the aerolysin's channel has a similar β -barrel structure, however in contrast to hemolysin the size of the barrel stays almost constant from top (cap) to base, with an average of ~1nm . In addition, a constriction site is observed at the channel's entrance (cf Figure 16.c), corresponding to rings of arginine (positively charged) and aspartic acid (negatively charged). In contrast to the previously described hemolysin toxin, aerolysin is highly charged due to several rings of basic and acidic residues. Figure 16.c shows that the channel's top part has two

successive pairs of rings of arginine and aspartic acid. These pairs are followed by one ring of lysines which are positive residues, itself followed by two rings of lysines, with a ring of glutamic acid in between. Below, inversely, are observed two rings of glutamic acid with a ring of lysines in between. In addition to the structure study, the insertion mechanisms of aerolysin in membranes has been explored.

In particular, degiacomi and collaborators [83] and Iacovache and collaborators [84] revealed that after activation of the proaerolysin monomer, a very particular mechanism of insertion happens: a swirling mechanism, followed by a vertical collapse of the β -barrel into the membrane allow for the protein to be incorporated into the membrane. The authors of ref. [84] also unravel the presence of a concentric β -barrel fold, held by high hydrophobic interaction, which conveys it a prion-like stability and makes it a highly stabilized structure. More specifically, upon binding of the proaerolysin to the membrane and following its activation, it is shown that formation of a prepore occurs: a secondary state of this prepore starts to form subsequently with a pre-stem loop that can insert into the membrane via vertical collapse, made possible by the torsion of two regions at the base of the β -barrel. This two regions twist consecutively while the β -barrel concentric structure is held intact, which causes the collapse of the inner β -barrel into the membrane.



Figure 15: Formation of the aerolysin pore. Different stages from prepore, to final pore, which forms after the collapse of the beta barrel (yellow and blue rods). The blue rods represent the concentric beta barrel described in the text. Extracted from [84]

As it has been shown, only few authors have tried to tackle the challenge of unfolding the structure and folding mechanisms of the final aerolysin pore [83], [84]. Similarly, this pore has been less investigated with analytes and it is only in the past decades that groups started to use this

pore for polymer/molecule investigation [35], [36], [85]–[87] and already through this few studies, this protein-pore shows to have many interesting properties, including its very narrow channel (~1nm diameter) which can increase the sensitivity in the detection of smaller molecules as it has been shown for small oligonucleotides [36], [47]. Its high charged state (Figure 16.b) can also be a very interesting property for example in the study of polyelectrolytes [85], [88].

This pore is also very stable once formed in the membrane, and can similarly to hemolysin be used for long hours of recordings. All these specificities, make aerolysin very interesting for our studies.

We will present next, the last pore that has been used in our studies, which is a mycobacterial porin. In contrast to PFTs, the purpose of porins is to transport molecules into mycobacteria, however the specificity and selectivity are still low in comparison to highly selective pores, like the previously described voltage gated channels. This pore therefore can be similarly used for the study of various molecules and polymers.



Figure 16: Simulation of the Cryo-EM aerolysin structure using Jmol (PDB). Cartoon structure of the secondary structure of: (a1) proaerolysin and its different domains, the N and C terminal are also shown. (a2) aerolysin pore in membrane. Color code per monomer unit, side and top view. (b) View of the charged state of the pore. Blue = positive amino acids, red = negative amino acids. From left to right: side and top view. (c) Sagittal section. From left to right: solvent accessible surface (one can notice a constriction ~30 Å below the channel entrance). The charged state and corresponding detailed charged amino acids.

III. Mycobacterial porins

Mycrobacteria are a particular family of are gram negative bacteria. Like most gramnegative bacteria, they contain an outer layer wall membrane, however, in contrast to the other bacteria, they possess the thickest cell wall known [89], with the characteristic of an uncommonly low permeability, which gives a strong protection against the external environment. One of the main examples of mycobacteria-caused diseases is *leprosy* and *tubercolosis* which is one of the main reasons that mycobacteria started to be studied [89]. This membrane is very important for their survival, as it helps them to be resistant to several antibiotics. Due to the length of the mycolic acids, the outer membrane reaches up to 9-10 nm thickness [89], [90]. The mycobacteria survival is conditioned by this low cell membrane permeability, however nutrients must also be able to traverse the membrane, and among the membrane proteins allowing this passage (of nutrients) are porins. Porins allow diffusion and selective passage of hydrophilic nutrients needed by the mycobacterium. Porins also seem to be the major pathway for drug delivery. As a consequence, research to explore the regulation mechanisms of these porins is of high relevancy for understanding their permeability in the context of efficient drug delivery.

In contrast to other pore proteins, porins appear to not be highly specific and seem to contain no ligand-binding sites [91]. They allow diffusive passage of some hydrophilic molecules through the membrane [92], the process of porin's selectivity is then determined by their size, which varies from one to the other, but is also dependent on the shape of the channel and the nature of the amino acids within it. Therefore the specific local charge, the hydrophobicity of the porin and other specific interactions, play a role in the molecule transport and modulate the degree of specificity [91].

The only porin used in our experiments is the *Mycobacterium Smegmatis porin A* (MspA). This porin has served as a model for the study of porins in general and has in consequence been more investigated than other porins which makes it more convenient to use in our experiments. In addition its particular geometry, with the very narrow constriction at its bottom, which we will explicit later in this chapter, is interesting to use in the frame of polymer sensing. However, this porin is more difficult to handle experimentally than the previously described PFTs (α -hemolysin and aerolysin). MspA is purified from *E.Coli* and subsequently attached to detergent micelle, which can disturb the membrane or increase the concentration of insertion in comparison to PFTs.

One has to consider additional parameters when performing experiments with such proteins. This is also why we performed a small number of experiments with this protein pore.

1. MspA

The MspA porin is secreted by the bacterium *Mycobacterium smegmatis*. It was the first mycobacterium porin for which the atomic structure was studied and documented using X-Ray crystallography [93]. This bacterium has the advantage of being less pathogenic than other bacteria like *Mycobacterium tuberculosis*, and can serve as lab-based model for the study of common aspects of mycobacteria like the cell wall structure and its permeability. MspA is the major porin involved in the transport of solutes for the *Mycobacterium smegmatis* [89].

The X-Ray study of its structure with a resolution of 3.0 Å [93] revealed an octameric oligomer, with an eightfold rotation symmetry. The porin has a goblet-like shape with its narrowest point at its base. The height of the porin is 96 Å and its largest outer diameter is 88 Å, (Figure 17.a). The inner diameter of the channel varies from 48 Å at its widest opening to 10 Å (Figure 17.c) at the constriction site.

The wild type channel is mostly negatively charged, especially near the narrow constriction site as it can be seen in Figure 17.b and c. This considerably limits its application for the study of negatively charged polymers. In particular, this is the case for DNA interaction. As both are negatively charged, repulsion between the pore and DNA becomes the dominant effect. As a consequence, for DNA sequencing purposes (see chapter 1 for further information on DNA sequencing) the MspA nanopore needs to be mutated in order to obtain residues of positive charges within the channel that would interact with the negatively charged molecule.

Authors of refs [45], [46], [94] showed different versions of the mutated pore, with different sets of charges. In particular, they demonstrated possible interaction of DNA with a mutant MspA, which mostly contains positive charges in the channel part, and that it may be possible to observe sufficiently long enough DNA produced interaction. This is important in the context of their studies because the resolution of detection of DNA bases, is dependent on the duration of the interaction. They subsequently used this pore for DNA sequencing purposes and it showed high potential for this purpose.







Figure 17: Simulation of the atomic structure of MspA using Jmol (PDB). Cartoon structure of the secondary structure side and top view (b) View of charged state of the pore. Blue = positive amino acids, red = negative amino acids. From left to right: side, top and bottom view. (c) sagittal section. From left to right: charged state and corresponding detailed charged amino acids.

IV. Summary

We have presented the three different pores and their various properties. α -hemolysin has been one of the main pores used for pore sensing purposes, however, because of their particularities both aerolysin and MspA have high potential in fundamental research and biotechnological applications. The high stability of α -hemolysin and aerolysin is an important parameter because it allows to use them in lipid bilayer experiments for long durations, and permits the gathering of more data. MspA shows also high potential in detection because of its very interesting geometry, it has already proven to be a powerful tool in the sequencing of DNA, and seems to be a good candidate for the study of polyelectrolytes in general. We will present these various pores further more in the next experimental chapter, and explicit their interest for the different molecules probed.

Chapter 3: Experimental methods

The experimental methods used in our experiments are electrophysiological measurements applied to *in vitro* planar lipid bilayers. Earlier methods of measurements of single ion channels consisted of patch clamp experiments, where measurements of single channels in single cells were carried out using a micropipette. In parallel the technique of in vitro lipid bilayer painting and measurements started to be developed in the 1960s and 1970s by Mueller and Montal [95], [96] and has been widely applied over the last decades for various types of applications, among which drug delivery tests and electrical characterization of various cell channels among which PFTs and porins described in the previous chapters . The experiments in this thesis discuss the application of such a method on the interaction of pore forming toxins (PFTs) and porins with various analytes, mainly biological or synthetic polymers. We will present and describe in detail this method in the following chapter.

Because lipids are one of the basis of our experiments, we will first discuss their selforganization and the various structures that are undertaken by lipid bilayers in general. In the second part we will explain in detail the formation of *in vitro* lipid bilayers and the lipid painting method starting by the classical method with the use of cuvettes, and subsequently a more modern method namely, the use of MECA-chips and orbit system which consists in the use of a chip-array that allows higher statistics and better resolution. Finally, we will put into perspective the use of these type of experiments in the measure of ion channels and polymer sensing, and explain the technique used for data analysis.

I. Lipid types and self-assembly

Lipids are molecules that are one of the main basis of living organisms, through selfassembly they create structures which act as a "shell" for the cells, and partly isolates them from the external environment, which makes the whole process of cell formation possible. They are also the basis of communication taking place between cells being the support of various transfer processes. Lipid molecules have an amphipathic nature, as they are constituted by a hydrophilic head and a hydrophobic tail. Therefore, they usually self-assemble in aqueous environments to minimize their configuration energy, and form structures for which they tend to minimize the contact of their hydrophobic tails to water.



Figure 18: Schematics of the various lipid shapes (left) and corresponding spatial organization in solutions (right). Figure extracted from [97]

The different structures that can be undertaken by the lipids depend on the geometrical state of the single lipid molecules, which plays a role for the final assembly and structure formation. The geometry of lipid and their structure depends on a set of interactions (hydrophobic, electrostatic, entropic, steric) [98], [99]. Single lipid geometries are classified into cylinders, cones and inverse cones, consequently leading to formation of lamellar, micellar or cubic organizations respectively (Figure 18). The classification of lamellar structures includes bilayer structures, which is a combination of two monolayers together. These structures can be described by their pressure profile depending on the distance from head to head.



Figure 19: Extracted from [99] (a) Schematics of symmetrical lipid bilayer with representation of repulsive and attractive attractions. Note the fact that attraction happens between the head and aqueous solution, and positive pressure arises due to surface tension of the hydrophobic lateral chains with water, whereas repulsion happens between similar lipid parts due to short range steric forces (head-head and chain-chain) (b) corresponding pressure profile along the bilayer with negative and positive values representing attraction and repulsion respectively.

In symmetrical bilayers the profile consists of a repulsive interaction due to steric effects and entropy between the heads of the lipids, attractive interactions between the head and tail of the lipids which translates into negative pressure values and again a dominant repulsion between the lipid tails [98], [99] (Figure 19). Asymmetrical bilayers mixtures can also be found in nature in cells [100], [101] or reconstituted experimentally using mixtures of different types of lipids, in this type of assembly of monolayers that are not of the same nature, the pressure profile can become asymmetrical along the lipid bilayer, and some rearrangements happen to make the bilayer stable. These rearrangements give rise to the curvature of the bilayer, forcing the lipid layers to rearrange and reorganize in order to have a symmetrical pressure profile which is more favorable energetically [98] (Figure 20).



Figure 20: extracted from [98] schematics of rearrangement of asymmetrical monolayers. Individual lipid molecules of the monolayer change their shape and fit a symmetrical pressure profile resulting in a curved bilayer (left side). The corresponding pressure profiles along the bilayer are depicted on the right side.

It is possible to reconstitute membranes of different types of phospholipids in our experiments [102], [103]) and depending on their nature, their formation which relies on their stability *in vitro*, will be more or less difficult. The choice of a certain type of lipids is mainly important for the aiming experiment. Other studies where the aim is to study to study reconstituted proteins in original cell membranes, one has to fuse the micelles with pre-formed lipid bilayers in vitro, in such a system, the membranes are then the asymmetrical previously described profiles takes place.

For our experiments, the main investigation is on the pore-polymer interaction, for that reason the lipids, need to be as stable and simple as possible and they need to be in symmetrical configuration because they are not the object of the study, for this reason we used the synthetic lipid 2-dyphytanoil-*sn*-glycero-3-phosphocoline (DPhPC), also called diphytanoyl phosphatidylcholine, provided by Avanti-Polar Lipids Inc (Alabaster, Alabama. USA) in powder form. This amphipathic lipid is neutral and constituted of choline headgroups and diphytanoyl fatty acid chains. To obtain lipid solutions, the powder is dissolved either in octane or decane to obtain final lipid solutions of 5-10mg/mL which is optimal in our experiment for the formation of bilayers and minimization of multilayers.

DPhPC are mechanically and chemically very stable and last over long periods of time (low degradation), and no gel (ordered phase) to liquid (disordered phase) transition was observed between -120°C to 120°C [104] which makes them very convenient for experiments involving complex polymer-pore interaction or extreme experimental conditions. It is possible for example to perform experiments with extreme pH conditions (very acidic pH~4.5 and or asymmetric pH conditions for example, as it has been shown in ref [76]), asymmetrical concentration conditions inducing high osmotic pressure on the membrane surface, or high temperatures as shown in ref. [78] with still stable membranes.



Figure 21: Chemical structure of the 2-dyphytanoil-*sn*-glycero-3-phosphocoline (DPhPC) molecule.

We just presented the essential features of lipids in general and the type of lipid we use in our experiments. In the following section we will focus on the techniques of lipid painting which is the basis of our experimentation and how the properties of lipids and their structure are involved in this technique.

II. Planar lipid bilayer painting and methods of current measurement

Lipid bilayer painting was originally developed in the early 1960s by Mueller *et al.* [95]. The system consists of two compartments, called *cis* and *trans* side separated by a wall of hydrophobic material that contains a small aperture (~50 to 200 μ m). Figure 22 shows the configuration of the experiment, the *cis* and *trans* compartments are filled with an electrolyte, which is a conducting solution of salt. In each compartment, a silver-chloride coated silver electrode (Ag-AgCl) is inserted (Figure 22) for the application of voltage and the measurement of

currents. The lipids, pre-dissolved in organic solvents, are spread with a "painting" tool (which can be a painting brush of a pipette tip) on top of the aperture, and spontaneously form a lipid bilayer (as shown in Figure 22). After bilayer formation ions do not diffuse freely through the membrane therefore no current is measured, however after ion channel insertion it is possible to measure a certain current (under a fixed applied voltage). In this paragraph, we will discuss standard measurements and the particular techniques used for lipid bilayer formation namely the classical technique and the more recent technique (MECA chip), and we will also explain the process of electrochemical measurements with AgCl electrodes.



Figure 22: The steps of lipid bilayer painting method. Figure not to scale. Both compartments are filled with an aqueous solution of salt. The first step of painting with the painting brush (equivalent to bubble lipid painting) and formation of bilayer, however this may give rise to multi-layer formation; the annulus (represented with the orange triangle at the lipid-support junction) represents the solvent stabilisation junction (see text for more details) between the septum and the bilayer. AgCl electrodes are inserted for current measurement and voltage application.

1. Classical measurements

In practice, for the classical measuring technique, we use cuvettes (Warner Instruments), made of poly(tetrafluoroethylene) (PTFE or Teflon) which is a hydrophobic material that serves as a surface for lipid adhesion. The cuvettes contain an aperture, as shown in Figure 23, and one cuvette is placed in a holder (black holders in Figure 23) in order to have two separate individual compartments. The lipids are dissolved in organic solvents. The choice of solvent is important, depending on the method and material used. In the case of PTFE cuvettes, decane is the most stabilizing for the lipids. The cuvette that contains the aperture is pre-painted with the same lipid solution, and left to dry for 30 min. This process helps for the adhesion of lipids later on during the painting process. The two cuvettes are then filled with 500 μ L to 1 mL of salt solution, depending on their volume.

We will next present the various steps and features of the lipid painting technique, and the behaviour of lipid bilayer under the influence of ion movement.



Figure 23: Warner Instruments cuvettes. Smaller cuvettes contain up to 500 μ L and bigger ones can contain up to 1 mL. Zoom out (right) represents the aperture placement. Black holders are the support of the cuvettes. Teflon coated magnets are for solution stirring (white cylinders) and orange cylinders are junctions to avoid liquid leakage.

1.1. The "bubble" painting method

In our experiments we use the "bubble" painting method which is equivalent to painting with a brush but minimizes the amount of lipid solution deposited on the surface. By dipping a pipette in the lipid solution without withdrawing a volume, it is possible to minimize the amount of taken lipid solution to ~ 0.1μ L, avoiding potential aperture clogging. A bubble is then created near the aperture where the lipids self-assemble on the septum, create a bridge between the two ends of the surface and in most cases this creates a thick film (multi-layer). The next step consists in using the Montal-Mueller technique to thin out the film [96]. This method lies in the principle of monolayers created at the water-air interface of each side of the aperture coming in contact, by lowering and raising the solution level. Indeed, when the volume is lowered, a monolayer is formed on top of the water surface and the thick film is usually destroyed. By raising the volume once more, the monolayers on the surface of water come in contact at the aperture level and form bonds between their hydrocarbon chains to finally form a bilayer (see Figure 24.b). In their work, Montal and Mueller assume that the bilayer is solvent free.



Figure 24: a) extracted from [96]. Original scheme of the in vitro lipid bilayer measurement technique. b) Lipid thining out technique, where the falling and rising of water level helps the bilayer formation.

However, it has been demonstrated later in ref. [105] that there is always remaining solvent at the septum edge, called the annulus (*cf.* Figure 25), creating a junction between the bilayer and the septum, which is of importance for bilayer stability. White *et al.* in ref. [105] have shown that

due to the fact that the septum is ten of μ m and the bilayer is only ~5nm thick, it is essential to have a transition at the septum's edge as shown in Figure 25, otherwise the membrane would not be able to form.



Figure 25: Extracted from [105] (a) representation of stages of bilayer formation from thick film to thinning out into a bilayer at the aperture level. (b) the critical point of view of the authors of ref. [105] on the description of the method of ref. [96] showing the fact that the septum is much larger than the bilayer thickness, and explicating the existence of an annulus and a curvature of the lipid at the septum-lipid junction.

1.2. Electrical and optical properties of the formed lipid membrane

As a consequence of the membrane formation, it is not possible for ions to move freely from one side to the other and the membrane acts like a capacitor with the accumulation of charges on its surface. In this way, it is possible to measure the membrane electrical properties and its thickness can be estimated through electrical measurements and automatized calculations using softwares, which provide real time values of the capacitance and resistance. The electrical response is shown in Figure 26.C and shows the typical observed response of a capacitor. The capacitance is usually of the order of ~pF and resistance of ~GOh. Thinner membranes usually see their capacitance values increase following the formula $C = k\epsilon_0 \frac{A}{d}$ (where k the dielectric permitivity of the membrane, ϵ_0 the vacuum permittivity, A the membrane's area and d its thickness), whereas lower resistance values point to thinner membranes (easier for ions to diffuse in thinner membranes).

The lipid bilayer also has some optical properties, that can be observed by placing the system under the microscope. The formed bilayer is called black lipid membrane (BLM), because it appears black when observed under a microscope as shown in Figure 26. As the bilayer is ~5nm thick, opposite sides of the bilayer interfere destructively and therefore appears to be black. It is also a good test (if an optical setup is coupled to the experiments) for checking if the bilayers are really formed (as opposed to multilayers which would not appear to be black but would reemit some light).


Figure 26: extracted from [106]. (A) and (B) Optical pictures of membranes. One can clearly see the black membrane characteristic. (C) Corresponding electrical response which is the typical one of a capacitor (3 is a zoom out of 1, and 2 represents the applied voltage).

1.3. Measurement using silver-chloride electrodes

In order to be able to follow the current travelling from both sides of the membrane and apply potentials one needs to use a special configuration because the measurements are performed in solution, thus the use of electrochemical-based electrodes. There exist various types of electrodes for this purpose (copper, zinc, silver) but one of the most convenient in this case are the silver-chloride electrodes, because of the use of chloride ions and the fact that they are less polluting for the solution. Measurements using silver-chloride electrodes rely on electrochemical measurements of reduction and oxidation reactions. The current transfers between the solution and the silver wires happens through the Red/Ox reaction, taking place at the interface of the silver chloride electrodes with ions in solution. The corresponding chemical reaction is written as below:

$$AgCl + e^{-} \rightleftharpoons Ag^{+} + Cl^{-}$$

The reaction transfers the chloride ions, which diffuse to the opposite electrode and where the inverse reaction happens. The silver electrodes are connected to a headstage, which in turn is connected to an electronic amplifier (Axopatch Texas instruments) and monitored by a software. The measured current is directly proportional to the chloride ions mobility and gives a direct measurement of the current through the membrane.

We overviewed so far, the classical technique for bilayer painting and all the parameters involved for these measurements, but more recently the classical technique of lipid painting has been developed by Ionera Technlogies (Freiburg, Germany) to rely on the use of chips and multichannel arrays (which is used in our experiments), simplifies the process of lipid painting, increases reproducibility and allows better statistics. This technique will be presented it in the following section.

2. Recent method of lipid painting: measurement with MECA-chips

The other method used, is similar in principle to the classical painting technique: there are two compartments separated by a lipid bilayer and the same electrochemical process of current measurement with AgCl electrodes is used. This technology, that was developed by Baaken and collaborators and described in [106]–[108], consists of a chip containing multiple cavities, called multi-electrical channel array (MECA) with a version (which is the one used in our experiments) containing 16 cavities in a 4x4 array. The bottom of these cavities consists of a glass base and on top, a titanium adhesion layer of 20 nm and a 200 nm gold layer. The whole chip is coated with the SU8 polymer (except on the top of the cavities) which is hydrophobic and serves for the lipid painting. A layer of silver is electrochemically deposited on the bottom of the cavities for the transfer of current from the solution to the electrodes. In all the experiments described in this thesis the diameter of the cavities is $50 \,\mu$ m.

The use of chip arrays in such experiments facilitates the whole process of bilayer painting. First, because bilayer formation is automatized (described below) which renders easier the process of lipid bilayer formation and offers more uniformity within membranes structures (thickness, stability, ... etc) this features represent improvements in comparison to the classical technique, where the membranes are formed through a two-step process (painting and thinning out) which takes more effort. Secondly, this method offers a higher throughput: it is possible to access 16 channels simultaneously using the multi-channel amplifier. In addition, the use of smaller cavities (50 μ m diameter, 8 μ m deep), in comparison to ~1 mL cuvettes induces an enhanced signal to noise ratio reaching an average of ~2 pA intensity noise value when no voltage is applied.

2.1. Lipid bilayer formation on a chip

The chip is placed on a holder, a PTFE chamber is placed on top of the chip with a silicone ring junction in between to avoid liquid leakage. The whole system is held by a board that connects the chip electrodes with the amplifier. The chamber can hold up to 400 μ L. However optimally, we use a solution of 150 μ L for best results in terms of membrane stability. The painting method described in ref. [106], consists in spreading the lipids horizontally using a magnet (instead of vertically in classical setups). A drop of lipid solution ~0.1 μ L is deposited on top of the chip, near the cavities (see white cross in Figure 27) and spread in a circular motion using a PTFE coated magnet moving under the influence of the magnetic force induced by a counter magnet placed under the chip holder. The experiments shown in our case were all performed for optimum results with lipids dissolved in octane solvent. However, it is also possible to use lipids dissolved in other organic solvents with this method if needed.



Figure 27: extracted from [106]. Scheme of bilayer painting on the chips. (A) The board (green and orange pins) is connected to the chip. The figure shows the placement of the stirring magnet and counter magnet. (B) axis and direction of rotation of the magnet on top of the chip. (C) Spreading and thinning out of the lipid (yellow) by the magnet (grey) of top of the cavity.

2.2. Platforms for electrical measurements

As described in Del Rio Martinez *et al.*[106], parallelized recordings are performed using an Orbit 16 platform developed by Nanion Technologies (Munich, Germany), dedicated to parallelized bilayer recordings on the MECA chip. Thus, parallelized measurements were achieved through the use of a Tecella 16-multichannel amplifier (Foothill Ranch, CA, USA) with a feedback resistor (MOh) integrated into the orbit 16. Measurements and voltage application were controlled via the Tecella Lab software vs 0.84.

High resolution recordings on individual array units were performed in the orbit 16 using an external patch clamp amplifier (Axopatch 200B, Axon Instruments, Foster City, CA, USA), that could be connected to each individual microelectrode via an electronic switch board built within the orbit 16 and controlled using a software designed by Nanion Technologies, allowing switching between single channels.



Figure 28: Orbit 16 (right) and the Axopatch amplifier (left - Texas Instruments) setup.



Chip holder

b



Figure 29: (a) Chip as placed on top of the holder. (b-left) board with connection pins containing PTFE chamber (white) with silicone ring (black) will be placed (b-right) on top of the chip.

High-resolution measurements were made on a board platform and with manual designed system for lipid formation (with an automatic system of rotation of counter magnet). Similarly, each individual microelectrode was measured independently by connecting the headstage (CV 203BU, Axon Instruments, Foster City, CA, USA) and manually switching the connection with the electrodes on the chip surface. (as demonstrated in Figure 31). The headstage is in turn connected to the same patch-clamp amplifier earlier described. The whole system is isolated from external noise (~50Hz noise) with a Warner Instruments Faraday Cage. This allows reducing the noise-signal ratio.

All high resolution current measurements were filtered with 8-pole Bessel filter at cutoff frequency 5-10kHz with all orbit 16 measurements and digitization (sampling rate) set to 200kHz using a NI-PCI 6221 AD-converter (National Instruments, Austin, TX, USA), and filtering of 50kHz with the manual high-resolution setup with a digitization of 500kHz.



Figure 30: High-resolution setup. The headstage is connected to an amplifier. Below the green board is connected to the chip electrodes. The platform is set on the rotor of the counter magnet (dark platform)



Figure 31: Close up into the board-chip and headstage connection. (top) view of the board, each number shows a connection to one of the chip electrodes. (bottom) zoom into the connection of the headstage to an individual electrode.

III. Nanopore insertion into membranes, analyte interaction

After lipid bilayer painting, a solution containing the purified monomers of the membrane proteins is added to obtain protein insertion into the membrane. Ion channels are subsequently formed post-membrane insertion. In the case of PFTs, the formation of channels is usually spontaneous: active monomers self-assemble after addition of the monomers solution.

As an example, Figure 32 shows the step-like profile of a PFT insertion into a bilayer. Indeed, penetration of each ion is monitored by a single step. Cumulative signal of conductance in the step-like profile indicates successive penetrations of each single ion into the bilayer.



Figure 32: Successive insertion of three hemolysin channels, the noise increases proportionally to the number of channel inserted. The insertion occurs in a discrete fashion. Extracted from [77]

In the case of hemolysin, aerolysin, the following resistances in various KCl concentrations are given in Table 2:

KCl	300mM	1M	3M	4M
α-hemolysin	4 GOh	1 GOh	500 MOh	340 MOh
Aerolysin	5 GOh	2 GOh	600 MOh	420 MOh

Table 2: different resistances of hemolysin and aerolysin in different concentrations of KCI

Addition of analytes which might interact with the pore can be observed subsequently, by a momentary increase of resistance (lowering of the current for the same applied voltage), also called an event (see Figure 33). The duration and amplitude of current blockade due the interaction or event at given conditions, are proper to each analyte. The statistical analysis of the events characteristics gives information about the type of interaction with pore and its dynamics [7] however they can also give insight into the type and behavior of analytes in solution. This is the case for the sizing of polymers in high salt concentration solutions [33], [36], [107], or for the conformation of polymers in various salt types [34], [37]. The inverse can also be useful such that polymers can be used to size the α -hemolysin geometry [29].



Figure 33: Extracted from [7]. Examples of events during interaction of different polymers with the α -hemolysin nanopore. (b) the interaction is short (few 100 μ s) (c) longer interaction (~s)

IV. Data analysis

Recordings were performed using Clampex from Molecular Devices Inc. for the classical setup, and Gepulse for the experiments with the MECA chip, which is a software developed by Michael Pusch from the Italian National Research Council. The measurement of current is performed via sampling (measurements of a point every t_i), the sampling rate for the classical measurements, was set to 250 kHz and to either 200 kHz or 500 kHz for the MECA chip measurements (depending on the filtering value of the external amplifier). The recordings are then imported in the .ABF format.

Blockade detection and characterization can be performed using various softwares (Matlab, Igor Pro, etc..). For these experiments, we used two different approaches in order to analyze the statistical data: an Igor-Pro algorithm as explained in [109] for the data that was generated in the university of Cergy-Pontoise, and a Labview-based software for experiments performed in the University of Freiburg described in [108] (Sup. Info.). This analysis generates in all cases characteristics of the recording and the blockades like: open pore current, noise of the measurement, dwell time and amplitude of each event and this analysis can be extended to the analysis of specific phenomena by the modification of the algorithms. We will present both methods and explain in each of them the various steps of this analysis.

1. Igor-Pro algorithm

The measured baseline current (open pore current), is characterized by an average value I_o and standard deviation σ . This algorithm is based on a first step of noise rejection, because the events happen outside the noise region, followed by the event detection using a second threshold. The rejection consists in setting an arbitrary threshold $n\sigma$, which permits to ignore all the fluctuations of current (in respect to I_o) with an amplitude smaller than $n\sigma$. A value of n = 3 or 4, in general can be used in most cases and allows to reject all the noise, however, if the recordings contains small amplitude events, then this value of n might be too high, so one has to take into accounts the particular characteristic of each recording. It is also possible to set a second arbitrary threshold after the step of noise rejection, in order to detect a specific type of events, and reject others of smaller amplitude, the value of the second threshold would then depend entirely on the type of events for the specific trace. In the case of big fluctuations of the baseline I_o (bigger than ~ 2 or 3σ), the method of detection described can be applied on small windows, whose size can vary as explained in ref. [109] (100, 1000 points etc..) depending on the type of recording.

The event is detected from its start t_o (onset) to its end t_f (end of transition) as shown in Figure 34. Each event is characterized by its duration or dwell time T_t and its amplitude δI or residual current I_b/I_o where the beginning of the event corresponds to the average value I_o just before the transition to I_b the value of blocked current. It ends at the end of the plateau of the blockade as displayed in Figure 34. The events can also be characterized by the inter event time T_i of which the inverse corresponds to the frequency.



Figure 34: Event detection using the Igor-Pro algorithm. Black trace represent current recording over time and points represent sampling. I_o, I_b respectively represent the open pore and blocked current. The threshold for noise rejection $n\sigma$ is also represented. t_o and t_f are the beginning and end of the event, correspondingly, δI , T_t and T_i are the amplitude, dwell time and inter event parameters respectively.

2. Labview-based software: DetectEvent

This method differs from the previous one, in the fact that the threshold is based on the analysis of the derivative of the current or in other word its variations (as opposed to a detection based on the values of current). The method is based on the rapid detection of negative and positive differences between the onsets and end of transitions (upwards and downwards variations respectively). Considering in this case that the fluctucations of the baseline are minor, in comparison to the fluctuations due to the event transition. The values of open-pore and blocked state are determined by averaging the points between the end of one state transition and the onset of the next transition which would correspond in Figure 35 to the points between the yellow circles and red squares respectively for each value. The dwell time corresponds to the differences between in time between the end of the transition state and the onset of the next, however it is possible to refine via subsequent analysis to correspond to the same T_t that was described in the "Igor-Pro algorithm" section.



Figure 35: Screen shot of the DetectEvent software. Green trace represent current recording over time. Red squares and yellow circles represent respectively the onsets and end of transition. The bottom trace in red, represent the rectified signal (derivative). The red line represent the threshold set for the detection of events: all event which transition lies below the threshold are not detected.

3. Current histograms and characteristic dwell time analysis

After the detection, the information of each event is stored in tables for the Igor algorithm or ".txt" files after the DetectEvent detection. It is then possible to generate histograms which correspond to the amplitudes of the analyzed events. Figure 36 shows an example of the resulting histogram of a polydisperse solution of PEGs with a PEG 28 labeling, using the DetectEvent software.

The dwell time characterization is subject to various theories, in most cases the dwell time corresponding to a first order kinetic reaction between the polymer and the pore of a binding/unbinding reaction described by the on-rate k_{on} and off-rate k_{off} (respectively proportional to the dwell time and frequency of events), corresponds to an exponential distribution [10]. It seems indeed that the duration of events in this case is a random process. Therefore the distribution of dwell times can be fitted with a monoexponential $exp(-t/\tau)$ where τ is the characteristic time. A second method for the determination of the characteristic dwell time for monoexponential distributions which is described in [110], [111] and supplementary info of [35], is to plot the distribution of natural logarithms of the dwell times and fit it with the relationship

 $exp(x - exp(x)/\tau)$. This yields in general powerful fits, even for a low amount of detected events (short recordings).



Figure 36: Statistical analysis of events. top) left) histogram of residual current for PEG 1500 Da and AeL with the PEG 28 label (most probable peak). right) Corresponding scatter plot of the dwell time (for all events) vs I/lo. bottom) left) monoexponential distribution of the same data, and fit, counts axis in logarithmic scale. right) natural logarithm of the same dwell times and corresponding fit.

This analysis allows then a complete description of the events characteristics and because the interaction of the various pores with the studied analytes can be studied via this analysis, one can retrieve fundamental information using this method. In the next chapters we will present experiments which put in perspective all these methods and analysis with three pores, hemolysin, aerolysin and MspA with three types of polymers:

- Synthetic neutral polymers -Polyethylene Glycols or PEGs- for purposes of investigation of flows through the pore as well as sensing applications.

- Synthetic charged polymers (polyelectrolytes) for purposes of precision polymer sequencing.
- Natural polymers : ssDNA oligonucleotides, and investigation of their dynamics of interaction with the pore.

For each of these chapters, different pores were used because of their particularities, it is optimal to use each of them for different purposes and with different analytes for best results. We will discuss in each case the final results in relation to preceding studies, and their relevancy in the advance of the field of biological pore sensing.

Chapter 4: Interaction of PEGs and cyclodextrins with two protein channels

It is of key importance to understand and ultimately to be able to control the dynamics of analytes with biological nanopores. This is especially crucial for single molecule sequencing and single molecule size discrimination [6]. In addition it is essential to fundamentally understand the mechanisms playing a role in the nanopore-molecule interaction. As explained in chapter 1, the flows in the nanopores play an important part in polymers transport in nanopore experiments, For this reason, investigating various flows and forces in various nanoporesis critical to fully comprehend and explain the interactions between pores and analytes. This is what we aim at achieving in this chapter by studying the transport of molecules with the aerolysin pore which has been less investigated, especially in terms of flow and molecular transport mechanisms.

As explained in chapter 1, the electro-osmotic flow (EOF) *i.e.* the net flow of fluid in the nanopore is usually important to take into account for highly charged channels, which for example, has been shown to play a significant role in DNA transport into oppositely charged nanopores [112], for the translocation of proteins through silicon nitride nanopores [113], but even in biological nanopores, which usually exhibit a lower net charge, it has been shown that EOF can play a role in the transport and interaction [19]–[24], [114]. EOF can play a role in the interaction of small charged peptides with biological nanopores: in the unfolding and transport of small peptides in α-hemolysin [20] it was discussed as a low impact parameter in the capture of charged peptides, or in the trapping of a peptide in the hemolysin pore against the electrophoretic force [22]. In the case of proteins, EOF was shown to help the trapping of a rigid protein into the ClyAnanopore [21] and lately it has been proved that an engineered FraC nanopore creates an EOF which allows capture of peptides and proteins regardless of their charge [24]. It was also demonstrated that EOF was strongly involved in the transport of a cyclodextrin into the CymA pore [23] but it has also been shown to play little role in the transport of polyelectrolytes in the α hemolysin pore [76]. These various studies show the importance of such experiments, and that this effect varies strongly depending on the type of pore, and its charge distribution.

As discussed in the previous chapters, α -hemolysin has been the most widely used biological nanopore for such type of measurements. In contrast, aerolysin has been much less investigated in terms of channel flow and interactions with different analytes. It is only recently that aerolysin

was explored in various experimental and theoretical work. Some of this work includes the interaction of this pore with PEGs[35], proteins[85], [86] or oligomers [36], [47], [115] and it is only recently that some insight about its structure was published using theoretical work and cryoEm first in ref [83] and later with a more complete study in [84].

In the following chapter, to investigate the dynamics of interaction with the aerolysin nanopore, two widely studied model molecules are used for these experiments: small rigid neutral molecules named the α -cyclodextrins and long flexible neutral polymers named polyethylene glycol (PEGs). We compare these results with α -hemolysin in previous similar conditions of [19], [38]. Because the salt plays a role in the interaction, and net flows of the pores, we performe the experiments in two different salt conditions with two different cation types LiCl and KCl, in order to investigate their influence, these two cations have has been previously demonstated to influence the interaction differently [37], [38]. At the end of the chapter, is also discussed the sensitivity of the aerolysin to the PEG polymerization degree (molecular weight) in these same salt conditions.

I. Analytes : Polyethylene Glycol and Cyclodextrines

1. Polyethylene Glycol

The PolyEthyleneGlycol molecule PEG or PEO (PolyEthylene Oxide for higher molecular weights), is a polymer of chemical formula $(CH_2CH_2O)_n$. It has been a widely used molecule, in different areas of polymer chemistry but also in biology. These polymers are highly biocompatible and they have often been tested in various pharmaceutical and medical applications [116]. It was commonly known that because of its low protein adsorption, PEG can be used to transport compounds into organisms for a longer time, these molecules are known to be very soluble in water, where they usually adopt a swollen configuration (good solvent), however, it has also been observed in the 70s and 80s that they are able to hydrophobically interact with unfolded proteins (in contrast to their hydrophilic folded configuration) allowing in the access to the hydrophobic regions of the unfolded state [117], [118], or that for example tryptophan (an amino acid with most hydrophobic side chain) is more soluble in PEG solutions [119]. These molecules which are very soluble in water seem to be partly hydrophobic too. This led to discussions about the fact that PEG might also interact with biological nanopores via hydrophobic interactions [28].



Figure 37: Polyethylene Glycol Chemical Formula

In our experiments, we use PEGs of molecular weight M_w =1500, in high salt concentration conditions of 3M, where the solubility highly decreases as opposed to pure water (optimal solubility for PEG).

2. Cyclodextrins

Cyclodextrins are cyclic oligosaccharides which are cone shaped. They were first obtained via the degradation of starch by the glucan transferase enzyme and they exist as three types, α , β and γ , with 6, 7 and 8 glucopyranose units respectively, their characteristics are shown in the table below [120]:

Cyclodextrin	Number of glucopyranose	Molecular weight	Outer diameter,	Cavity d	iameter (nm)
	units	(g/mol)	(nm)	Inner rim	Outer rim
α	6	972	1.46	0.47	0.53
β	7	1135	1.54	0.60	0.65
γ	8	1297	1.75	0.75	0.83

Table 3: Physical charateristics of cyclodextrins

These molecules are hydrophilic on the external part of their structure and bear a hydrophobic core. These properties make them good candidates for coupling and transport of hydrophobic compounds in solution.

In the case of nanopores, β -cyclodextrins have been previously used as described in chapter 1, in the case of engineering the α -hemolysin pore to enhance DNA detection [121], they have also been used as neutral molecules to probe the electroosmotic force in the hemolysin pore in ref [19] and used to investigate the influence of alkali ions on the EOF through the same pore in ref [38] and recently α -cyclodextrins have been used to show the in the permeation of a neutral molecule through the CymA pore [23].



Figure 38: Cyclodextrin structure. chemical formula structure and side view (top) [122].chemical model obtained using Jmol (bottom)

II. Results

1. Interaction of cyclodextrin and aerolysin (AeL)

We started by investigating the directionality and voltage dependence of interaction of the cyclodextrin molecules in two salt conditions of the same ionic force: KCl and LiCl 3M. We probed the interaction of the rigid small α -cyclodextrins (α -CDs) in both salts, distinguishing between the two cases of addition from *trans* side of the membrane (*stem* and *cap* side of the pore).

First, one can observe in Figure 39 an asymmetrical interaction of the molecules with aerolysin: whatever the applied voltage (-130 to 130mV) there were no cyclodextrin-induced current blockades from the stem part (*trans* side of the membrane), and this in both salts. In contrast if the cyclodextrins were added on the cap side of the nanopore (*cis* side of the membrane) the majority of the current-blockades were observed for positive voltages, with almost no detected events in negative voltages. In Figure 39 (bottom) we observe long blockade duration in LiCl in negative voltage, but this kind of gating behaviour is not uncommon for the pore even in the absence of any kind of analyte. In our studies, we have systematically discarded such events, as they are not related to the interaction of the α CD with the pore.



Figure 39: Current–voltage (I–V) curves through AeL in 3 M KCI (top, blue) and in 3 M LiCI (down, red), in the presence of α CD added either from the trans-side (left) or from the cis-side (right) of the membrane. Schematics of recording the configuration are depicted on top of each curve. In each case the voltage was varied between –130 mV and +130 mV. Note the slight trans-ward rectification indicating the physiological insertion of the AeLnanopore from the cis-compartment. (*cap* of the pore in *cis* side)

One can notice already from Figure 39 seemingly longer α CD induced blockades in LiCl than in KCl, further analysis of the duration of blockades (dwell times, Figure 40) in both salts, first shows a linear dependency of dwell times as a function of voltage, independently of the salt type, we can remark though, different orders of magnitudes. In LiCl the dwell times are 10 folds more than in KCl (11ms *vs* 760µs at +90mV).A greater increase of dwell times*vs*voltage in LiCl

than KCl can also be seen (3fold/50mV *vs* 1.11fold/50mV). These results suggest a stronger force acting on the neutral molecules in LiCl.

First, the observation of asymmetrical interaction depending on the direction of interaction (*stem* or *cap*) with the pore suggests that the blockades are a result of reversible binding/unbinding reactions (first order kinetic reactions). This reactions are usually described by the on-rate (k_{on}) and off-rate (k_{off}) which are directly proportional to the dwell times and frequency of events respectively. It is customary in these experiments therefore to describe these reaction using the dwell time and frequency. The cyclodextrin binds in the pore from the cap side but is unable to be completely accommodated in the pore so it unbinds and returns into the bulk after dissociation directly from the cap side again, this last interpretation can be corroborated by the fact that the outer diameter of the α CD is 1.46nm [120] whereas the aerolysin pore was found to have a radius of 0.5nm [83]. Furthermore, the voltage dependency of interaction suggests that while the molecule is neutral, there is a dependency on the applied electrical field. This points to a possible driving force depending on the electrical field but that is not an electrical force, this force then can be explained by an electroosmotic flow (EOF) generated in the AeL nanopore.



Figure 40: AeL/αCD interaction as a function of voltage and salt nature.αCD were added from the cis-side of the membrane. (a) Nanoporecurrent traces at a constant applied voltage of +90 mV in 3 M KCI (blue) and in 3 M LiCI (red). (b) Blockade duration as a function of voltage in 3 M KCI (blue) and in 3 M LiCI (red). Insert on the right is an enlargement of the blockade duration vs. voltage curve in 3 M KCI. Dashed lines are linear fits

The pore lumen is charged at a pH=7.5 (which corresponds to our experimental conditions) and has a net charge of +7, coinciding with one positive unpaired charge per monomer and these charges play a role in the transport of ions and thus water molecules through the pore. As explained previously, the electroosmotic flow generates a plug flow in channels with uniformly distributed charges. However, biological nanopores are not uniformly charged, as it is formed of different amino acids. The diversity of amino acids explains the distribution heterogeneity of charges that is not found in the case of charged artificial channels. but biological nanopores are still ion selective because of this heterogeneity. As the ions moving inside the pore are transporting water molecules, the net flux of ions through the pore gives consequently the

direction of the electroosmotically driven water flow. Then, to determine this flow we have to measure the channel's ion selectivity.

2. Measurement of ion selectivity

2.1. Experimental setup

Experimentally, we measure the reverse potential to determine the ion selectivity in asymmetrical conditions of salt from each part of the membrane. In this conditions even if no potential difference is applied across the membrane we still measure a current flux, the reverse potential is deduced (in such conditions) from a difference between the value of current without any external applied potential and the value of necessary applied external voltage to reach a current value of zero.

One of the compartments is filled with a 1M (*cis*) and the other at 200mM (*trans*). The electrodes immersed into the medium need to be in agar salt bridges (and not directly in contact with the bath solution). We use agarose salt bridges containing 3M KCl to stabilize the avoid drifts. This gives rise to a junction potential at the interface between the salt bridge and the solution that we need to take into account to calculate the reverse potential V_r . We use for that purpose the Clampex 10.5 Junction Potential Calculator which takes into account the type of ions and their concentration in the solution and in the salt bridge, whichallows then the calculation of V_r through the Henderson equation.



Figure 41: Schematics of reverse potential measurement (not to scale)

The membrane potential V_r , that is deduced from the measured value V_E on the external amplifier (Axopatch Texas Instruments), is given by this relationship:

$$V_r = V_E + (V_L^i - V_L^o)$$
(38)

Where V_E is the experimentally measured potential on the axopatch, and $V_L^i - V_L^o$ is the potential difference due to the salt bridges in respectively *cis* (*1M*, V_L^i) and *trans* (0.2M, V_L^o) compartments.

2.2. Calculation and equation of J_w and P_+/P_-

We follow the same reasoning than in ref [19] for the determination ions permeability. As explained in the first chapter in equation (39), the electroosmotic velocity in a charged channel can be written as:

$$v_{EOF} = -\frac{\epsilon_{\omega}\epsilon_0}{\eta}\zeta E$$

In the case of narrower channels, it is not possible to reach the plug flow, and the velocity keeps increasing, never reaching its limit value described in chapter 1, which significantly reduces the overall flow.

Due to the difficulties faced in determining the value of the potential ζ , other approaches can be used, relating this velocity to the electrical flux through the pore's cross section instead of the applied electrical potential, as it has been suggested in ref [19]. v_{EOF} can thus be written as:

$$v_{EOF} = \frac{J_w}{c_w \pi r^2} \tag{40}$$

Where J_w is the water flux, r the radius of the pore, c_w is the number of water molecule per unit volume

It has been shown in the case of some cation selective channels that up to 7 water molecules move with the cation inside the pore [19]. Comparatively, in our case, ions move through wider channels like α -hemolysin and aerolysin. One can then consider that the ion moves in the channel with its primary hydration shell.

Therefore, it is possible to relate the water flow to the electrical flux of ions as below

$$J_w = N_w J_{\pm} \tag{41}$$

Where N_w is the number of water molecules surrounding the ion and J_{\pm} the flux of cations or anions. We can write the permeability ratios of cations and anions using the Goldman-Hodgkin-Katz (GHK) equation for reverse potential as:

$$\frac{P_{+}}{P_{-}} = \frac{[Cl^{-}]_{cis} - exp\left(-\frac{eV_{r}}{k_{B}T}\right)[Cl^{-}]_{trans}}{exp\left(-\frac{eV_{r}}{k_{B}T}\right)[K^{+}]_{cis} - [K^{+}]_{trans}}$$
(42)

Assuming that the fluxes of ions are independent (assumption of the (GHK) equation below), the net movement of ions through the pore becomes

$$J_{w} = N_{w} \frac{I}{e} \frac{(P_{+}/P_{-} - 1)}{(P_{+}/P_{-} + 1)}$$
(43)

2.3. Aerolysin selectivity and electroosmotic flow and comparison with α-hemolysin



Figure 42: Measurement of reverse potential in aerolysin.Experiment performed in assymetrical conditions (0.2M *vs* 1M at opposite sides of the membranes). Measured current *vs* voltage in KCI (top) and LiCI (bottom). Curves represent four independent measurements for both conditions.

From these results obtained in our experiments we deduce a selectivity of the pore for anions (results shown in Table 4), which means that the net electroosmotically driven water flow is in the direction of the chloride ions.

	$V_r(\mathrm{mV})$	P ₊ /P_
KCl	-20.1	0.27
LiCl	-24.5	0.19

Table 4: Values of reverse potential and ionic ratios through aerolysin in KCl and LiCl 3M

In our experimental setup the cyclodextrins are driven in the direction of the anions flux (interaction at positive voltage), which seems to confirm the fact that the EOF is the main driving force for these neutral molecules. As described in Table 4, we measure values $P_+/P_-(LiCl) =$ 0.19 and $P_+/P_-(KCl) = 0.27$, and using equation (22) for electroosmotic velocity, we obtain at 100mV $v_{EOF}(LiCl) \approx v_{EOF}(KCl) \approx 3cm. s^{-1}$ assuming the three types of ions (Cl⁻, K⁺, Li⁺) bear the same number of water molecules $N_w = 10$. This results which gives an order of magnitude, doesn't account for the longer duration of current blockades in LiCl than in KCl. The latter effect is, probably related to the fact that Li⁺ and K⁺ have a different hydration number. This value has often changed over the past decades:, experiments and calculation have shown in various references [123]–[127], a hydration shell number for Li⁺ between 4-6 whereas a hydration shell number between 4 and 8 for K⁺ [123], [125], [126] has been reported. Many of these values were measured using computer simulations (MD or MC) [123], [127]. A few theoretical calculations have shown a higher hydration number for Li⁺ of ~5 and for K⁺ of ~3 [124], [127]. As it has been discussed in the review of ref [125] experimental values measured using X-Ray and neutron scattering were often overestimated.in years before 1996. More refined analysis sometimes for pre-existing data finally showed a hydration number of 4 for Li⁺ and 6 for K⁺. Confinement may also play a role in the hydration of these ions, which is in that case is different from the bulk. Therefore one has to take into account these important factors.

3. Interaction of PEGs and aerolysin

Next, we performed experiments with PEGs in the same experimental conditions.We measured the interaction of the polymer depending on the direction of the electrical field.,First, similarly to the α -cyclodextrins, interactions are only observed when the polymer was added in the *cap* side of the pore and never from the *stem* side. However, in contrast to the previously discussed experiments, these interactions displayed opposite voltage-dependent behaviour in the two different salt types as depicted in Figure 43. In KCl, it is only in the case of a negatively applied voltage,that current blockades were observed. Inversely, in LiCl, the blockades were observed only with positively applied voltages as it was the case for the previously measured cyclodextrins. These results could suggest that, while the site of interaction might be the same for the PEG in this nanopore (interaction only from the cap side) the forces acting on PEG in the two different salt types are of a different nature and the complexation depends on the type of cation in solution too [37].



Figure 43: Current–voltage (I–V) curves through AeL in 3 M KCI (top, blue) and in 3 M LiCI (down, red) in the presence of PEG 1500 added from the cis-side of the membrane. Schematics of recording the configuration are depicted on the left. In each case the voltage was varied between –130 mV and +130 mV. Cartoons on top of each curve illustrate the direction of the cation flow and of EOF according to the voltage polarity.

To further investigate the different forces acting on the polymer we analysed the frequency and dwell time (characteristic duration of the blockades) as a function of the voltage. The obtained results shown in Figure 44 showed an exponential dependency of the frequency and dwell time in KCl and a linear dependency of those parameters in LiCl. This additionally indicated different mechanisms for the forces acting on the polymer in the different salts. First, in KCl the measured blockade duration is in the order of the millisecond, in contrast to the tens of microseconds in LiCl, suggesting an overall stronger binding to the pore wall in the case of KCl. In addition, the blockade duration increased exponentially of about 10fold in the case of KCl (~0.2ms at -40mV to 2.5ms at -160mV) and the frequency increased of about 4fold from -40mV to -160mV. However, the rate of events saturated above -130mV at a value of ~120Hz. This saturation coincided with a coincided with a value of dwell time of events that reached the same order of magnitude than the average inter-event time (*i.e* ~3ms vs 8ms) which indicated a limit in the rate of interaction (frequency). In LiCl, the dependency of the frequency and dwell time are linear with the applied voltage, the duration of events in creased of about 3 fold (40μ s at 60mV to 120μ s at 200mV) while the frequency of events increased of about 5 fold (80Hz at 60mV to 400Hz at 200mV). So here we can clearly observe a stronger dwell time increase for KCl vs LiCl (4.2 vs 1.07 fold per 50mV).but a similar increase for frequency (1.7 in both salts).

The exponential dependency of interaction of the frequency and dwell time and the direction of interaction (events observed only for a negative applied voltage) suggest an electrophoretic dominant driving force. Indeed, it is known that in salt solution PEGs might complex cations in various molecular dynamics (MD) and theory studies [128], [129] and it has been suggested in previous studies that the force acting on the PEGs is of an electrophoretic nature in the α -hemolysin nanopore. Consequently, one can also fit the frequency or rate of events using an Arrhenius rate equation as previously shown in the study by Kasianowicz *et al.*[39] and using the relationship (37), $R = R_0 e^{\frac{\Delta U - U_0}{k_B T}}$ described in the first chapter where ΔU is the electric energy of the polymer $|ze\Delta V|$. On the other hand, the dependence of interaction of PEGs on the positive voltage in LiCl, means thay they are transported similarly than the direction of the Cl⁻ flow, thus the EOF flow. This in addition to the non-exponential dependency of the frequency and dwell time (linear dependency) suggest an electroosmotic driving force in these conditions. As, it has been previously suggested in previous studies, the PEG has a weak affinity to complexation with Li⁺ cations, which would make it in a neutral polymer in LiCl. Consequently the EOF seems to be the only candidate which is still strong enough to drag the polymer inside the pore.



Figure 44: Dynamic properties of the AeL/PEG 1500 Da interaction as a function of voltage and salt nature. PEGs were added from the cis-side of the membrane. Schematics of recording the configuration are depicted on the left, with arrows indicating the direction of the cations flow and of EOF according to the voltage polarity. (Top, blue) Blockade duration and frequency as a function of negative voltage in 3 M KCI. Dashed lines are exponential fits. Note that for negative voltage magnitudes greater than130 mV, the blockade frequency saturated, as the mean inter-blockade duration (\approx 8ms) became of the same order of magnitude than the mean blockade duration (\gtrsim 1 ms). (Down, red) Blockade duration and frequency as a function of positive voltage in 3 M LiCI. Dashed lines are linear fits.

4. Interaction of PEGs and α-hemolysin

We reproduced the last series of experiments (with PEGs), using the α -hemolysin pore, for direct comparison with the previous experiments. Similarly to AeL, the measured magnitude of the electroosmotic velocity and the results of reverse potential (showing also a selectivity of the pore for anions), was previously described in ref. [38]. The measured EOF in α -hemolysin gave the following results of $v_{EOF}(LiCl) \approx 3v_{EOF}(KCl) \approx 3cm^{-1}$ at 100mV which are of the same order of magnitude of the results obtained with AeL. Figure 45 shows the IV curves of PEG and hemolysin added to the *cis* side of the membrane (cap side of the nanopore). First, in 3M KCl, in negative voltage corresponding to the direction of cations flow, higher frequencies of events were observed. In contrast, only few blockades were seen for positive voltages (top of Figure 45). In 3M LiCl, interactions were also observed however, in both directions of the applied voltage.



Figure 45: Current–voltage (I–V) curves through aHL in 3 M KCI (top, blue) and in 3 M LiCI (down, red) in the presence of PEG 1500 Da added from the cis-side of the membrane. Schematics of recording the configuration are depicted on the left. In each case the voltage was varied between –120 mV and +120 mV. Cartoons on top of each curve illustrate the direction of the cations flow and of EOF according to the voltage polarity.

The same analysis than in the case of aerolysin for the frequency and dwell time, showed an exponential behavior *vs* voltage in KCl (see Figure 46) similarly to the case of aerolysin in the same conditions. In LiCl the frequency analysed for the positive voltage (in order to have a direct comparison with aerolysin) showed a linear increase of about 3 fold from 40mV to 120m, with a following steady value of ~200s⁻¹ up to 200 mV. In KCl however, the frequency increases exponentially. The dwell times, in contrast to the frequency, decreased in KCl, while in LiCl there

was no clear change in the characteristic times (same value ~ 60μ s). The observed negative voltage-dependent interaction of PEG and the exponential behaviour of dwell times and frequency vs voltage suggest that an electrophoretic driving force in KCl acts on the PEG. These effects were also observed in other studies [37], [38], [72] and similar explanations were suggested: in all the references the observed exponential behaviour was explained by an electrophoretic dependent interaction and cation-PEG complexation. For all the cited references, the PEG was added from the stem side of the nanopore and effects similar to our observations for this interaction were reported in all cases. In ref.[37], the decrease of dwell times following its increase up to 50mVfor PEG 1500 Da in 4M RbCl, was suggested to be due to an energy barrier. This energy barrier is interpreted to be required for confinement of the whole polymer inside the pore, above which the whole polymer is accommodated inside the pore and is able to translocate. In ref [72], the longer dwell times for lower voltages, were interpreted as the enhancement of the PEG-pore binding due to the cation-PEG complexation, which suggests a dissociation of the cation as the voltage increases. However, the earlier described results of the interaction of AeL and PEGs (continuously increasing dwell times in KCl) tend to contradict parts of these interpretations, and suggest a complex mechanism that involve peg-cation complexation but also pore-dependent interaction influencing the decrease of dwell time in the case of hemolysin.



Figure 46: Dynamic properties of the aHL/PEG 1500 Da interaction as a function of voltage and salt nature. PEGs were added from the cis-side of the membrane. Schematics of recording the configuration are depicted on the left, with arrows indicating the direction of the cations flow and of EOF according to the voltage polarity. (top, blue) Blockade duration and frequency as a function of negative voltage in 3 M KCI. Dashed lines are exponential fits. (Down, red) Blockade duration and frequency as a function of positive voltage in 3 M LiCI. Dashed lines are linear fits.

In contrast to Ael, in LiCl, the observed behaviour of non-dependent voltage interaction in addition to the small increase of dwell time and frequency suggest that the PEGs behave as neutral molecules and that what we are observing could be mainly diffusion effects. The fact that blockades are observed in negative (in opposite direction of EOF) and positive (in the same direction of EOF) driving electrical fields, does not show clear indication that electroosmotic flow might play any role on the driving or interaction of the molecules. This suggests that the electroosmotic flow could be very weak in hemolysin or that the binding reaction between PEG and hemolysin in the presence of lithium cations is much weaker than in the case of potassium cations. This result proves again that the pore structure and the type of interaction with the polymer plays a crucial role in the influence of frequency and dwell time that we get.

In any case the interpretation of PEG-cation complexation is debatable as we also observe blockades for all directions of voltage whatever side of α -hemolysin the PEG is added as shown in

Figure 47. When PEG is added from the cap side (Figure 47) a steady frequency of interaction is observed at about ~ $20s^{-1}$ for positive voltages. Nevertheless, one can still observe a non-negligible amount of blockades for ΔV >0. The dwell times show a non-monotonic evolution with a maximum at 40mV similarly to when the PEG is added on the stem side. Moreover in ref [34] (when PEG is added from the stem side) whereas the dwell time shows a similar behaviour (maximum increase), the frequency seems to increase for positive voltages *i.e.* in the direction of anions flow, which contradicts the previous interpretation. Indeed, as shown here in both cases (addition from either side of the pore), there is always a non-negligible number of blockades in the direction of anions. One could then argue that in the first case here (addition from the cap side), the diffusion is still strong enough to push the polymer towards the pore (in the direction of anions) even if it is considered to be positively charged. But the observed increase in frequency for blockades in the direction of anions (PEG added from the stem side) observed in ref [34], cannot only be explained by a diffusion phenomenon. Additional experiments on this subject completing all the past experiments are required to explain the behavior of PEG.


Figure 47: Interaction of PEG with α -hemolysin from *stem* side, in various KCI concentrations from [34] (top) (A) k_{on} (dwell time normalized by the concentration) in s.M⁻¹ (B) k_{off} (frequency) in s⁻¹ vs voltage (from -0.2 V to 0.2V). Interaction of PEG (28 monomers) and α -hemolysin from *cap* side in KCI 3M (bottom) (C) Dwell time in μ s and (D) frequency in s⁻¹ vs voltage (from -0.09 V to +0.09 V)

5. Effect of salt on the discrimination of PEGs in aerolysin

We also performed the analysis of the various resistive pulses of the obtained data with polydisperse solutions of PEG1500 Da in 3M solutions of LiCl and KCl with the aerolysin pore. We obtained similar histograms in both LiCl and KCl than in ref [35], [33]. It seems that it is possible to obtain a precise polymer size discrimination at a lower polyelectrolyte molarity.

Figure 48 shows this analysis in and 3M LiCl. It seems still possible to discriminate up to 16-17 monomers, but the pore seems to be a little less sensitive to the bigger sizes. Here, we are

able to discriminate up to 45 monomers, but the peaks in the histogram get less sharp as the value I/Io due to polymer blocking the pore, gets smaller (i.e. more blocked current). It could be due to the fact that the polymer is a little more swollen in 3M KCl in comparison to 4M and that the pore is not able to accommodate the bigger sizes (above 45 monomers). In contrast, in 4M in ref [35] it was possible to observe up to 48 monomers size. Nevertheless, the discrimination is still almost as powerful in our conditions, probably because in 3M we are close to the limit of solubility: 34% in water for KCl [130] at 20°C corresponding to 4.58M, which seems to enhance the interaction of the polymer with the pore.

Surprisingly, we also observe size discrimination between the polymer molecular weights in LiCl, starting from a voltage of ~100mV. This discrimination becomes more significant as the voltage increases with the sharpest detection at 190mV (it is possible that it becomes better at voltages higher than 190mV but in order to avoid the membrane breakage we didn't go higher than this voltage, though it is also possible that the discrimination reaches a maximum optimum voltage as in the case of KCl 4M [35]). We remark that the pore is less sensitive to bigger polymers (>38monomers) in these conditions. But the pore seems to be even more sensitive for the discrimination of smaller polymers, than in the case of KCl 4M. The discrimination in LiCl reaches polymers that contain 12 monomers at 190mV, which corresponds to $M_W \approx 530Da$. This is even more surprising as the order of magnitude of blockade durations are 10 times less in LiCl than in KCl (100µs *vs* 1ms).

The first conclusion that we can draw from these results is that the type of interaction might be similar in both salts, or in other words, the same binding reaction is happening between the polymer and the pore in both salts as the discrimination still happens. However it is fairly obvious that the interaction is strongly dependent on the type of cation. Secondly, it is very surprising that in LiCl, the aerolysin is more sensitive to smaller polymers, given the fact that the LiCl solubility is higher than KCl in those condition (LiCl has a solubility of about 80% [130] at 20°C corresponding to 18.8 M, and that the observed dwell times are much lower than in KCl. The longer dwell times could be explained again by the fact that the polymer is more soluble in 3M LiCl than in 3M KCl. Therefore in KCl, a large proportion of water molecules is taken by ions which would make the polymers salt out of solution. This in turn makes the binding reaction stronger in KCl than LiCl. The latter effect of sensitivity for the various polymer sizes in LiCl would again be explained by a higher hydration of the polymers. First, the decrease of sensitivity for bigger polymers might be explained by the fact that PEG is in the conformation of a chain in good solvent (swollen): in that case it is more difficult to accommodate bigger chains into the pore. The increase of PEG size makes the dwell times shorter, which can be related to a similar observation described in ref [32]: an entropic-spring pull-out, where bigger polymers only enter the pore partially and this decreases the sensitivity of the pore for the bigger sizes in LiCl, and shortens the dwell times. In Figure 48, for LiCl 3M at 140mV and 190mV, one can clearly observe the decrease of dwell times as the polymers get bigger. In addition, the sensivity to smaller polymers could also be explained in such fashion where smaller polymers which are not possible to observe in the same KCl molarity, are probably in a more extended conformation in LiCl: this in turn suggests more blocked current in the pore. This latter effect would make their detection possible even though the dwell times are short.



Figure 48: Discrimination of polydisperse PEG ($M_w = 1500$ Da) with aerolysin in LiCl and KCl 3M, at 140mV and 190mV. Black curve (I/Io) represent residual conductance histograms with corresponding event counts (right axis). Highest peaks (28 and 32 monomers) are monodiperse markers for the detection. Red scatter plots represent dwell times (in ms, left axis in logarithmic scale) *vs* residual conductance. Dwell time axis is a logarithmic scale.

It is clear here that the mechanism of interaction is strongly dependent on the type of ions, but only a few studies give some insight to the influence of ions on the interaction. In ref [75] Rodrigues *et al.* studied the effect of halogens on the PEG- α -hemolysin interaction, using four different types of halogens (F, K, Br, I). They observed a Hoffmeister effect with an effect of hydration-dehydration, where the PEG and anions compete for water: as the ions get more hydrated, the PEG is salting out of solution more and the influence of these anions follows the Hoffmeister ranking according to their influence on the dwell time (increasing on rate constant/ dwell time). In Breton et *al.* [37] it is shown that the alkali ions type also influences the interaction, suggesting either a neutral or polyelectrolyte-like behaviour of the PEG. But more effort needs to be put into the explanation of such effects that need to be further investigated. It would also be interesting to probe the types of intermolecular interactions that are involved at the amino acid's scale in such type of interaction, this later idea could be further developed using MD simulation, as it has for example been shown in the case of DNA and hemolysin [131].

III. Summary

We have shown in the following chapter the importance of understanding the mechanism of interaction and transport of simpler model molecules with the nanopores. Studying the flow mechanisms in aerolysin showed different forces involved in the transport of these molecules, depending on the salt type and the chemical nature of the molecules. The use of cyclodetrines seemed to show that the EOF in aerolysin is of a much higher magnitude in comparison to what was previously demonstrated with α -hemolysin, and plays an important role in the transport of neutral molecules. This seemed to be validated for PEG in LiCl in contrast to KCl where the PEG behaved similarly to a positively charged molecule. Following this observations with PEG, some theories of complexation seemed to be confirmed but are still debatable, as we also see discrepancies, depending on the pore, and direction of interaction. More studies on the influence on the ion type, the interaction and properties of the pore are necessary in order to give a complete scheme of the interactions. This is especially important for applications of nanopores as sizing tools. Investigating the exact nature of the interaction (hydrophobic, electrostatic, etc...) is of high importance for this kind of experiments and possible future applications.

Chapter 5: Poly(phosphodiester)s detection and characterization

Until the last decade advances in polymerization processes and monomer sequences, were commonly focused on over the past 50-60 years over synthesis of homopolymers with the same monomer unit or minor developments of copolymer polymerization, like alternating, periodic, random or block copolymers [132]. However these processes lacked a certain degree of complexicity, [132]–[134]. In the aim of controlling the monomer sequence more precisely within synthetic polymers, nature has inspired polymer chemists into the development of complex polymers, similar to DNA, RNA or peptides: precision polymers appeared as the next development in polymer chemistry.

Following that the sequencing of precision polymers has seen a grown interest over the past decades. The possible applications related to such polymers gave a high appeal for researchers, with the ultimate goal of obtaining the most precise molecules which can be used in various fields, from bio-inspired molecules, which could reproduce folding processes of proteins for example, to data storage and molecules which in general might offer many prospects in material-science applications [132]–[134].

An interesting aspect of such polymers is the writing and storage of information. Indeed these precision polymers, are prepared using polymerization techniques of step by step growth and therefore, it is theoretically possible to write information with a high degree of precision and one of the major challenges is to be able to retrieve this information in the most precise way possible without destroying the molecule's structure. This gave rise to the idea of using nanopore tools in order to read or "sequence" these molecules as it has been done and still carried on with DNA.

In this contest we began experiments in collaboration with polymer chemists from the Institut Charles Sadron in Strasbourg (France), with whom we attempted at reading out the sequence of precision polymers. As explained in this chapter, because this field is relatively recent, the main aim was to find the best fitted polymers and conditions, for sequencing with biological nanopores. We carried nanopore experiments with various molecules and nanopores and we coupled this study to their characterization using various scattering techniques (X-ray and DLS) to investigate their conformation in solution and correlate it with the electrical characterization.

In the following sections we will begin by describing the process of synthesis of the polymers used in this study and explaining the scattering techniques used. We will proceed by first describing the various polymers studied using electrical detection, followed by the wave-scattering characterization for each molecule and finally the correlation between the two techniques.

I. Sequence controlled polymers

Sequence controlled polymers, has known increasing interest in the past decade. Various processes have been used to the preparation of these polymers. Polymerization techniques like chain-growth or step-growth polymerization were used to control comonomer sequences, however polymers obtained showed high polydispersity. New methods like solid phase iterative step growth have been used to control polymer sequences and their dispersity. In this process the chain is grown on a solid support and monomers attached one by one after the other [133]. Nowadays, oligonucleotides synthesis is based on this method (solid-phase synthesis) coupled to phosphoramidite chemistry. In the latter chemistry, a phosphate linkage is obtained by the reaction of a phosphoramidite with a hydroxyl group, giving a phosphite which is in turn oxidized to make a phosphate [135]. The whole process of oligonucleotides synthesis is nowadays quite fast, controlled and cheap especially compared to enzymatic synthesis, and one can synthesize specific sequences of various lengths with different mixtures of nucleobase and this usually with a good dispersity index. In our experiments we use polymers obtained with phosphoramidite coupling reaction. This chemistry has been developed and perfected over the past years in the area of biochemistry to synthesize oligonucleotides of DNA and RNA, this offered high potential for its use in synthetic polymer science. As described in Al-Ouahabi et. al [135], using AB-monomer type containing a phosphoramidite and dymetoxytrithyl (DMT) protective hydroxygroup, controlled muti-step growth can be achieved on a solid support, the whole process consists then of a phosphoramidite coupling, phosphite oxidation and deprotection of DMT. Monomers of this type are shown in Figure 49. The model polyphospho(diester)s of Al-Ouahabi et. al, were then further more developed and longer chain lengths have been synthesized with up to 100 monomers per chain [136].



Figure 49: extracted from [136] (a) molecular structure of the different comonomers used. (b) steps of the solid-phase step growth (the chain is grown on the red bead)

It appeared that such polymers possess a high potential in the storage of information, as shown in ref [136] and finding precise techniques for read out is a key importance, the authors used Electrospray Ionization Mass Spectrum analysis (ESI-MS) for this purpose, but this technique is desctructive, the polymer's link are broken on order to be read. In this context, it seemed that nanopores could be interesting candidates in the reading of such polymers and that, similarly to DNA, this technique might have high potential for the sequencing of synthetic polymers.

One of the major first challenges in this project was to find appropriate polymers, that would interact with biological nanopores which was not a trivial task. Throughout the project we have tried out various types of polymers, with various backbones and side chains, of which one type seemed to interact strongly with the nanopore was the polyphospho(diester)s. The first step in this study was to analyze and understand the interaction of a simple form of this polymer: the homopolymer containing the "0" monomer type (see Figure 49). This was followed by the study of similar structures with the introduction of more complex comonomers. We will next briefly

describe the scattering techniques of characterization used to describe the polymer dimensions in solution, and present subsequently for the each polymer that was studied the experiments with nanopores followed by their characterization using wave scattering.

II. Scattering Techniques for characterization

We combined DLS and XRay scattering for the characterization of the polymers. We describe both method and how to extract useful information about the polymer's conformation and dimensions.

1. Dynamic Light Scattering (DLS)

Dynamic light scattering is a method relying on quasi-elastic light scattering and can be used for the obtention the hydrodynamic radius of the polymers. The principle is that light (a laser) of wavelength λ is passed through a solution, and the intensity is collected at a fixed angle θ and the scattered light is recorded over time. The corresponding function I(t) is related to the concentration fluctuations of polymer in solution due their brownian motion. By computing the correlation of temporal fluctuations of intensity, it is therefore possible to extract a distribution of relaxation times of objects in the solution (related to the brownian motion) which is related to the diffusion coefficient and thus the hydrodynamic radius of the objects.

We have performed these experiments in collaboration with François Schosseler at the Institut Charles Sadron in Strasbourg, France. We detail below the main part of analysis and data analysis also described in ref. [137] used for the results of our polymer solution analysis :

The intensity correlation function (ICF) can be written as

$$g^{(2)}(t) = \frac{\langle I(t')I(t+t') \rangle}{\langle I(t') \rangle^2}$$
(44)

The corresponding field-field correlation function is written as below

$$g^{(1)}(t) = \frac{\langle E(t')E^*(t+t') \rangle}{\langle I(t') \rangle^2}$$
(45)

Where E(t) is the electromagnetic field function, and $E^*(t)$ its conjugate.

The two functions are related by the Siegert relationship

$$g^{(2)}(t) = 1 + \beta^2 \left| g^{(1)}(t) \right|^2 \tag{46}$$

Where β is a coefficient depending on the geometry of the measurement.

For polydisperse solutions the field-field correlation function is related to the distribution of correlation times R(t) as :

$$g^{(1)}(t) = \int_{t}^{+\infty} R(\tau) e^{-t/\tau} d\tau$$
(47)

Using CONTIN method, it is possible to transform the previous equation into its discrete form and use a pseudo-distribution of correlation times $P(\tau_i)$ over a chosen number of gridpoint values N_q

$$g^{(1)}(t) = \sum_{i}^{N_q} P(\tau_i) e^{-t/\tau_i}$$
(48)

Experimentally the ALV/DLS/SLS-5020F setup was used (ALV Laser Vertriebsgesellschaft mbH, Langen, Germany), which has a Ne-He laser (22mW) that was directed at the sample with a wavelength of $\lambda = 632.8$ nm. The scattered intensity was collected over time at a fixed angle $\theta = 90^{\circ}$ on a detector. The temperature was regulated to $25.2 \pm 0.1 \,^{\circ}$ C.

2. X-Ray Scattering

It is also possible to use X-Ray scattering for the characterization of the polymers in solution. The principle of this characterization also lies in the choice of the angle, in order to observe the polymer and its conformation in the given solution. We have performed a long range scattering analysis, Small and Wide Angle X-ray Scattering (SAXS and WAXS) however, we only analyzed the SAXS data which is sufficient for the first description of polymer's dimension and comparison with the nanopore experiments.

We have also performed these experiments at the Institut Charles Sadron in collaboration with Guillaume Fleith (experiments) and Michel Rawiso (data analysis and interpretation). The scattering technique was performed in a lab-based facility, which is powerful enough for simple characterization (gyration of radius, mass, etc..), but requires long duration of recordings to gather enough data points. In contrast, for information on the complete conformation of the polymer it is easier to use a synchrotron facility, where the radiation is much more powerful.

SAXS relies on coherent X-Ray scattering (Thompson)[138] the intensity of the diffraction is dictated by the electrons of the atom, acting as secondary sources of diffusion of the X-Ray waves as they oscillates under the influence of the beam's electrical field, reemitting it with the same wavelength but out of phase. The atomic scattering factors are dictated by the number of electrons, and the electrons orbitals. The intensity profile of the diffraction pattern as a function of the wave vector $q = \frac{2\pi}{\lambda}$ (λ is the used wavelength for the measurement, and $\lambda \approx 10$ Å) is written as:

$$I(q) = K^2 C[f_1(q) + C f_2(q)]$$
(49)

Where K^2 is the contrast and is written as a function of the electron density of the solvent ρ_s and of the polymer ρ , $K = (\rho - \rho_s) \frac{v}{N_a}$, v is the molar volume of the elementary diffuser (solvent molecule or monomer) and N_a is Avogadro's number, the formula for the electron density is $\rho = \frac{a}{v}N_a$ where $a = 0.282 \times 10^{-12} z$ (*cm*) is the and z is the number of electrons of the diffuser. C is the concentration of the polymer, $f_1(q)$ and $f_2(q)$ respectively represent the intramolecular (form factor) and the intermolecular interaction. At dilute regimes, we can neglect the interaction between molecules described by $f_2(q)$, therefore, in the following experiments we have set the concentration of the polymer at relatively low concentrations (0.5 to 1% which is much below the critical concentration of polymer overlapping c*), in this case the intensity of diffraction is only dependent on the isolated molecules in solution, and after normalization equation (49) becomes:

$$\frac{I(q)}{K^2 C} \approx f_1(q) \tag{50}$$

It is then possible to relate the experimentally obtained scattered intensity to $f_1(q)$. There is a number of theoretical description of the form factor $f_1(q)$, which can be used to fit the experimental data, depending on the range value of q which varies as a function of the angle of measurement.

X-ray analysis was performed over a long range for all the polymers. performing this analysis at small angle and wide angles (SAXS and WAXS) permits to acquire data from which it is possible to extract a full description of the polymer behavior in solution. Throughout this chapter, we have shown two different methods to fit the form factor curves :

- Using the Guinier approximation [138], [139] :

$$\frac{1}{f_1(q)} = \frac{1}{N} \left(1 + \frac{q^2 R_g^2}{3}\right) \tag{51}$$

and by plotting $\frac{1}{f_1(q)}$ as a function of the squared scattering vector q^2 , it is possible to deduce values of N and R_g , this approximation is valid for ranges $qR_g < 1$.

- Similarly, we can use the Debye approximation [138], [140]:

$$f_1(q) = NP(q) = \frac{2N}{q^2 R_g^2} (q^2 R_g^2 - 1 + e^{-q^2 R_g^2})$$
(52)

describing the form factor of an unfolded polymer. After plotting $f_1(q)$ as a function of q^2 it is possible to also obtain values of N and R_g . This is valid for ranges $qR_g < 3$.

III. Results

We will present in the following parts the interaction of poly(phosphodiester)s (homopolymers and block copolymer). Their comonomers are classified depending on their side groups, all comonomers contain a phosphate group as explained previously, we will present 3 different comonomer types:

- (0) Hydrogen, R=H
- (1) Alkyne, $R=CH_2-C\equiv CH$
- (2) Triazole, $R = CH_2 C_3H_3N_3$



Figure 50: The various structures of sequence controlled polymers studied. Polyphosphodiesters are classified by rows depending on their side group (R), all comonomers contain a phosphate. (0) Poly(phosphodiester)s with comonomer '0' (R=H), all are homopolymers. P1a and P1b contain the 3-thymidines primer sequence and respectively 56 and 104 comonomers, P1c contains 1-thymidine primer sequence and 64 comonomers. (1) Poly(phosphodiester)s with the 1-thymidine primer sequence comonomer '1' (R=CH2-C≡CH - alkyne). P3 is a homopolymer with 64 comonomers, B1 is a block-copolymer, alternating 4 blocks of comonomers '0' and '1' each of 16 units. (2) Poly(phosphodiester) with the 1-thymidine primer sequence and 64 comonomers '2' (R=CH₂-C₂H₃N₃ - triazole ring)

1. Poly(phosphodiester)s with comonomer (0) and the primer sequence of 3-thymidines (P1a and P1b)

1.1. Interaction with α-hemolysin and aerolysin

We first investigated the interaction of simple homopolymers with both nanopores, alphahemolysin and aerolysin. As described in ref [136] the polyphospho(diester)s contain 3 thymidines nucleobases as a primer-sequence followed by either 56 or 104 monomers of type (0) as the two side groups are hydrogens (see Figure 51). Because the phosphate group (PO_4^-) has a pKa of 2-3, it is expected to be completely deprotonated in our experimental conditions (physiological pH~7.5), therefore in the form of a polyelectrolyte that consequently can be driven under the influence of the electrical field, an important fact if one is to use nanopores for sequencing purposes.



Figure 51: extracted from [136]. Step growth on solid support of poly(phosphodiester)s with the 3-thymidine primer sequence.and comonomers "0" (side group R=H) or "1" (side group $R=CH_3$).

The experiments were carried out in low salt solution conditions (KCl 1M or 300mM solutions), where the polyelectrolyte is expected to be more soluble, therefore, to take a more stretched configuration, as opposed to a globular or more collapsed one. The globular conformation, usually is adopted in bad solvents or conditions of high salt concentration (ex: KCl 4M) especially for polyelectrolytes, where charges are highly screened by counterions, reducing monomer repulsion leading to chain size decrease described by simulations and experiments in various experimental conditions [141]–[146]. This conformation is obviously not favorable for

sequencing purposes because the polymer is collapsed and the information written on the chain is consequently difficult to access. Traditionally, the first DNA experiments in hemolysin, were carried out in 1M conditions [3] as well as the MspA experiments[45], [46], however some more recent experiments in both pores were carried out in lower salt concentration[41], [147]. In our case we already observed strong interaction under these conditions, with 50% current blockades that were only terminated by switching off or reversing the potential polarity, for these reasons we chose to reduce the concentration early on in the study to 300mM KCl but we will also discuss experiments we peformed in 1M KCl.

Figure 52 shows the interaction of polymer P1a (see Figure 50 and Figure 52 for polymer structure) with both α -hemolysin and aerolysin in 300mM KCl at ± 100 mV and the polymer was added to a final concentration of respectively 15 and 40μ M in the cap side of α -hemolysin and aerolysin. It is important to notice as showed in Figure 52 that first, interaction only happens at trans-positive voltages for both pores, and there is no observation of events at trans-negative voltages, this seems in line with the fact that the polymer is a polyanion. Secondly, it is observed in Figure 52.a that for α -hemolysin, the interaction is very long (recordings are shown over several seconds), we observe a long lasting blockade, which is experimentally quasi non-reversible, even at the lowest KCl concentration. In these experiments, such blockades usually happen, and sometimes they are due to the empty pore's gating, or a polymer strongly bound to the pore and one can reverse this interaction by reducing the applied voltage magnitude, going to zero voltage or in extreme cases reversing the potential polarity. Here, controls of the empty pore showed no such types of gating behaviour (at 300mM KCl hemolysin the pore does not spontaneously gate) and such effects were observed only after addition of the polymer into the solution. These long lasting blockades started to happen, and lasted for seconds to minutes and it was only possible to return to the open-pore current configuration by switching off the applied voltage or reversing the voltage polarity. Under these experimental conditions, it is very difficult to carry-out such measurement and much less to extract statistically relevant information about the interaction with the pore. As a comparison it has been seen that the interaction of polystyrenesulfonate (NaPSS) with α -hemolysin in 1M KCl, lasts for a few milliseconds (in the order of 1 to 10ms) depending on their average molecular weight and voltage applied [76], [148].



Figure 52: Interaction of two biological nanopores with the P1a sequence-encoded polymer in 300 mM KCI. Current traces are shown over 7 s, at +100 mV and -100 mV for (a) α -hemolysin and b) aerolysin. Zero current represented by a dashed line. The polymer was added to a final concentration of (a) 15 μ M and (b) 40 μ M. Insets show schematics of electrophysiological recording situations and a representation of the fully deprotonated P1a structure (top right).

In Figure 52.b interaction of P1a with aerolysin exhibits a more "traditional" interactiontype. Blockades which also happen only at *trans*-positive voltages, last for a few milliseconds and are reversible, they have a reasonable frequency at such a concentration and it was therefore possible to perform statistical analysis. For the reason that even at low salt concentration, the interaction with α -hemolysin was too long and difficult to analyse, we proceeded in using only aerolysin to characterize the rest of the polymers.

Figure 53, displays the detailed analysis of P1a interaction with aerolysin in different salt concentrations, at +120mV. The polymer was added in the solution at a concentration of 40 μ M and 20 μ M in 300mM and 1M KCl respectively. The events shown at 300mM for polymer P1a, were observed with two residual conductance distributions. It is possible to see, at first glance this two main levels in an all point analysis of 4s traces (Figure 53.a1). Whole trace analysis (Figure 53.a2) shows similar peaks, a main distribution I₁ and secondary distribution I₂ with peak values respectively at I/Io ~ 0.66 and 0.81. The deeper level "I₁" (with peak value at 0.66) appeared to be predominant. Scatter plots (in red Figure 53.a2) representing dwell time *vs* voltage show that the distributions of dwell times for the main deeper level I₁ are overall longer than the secondary level I₂. Analysis of the time distributions, showed mono-exponential distributions (as seen in Figure 58) and fittings of these distributions gave characteristic values of 27.92 ms and 0.48 ms for I₁ and I₂ respectively.

In 1M, a similar behavior was observed. The events were distributed in two main distributions, with values of residual conductance of main level I_1 at 0.65 and secondary level I_2 at 0.81 (Figure 53.b2). Similarly, the dwell times, longer for the main level, had characteristic values of 67.53 and 0.99 respectively. One can also notice that both these values of dwell time are longer in KCl 1M than in 300mM which shows that the salt promotes the interaction of this polymer with the pore.

The next experiments consisted in the study of the interaction of longer same-type polymers (comonomer (0)) with the aerolysin nanopore, observe how the length of the polymer promotes the interaction. This polymer called P1b contained 104 comonomers of type (0). As displayed on Figure 54.a it seems that the polymer is interacting in 300mM KCl in a similar way than P1a. The residual conductance histogram shows two preferred distribution (Figure 54.a2) with a main distribution I_1 at 0.69 and a secondary distribution I_2 at 0.82. Characteristic dwell times, distributed in mono-exponential fashion (see Figure 58), were respectively of 31.7 and 0.62 ms for

the main and secondary peak. Although the residual conductances here exhibit no significance differences than with the polymer P1a in the same experimental conditions, duration of blockades, appeared to be overall longer for P1b, which seems consistent with the fact that it is a longer chain.

P1b at 1M KCl, it showed interaction with the pore in these conditions, however as shown in Figure 54.b, this interaction showed similarities to the case of P1a α -hemolysin, especially at higher voltages (>120mV), although, to the contrary α -hemolysin, here shorter-lived events were possible to observe. Longer interactions (several seconds), were of deeper residual conductances than α -hemolysin, but showed a similar quasi-irreversible behaviour. This made the statistical analysis of the reversible events difficult and non efficient (it was diffucut to obtain enough events and have significant statistics).

The observation of the two distributions is surprising, giving the fact that we have only one polymer type in the solution, one possible explanation is that the polymer is entering in two different orientations in the pore. It has been shown in DNA interaction with the α -hemolysin pore that similarly two distributions of residual current could be observed depending on the 3', 5' orientation of the polymer at the pore's entrance [3], [48], [49]. In our case the polymer is assymetrical, as it contains a primer sequence on one end but not the other, it is possible that the polymer gives two different distributions of residual conductance with respect to either the 3T sequence or the phoshodiester facing the pore. However, to validate this hypothesis a number of factors need to be taken into account, we need to rule out the presence of impurities or aggregates and to consider the dimensions and conformation of the polymer in the solution (coil or stretched or in a good solvent), this is what we try to achieve by studying the polymers in solution using scattering techniques.



Figure 53: Interaction of P1a with the aerolysin pore at +120 mV in (a) 300 mM KCl and (b) 1 M KCl. The polymer was added to a final concentration of (a) 40 μ M and (b) 20 μ M. (a1,b1): Current traces shown over 4 s. Superimposed on the traces is an all point histogram (red line), the dashed black line represents zero current. Dashed red lines (I1, I2) represent, respectively, the main and secondary levels (see the main text). (a2,b2): Histogram of residual conductance (*II/I*0, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots). (a3,b3): Superposition of ~30 for I1 (black) and I2 events (red). Insets are the same as in Figure 52.



Figure 54: Interaction of P1b with the aerolysin pore at +120 mV in a) 300 mM KCl and b) 1 m KCl. The polymer was added to a final concentration of 13 μ M. (b). (a1,b): Current traces shown over 4 s. Superimposed on the traces is an all point histogram (red line), and the dashed black line represents zero current. Dashed red lines (I1, I2) represent, respectively, the main and secondary levels (see the main text). (a2): Histogram of residual conductance (*III*, 0, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots). (a3): Superposition of the onset of \approx 30 for I1 (black) and I2 events (red). Insets as the same as in the previous figures.

1.2. Scattering experiments and characterization : P1a

It is important to understand if this two current distributions are due to the same polymer or to eventual aggregates. For this reason, it is important to perform physical analysis of these molecules in solution in parallel to the nanopore experiments. Analysis of these polymers performed using ESI-MS with chemist's collaborators, coupled to DLS and X-Ray scattering performed at the institut Charles Sadron in Strasbourg, were performed to check for the purity of the samples, and to study their conformation in solution under similar conditions than the nanopore experiments, and ultimately correlate them to the obtained results with nanopores.

ESI-MS was performed on P1a, in order to check for the validity of the synthesized polymer, and showed pure samples [136], however this analysis was performed in different conditions than our experiments (no salt added, different pH, etc..), so one cannot directly correlate this result to our experiments.

DLS analysis was performed for P1a polymer in 300mM KCl, 10mM Tris at pH=7, in order to be directly correlated to our experiments. The polymer here was added at a concentration of 0.5% corresponding to 27µM, it is here in a dilute regime, so we can neglect intermolecular interactions and these conditions are similar the nanopore's experimental ones. Figure 55 shows the correlation functions g_1 and g_2 of the intensity and the electromagnetic field. The distributions of time (black) showed in Figure 55, shows a main distribution with a hydrodynamic radius $R_{\rm H} \approx$ 2.84 nm. The whole curve seems to have almost one main distribution. Peaks at bigger R_H values of 112 nm and 404 nm can also be observed, this can be due to aggregates of polymers or impurities in the solutions however as their amplitude is quite low (bigger polymer if present in high enough concentration in solution, scatter more light and their amplitude should be higher, which is not the case here), it suggests that they are in very low amounts in the solution. This implies that the sample is quite pure in this case, with what seems like one type of polymer, in addition all the repeated measurement on the sample gave little disrepancies over the hydrodynamic radius value. Other distributions (for smaller R_H than 2.84 nm) can be discarded as they are difficult to discern from density fluctuations of the solution (the solution had a low polymer concentration which causes such artifacts to happen more often).



Figure 55: Dynamic light scattering of P1a at a concentration of 27 μ M (0.5%) in 300mM KCl. Blue and red plots represent respectively the correlation functions of the intensity and electromagnetic field. Black plot is the distribution of relaxation times, which can be directly correlated to the hydrodynamic radius, as shown in the top axis. Amplitude of correlation, time and radius axes are in logarithmic scale.

X-ray analysis was performed over a long range for all the polymers, performing this analysis at small angle and wide angles (SAXS and WAXS) permits to obtain information on polymerization degree and Radius of gyration at small angles. For P1a, Figure 56 shows X-Ray diffraction for a solution at a low concentration of 0.47% i.e 26.4 μ M of P1a, this concentration which corresponds to a dilute regime, allows the use of the intensity function to retrieve the form factor as explained earlier, at dilute regimes, we can neglect the interaction between molecules. We have showed two different methods to fit this intensity using the Guinier and the Debye approximation. Values using both methods for all analyzed polymers are summarized in the following table.

	R_H (nm) - DLS	X - Ray	N	<i>Rg</i> (nm)
Homopolymer		Guinier	68.5	4.88
	2.84			
P1a		Debye	66.3	4.22
Homopolymer	27.12	Guinier	3926.7	21.1
	35.12			
P3		Debye	2554.2	14.0
Block Copolymer		Guinier	150.9	4.30
	2.84			
B1		Debye	141.8	3.68
Homopolymer		Guinier	38.0	4.25
	/			
B1		Debye	35.7	3.61

Table 5: Summary of the values of the hydrodynamic and gyration radius (R_H and R_g) of DLS and X-ray scattering measurements. X-Ray measurements were fitted with the Guinier and Debye approximation.

We can clearly see in Table 5 that the values are very close to each other for both equations (X-Ray), and that the fits for P1a work quite well in Figure 56 using either method. We can also deduce a ratio between hydrodynamic radius R_H and gyration radius R_g taking an average value of 45 Å for R_g , $R_H/R_g \approx 0.631$, which is close to the theoretical value. However, we notice that we have obtained a polymerization degree with an average of 67 which in this case is of the same order of magnitude than the value N of the awaited polymer but doesn't correspond to the theoretical value of N for the polymer measured which is supposed to have N = 59 (56 comonomers (0) + 3 thymidines). Calculation of the contrast K was done with the hypothesis that the polymer is completely deprotonated, this could not be true and give artifacts in the measurement of N. In order to obtain a more accurate value it is needed to know the exact counter-ion and the exact deprotonated fraction, we also neglect the thymidine sequence in the calculation, which might influence the result as well. However in first approximation here for the characterization of the polymer's dimensions, these simple assumptions are good enough.

The characterization results for this polymer P1a suggest, in our experimental conditions, that we have mainly one conformation of the polymer, and no aggregates of big size or high concentration that produce the events that we are observing. This in turn, points to the fact that the two residual current distributions are due to the same polymer and not to impurities. The fact that the polymer's dimensions are much bigger than the aerolysin's channel diameter (~1nm) suggests that the polymer cannot enter as a coil in the pore but more in a stretched configuration. As previously discussed, the most likely hypothesis is that the polymer is interacting with two different orientations, depending on the side it interacts with the pore, either with the 3T primer sequence or with the opposite end of the polymer. To validate this we have observed the interaction of only a polythymidine with 3 bases with the aerolysin nanopore.



Figure 56: X-Ray scattering measurements for P1a at a concentration of 27 μ M (0.5%) in KCI 300mM. Both curves are plots of the form factor $f_1(q)$ extracted from the normalized intensity. the top plot corresponds to $f_1(q) \alpha q^2$, used to fit the Debye equation, valid only for $qR_g < 3$. both axis are in a logarithmic scale. Bottom plot also called the Zimm representation is the inverse form factor $1/f_1(q) \alpha q^2$, used to fit the Guinier equation, valid for the range $qR_g < 1$.

1.3. Interaction of T3 with aerolysin

In order to furthermore check the orientation-dependent interaction hypothesis, we performed experiments to probe the interaction of oligonucleotides sequences, containing 3 thymidines with aerolysin. As one can see in Figure 57.b the amplitude of interaction is quite similar to the previously observed main level of interaction I_1 of P1a and P1b, at a value of

I/Io \approx 0.63, and seems to validate the previous hypothesis, however the dwell times, are much shorter (~100 µs). These observations suggest that this ssDNA sequence seems to be responsible for the main residual conductance value of the polymer, however the whole poly(phosphodiester)s, plays a role in promoting the interaction, but doesn't significantly change the conductance when it enters in this orientation.



Figure 57: Interaction of ssDNA 3T sequence with the aerolysin pore at +120 mV in 300 mM KCI. The polymer was added to a final concentration of 1 μ M. (a): Current traces shown over 4 s. The dashed black line represents zero current. The dashed red line (I₁) represent, the main level of interaction. (b): Histogram of residual conductance (*III*₀, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots).

1.4. Dwell time analysis

The obtained values of R_H and R_g (see Table 5) suggest that the polymer doesn't enter the pore in a coil state, as these dimensions are much greater than 1 nm. However it doesn't validate complete translocation of the polymers through the pore. Therefore, an analysis of the dwell-time,

was performed, as evolution of dwell time *vs* voltage can give information on the type of interaction, as we previously showed in chapter 4.

Figure 58, represents the evolution of dwell time for P1a and P1b over a voltage range of 70mV to 190mV. The dwell time were analyzed for the main and secondary distribution of events. First as displayed in Figure 58.a, the histograms which represent a distribution of the natural logarithm of characteristic times, showed a monoexponential distributions behavior for the characteristic times. As one can observe in Figure 58.b, the evolution of dwell time for the main distribution shows a biphasic distribution, with increase of sevenfolds for voltages between 70 mV and 150mV, followed by decrease for voltages above. The analysis of the secondary peak shows only the increase of dwell time *vs.* voltage.

The above results are suggesting a barrier for translocation, for the polymer when it enters with the T3 primer sequence, and that the decrease of dwell times might reflect translocation for higher voltages. It has been previously discussed [76] for polyelectrolytes, that the dwell time is inversely proportional to the voltage and suggests successful translocation. The magnitude of the electric force increases with voltage, in absence of any other force influence, the velocity of the charged polymer increases therefore reducing the characteristic dwell time. However if other factors are involved for example, energy barriers, like an entropic barrier related to the polymer's conformation or an enthalpic barrier, the evolution of dwell time can be increasing with the voltage. The fact that characteristic times decrease, thus a possible translocation, only for very high voltages above 150mV, suggests that the barrier of translocation for this polymer, is higher than observed for other polyelectrolytes, like for example NaPSS with α -hemolysin [76] or ssDNA (~50 bp) where the duration events with aerolysin was of the order of ~10ms decreases for voltages 100 mV and above, where no such biphasic behavior was observed. A possible explanation of this barrier is that the polymer is highly flexible in these experimental conditions, which increases the entropic barrier for translocation. In addition, monoexponential distributions usually are a reflection of first order binding reactions, in contrast, electrophoretically driven translocation is theoretically described to exhibit a more deterministic distribution [10]. In our case it can point to the fact also that the polymer is binding to the pore, without ruling out translocation at higher voltages.

The continuous increase of dwell times for the secondary peaks, suggesting no translocation for these types of events might be explained by the fact that the barrier of translocation for the polymer, from the other orientation is too high and doesn't allow for a translocation to happen for the range of voltages possible to apply.

These result suggest that in these experimental conditions, the polyelectrolyte is behaving as a chain in a good solvent, and is very flexible. Changing the solvents for the nanopore experiments in association with light scattering and X-Ray study in pure water and other solvents, might give insight into the best conditions for translocation experiments. Finding the optimum conditions is highly important for the nanopore experiments, for example if the polymer is needed to be in a rod-like conformation, the charges must be less screened. Mixture of organic solvents with water, might be used for example, to achieve this purpose, however one has to take into account factors of electrical conduction, and membrane stability.



Figure 58: Analysis of the dwell time of resistive pulses during interaction of poly(phosphodiester)s P1a and P1b with the AeL pore in 300 mM KCI. a) Histograms of the natural logarithm of dwell times of 11-events (left) and I2-events (right) for P1a (light blue) and P1b (dark blue) at 120 mV with superimposed single-exponential fits. b) Voltage dependence of characteristic dwell times for I1 events (left) and I2 events (right), for P1a (red) and P1b (black). Values are averages (error bars: standard deviation) obtained from at least two separate pores. Black and red curves represent Gaussian (left) and a linear (right) fits.

2. Poly(phosphodiester)s with comonomer (0) and the primer sequence 1-thymidine (P1c)

2.1. Interaction with aerolysin

We tested a similar polymer in the same conditions. The polymer P1c contains one thymidine as a primer sequence and 64 comonomers (0). Figure 59 shows interaction at 120mV. The first striking effect is that interaction is happening at a much lower frequency (even though the concentration is similar to P1a and P1b) and dwell times are much lower (~1 ms) in comparison to previously described polymers. Somehow, reducing the number of thymidine bases seems to reduce the strength of the interaction. To the contrary of the previously described experiments, we do not observe a main or secondary peak, instead a broad distribution of events, which suggests no preferred interaction. This differences in results validates the fact that the thymidine sequence plays an important role in the interaction of these polymers, and probably promotes it. However, one should expect to observe a similar distribution of events than the secondary distribution previously observed (interaction of comonomer (0) tail at nanopore entrance), which is not the case here, so even for this orientation the thymidine sequence seems to play a role in the interaction. A possible explanation considering the results with the P1c polymer is that the polymer's part with the (0) comonomers doesn't bind strongly with the pore and that the ssDNA sequence is mainly responsible of the binding interaction, whereas the rest of the chain (poly(phosphodiester)s), mostly promotes the duration of the interaction. This secondary peak of interaction might be not observed for ssDNA 3T sequence because of the fact that its length (\sim 3 x $0.33 \approx 0.99$ nm) is much smaller than the length of the channel (~9 nm) and the same order of magnitude than the pore's diameter (~1nm) which suggests that the polymer can possibly still turn in the pore and that the orientation doesn't matter in this case as suggested by simulations of [149]. However these hypotheses need to be checked by molecular simulations as this effect is only an indirect observation of the phenomenon and interpretation is merely a possible interpretation.



Figure 59: Interaction of P1c with the aerolysin pore at +120 mV in 300 mM KCl. The polymer was added to a final concentration of 100 μ M. (a) Current traces shown over 4 s. The dashed black line represents zero current. Red plot is an all point histogram over 4s. (b): Histogram of residual conductance (*II*/*I*₀, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots).

The next experiments performed were with similar polymers. For sequencing purposes these polymers need to contain some kind of infomation, therefore, to have other types of comoners, which would contain side chains. In order to investigate such effects it is important to observe the influence of the addition of side chains into the interaction with the protein pores.. We will present here results with three types of side chains (comonomers (0), (1) and (2)), bore by homopolymers, and blockcopolymer.

3. Poly(phosphodiester)s with comonomer (1) and the primer sequence 1-thymidines (P3 and B1)

3.1. Interaction with aerolysin : 300 mM KCl

We tested the homopolymer P3 and block copolymer B1, in the same conditions as previously described (KCl 300mM, Tris 10mM, pH 7.5). These polymer are derivatives of the P1c polymer, with only one thymidine, they contain 64 comonomer units of type (1) with side groups containing alkynes (CH₂-C \equiv CH). P3 is the homopolymer and B1 is the block copolymer of the same type, containing 4 alternating blocks of comomoners (0) and (1) respectively, with 16 monomers each. Figure 50 and Figure 60 shows the structure of these polymers and Figure 60 similar histograms than the previously studied polymers.

First we notice that the events are observed only at trans-positive voltages, the current trace here shows interaction at 120mV. The all point histograms of Figure 60 a1 & b1 already indicates the presence of specific levels of interaction. The analysis of the whole trace for one experiment, much like the previously analyzed P1a and P1b shows two main distributions of interaction, at $I/Io \approx 0.45 \& 0.7$, respectively with dwell times of 0.88 ms and 0.31 ms. In contrast to the other polymers, for P3, both these distributions are broader and the dominant distribution, is the one that is correlated to more residual current (I/Io ~ 0.7) (shallower blocks happen most often). Thereafter, the block-copolymer B1 shows similarly two main broad distributions I/Io \approx 0.5 & 0.78 with characteristic dwell times of 0.25 ms and 0.23 ms respectively, however neither of the distributions in this case is dominant. We remark that the values of residual current for P3 and B1 are close to each other, on average, magnitudes of blockades amplitudes, are shallower for the block copolymer B1 in comparison to P3 however the difference between these values is not significant.

The motivation for using this structure of block-copolymer, was for the attempt of sequencing simple enough structures. Indeed one might expect with this type of structure and the asymmetry of the polymer B1, blockade events which exhibits current steps within them, and depending on the orientation of the polymer in respect to the pore, would be mirrored in shape. However, we see no such patterns in these experiments, it seems, that the blocks don't interact long enough with the pore to give significant change to the shape of the current blockades signature and that the polymer participates as a whole structure to the depth of blockades, to the contrary of the expected single block interaction. It is possible that the events are too fast to

distinguish such patterns, and for each block of the polymer to produce a significant change in the current passing through the pore. It also seems that the homopolymer produces similar events, indicating that the difference between the two structures is not significant enough to exhibit differences in the current passing through the pore.



Figure 60: Interaction in 300 mM KCI with the aerolysin pore at +120 mV of (a) P3 and (b) B1. The polymer was added to a final concentration of (a) 450 μ M and (b) 130 μ M. (a1, b1): Current traces shown over 4 s. Superimposed on the traces is an all point histogram (red line), the dashed black line represents zero current. Dashed red lines (I1, I2) represent, respectively, the main and secondary levels (see the main text). (a2,b2): Histogram of residual conductance

(*I*/*I*0, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots). (a3,b3): Superposition of ~30 for I1 (black) and I2 events (red). Insets are the same as in Figure 52.

3.2. Interaction with aerolysin : 1M KCl

We performed the same experiments with polymers P3 and B1 in 1M KCl, (see Figure 61). Analysis shows again two dominant distributions. Overall, the polymers seem to block more current, however the previously observed distribution in 300mM KCl , seem broader. With the homopolymer P3, the pore started to gate at voltages higher than 100mV, therefore in 1M, we recorded traces at lower voltages of 60 and 80mV. The deeper distribution in 1M at 80mV has a peak value at I/Io ~ 0.2 with a characteristic dwell time of 6.13 ms. The shallower distribution which is broader has a most probable value at I/Io \approx 0.61 with a characteristic dwell time of 0.80 ms. An interesting observation is that in 1M, it is possible to observe subpeaks within the shallower distribution of P3 which might indicate some kind of polydispersity in the solution. The copolymer B1 also shows two main distributions, in this case both distributions are not as broad as for the homopolymer P3 and the deeper peak located at I/Io = 0.29 with a characteristic dwell time of 4.37 ms, corresponds to the most probable events, in addition it is possible to observe a distribution at I/Io \approx 0.43 of residual current with a characteristic dwell time of 0.91 ms. We can also see here that in comparison with the dwell times in 300mM KCl the characteristic times here are longer, showing again that the salt promotes the interaction.



Figure 61: Interaction in 1 KCl with the aerolysin pore at +120 mV of (a) P3 and (b) B1. The polymer was added to a final concentration of (a) 200 μ M and (b) 100 μ M. (a1,b1): Current traces shown over 4 s. Superimposed on the traces is an all point histogram (red line), the dashed black line represents zero current. Dashed red lines (I1, I2) represent, respectively, the main and secondary levels. (a2,b2): Histogram of residual conductance (*II*/0, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots).

3.3. Scattering experiments and characterization

The conformation and size of the polymers is also important to take into account for these polymers, and how they behave in such conditions, especially in these case where we observe multiple and broader distributions. For this matter, we also performed physical analysis on this polymer, again using DLS and X - Ray analysis.

DLS analysis of polymer P3 is shown in Figure 62. The first striking observation for the distribution of relaxation times, is that the dominant peak here has a most probable value at $R_H \approx 35.1$ nm, it is also possible to observe a distribution with a most probable value $R_H \approx 3.2$ nm. It appears that for this preparation and in these conditions (300mM KCl, Tris 10mM, pH=7.5) most of the observed objects are quite voluminous.

The SAXS analysis of this polymer, shown in Figure 63 using both Guinier and Debye equations, gives similar results than the dynamic light scattering. Indeed, as summarized in Table 5 the values of the gyration radius in this case is estimated to be 14 to 20 nm, depending on the

fitting method used, the polymerization number here is estimated with values between 2500 and 3900 monomers.



Figure 62: Dynamic light scattering of P3 at a concentration of 20 μ M (0.5%) in 300mM KCI. Blue and red plots represent respectively the correlation functions of the intensity and electromagnetic field. Black plot is the distribution of relaxation times, which can be directly correlated to the hydrodynamic radius, as shown in the top axis. Amplitude of correlation, time and radius axis are in logarithmic scale.

Results of DLS and X-Ray analysis clearly show first, that the polymer P3 seems to be mostly aggregated as the expected number of monomers of this polymer is 65, given the values of N_{W} , indicates aggregates of 40 to 60 chains. This fact is validated for the very high obtained values of hydrodynamic and gyration radius, which are respectively 12 times and 4-6 times more than the previously estimated values of hydrodynamic and gyration radius of P1a. However, it is still possible to observe a distribution of smaller objects, through the DLS analysis, as shown, with the distribution of radius 3.2 nm, this latter might represent single chains as this dimension seems to be coherent with such a hypothesis. This fact could explain, why even though the preparation is dominated by aggregates, we are still able to observe interaction with the nanopore.


Figure 63: X-Ray scattering measurements for P3 at a concentration of 40 μ M (1%) in KCI 300mM. Both curves are plots of the form factor $f_1(q)$ deduced from the normalized intensity. The top plot corresponds to $f_1(q) \alpha q^2$, used to fit the Debye equation, valid only for $qR_g < 3$. both axis are in a logarithmic scale. Bottom plot also called the Zimm representation is the inverse form factor $1/f_1(q) \alpha q^2$, used to fit the Guinier equation, valid for the range $qR_g < 1$. For low q values we are clearly not et the expected plateau regime, pointing to the fact that the polymer is not in a single chain conformation.

The analysis of polymer B1, shown in Figure 64 for DLS and Figure 65 for X-Ray, Suprisingly shows much smaller dimensions of hydrodynamic radius, the estimated value of 2.84 nm which is the same as the one of P1a. Values of X-Ray diffraction also show a radius of gyration with dimensions of 4.3 nm and 3.68 nm depending on the fitting method. These dimensions are a little smaller than for P1a here, however the estimated number of monomers for these measurements, is ~150 monomers, which is surprising considering the obtained dimensions, however, seems to be in line with the chain aggregation similarly than polymer P3, yet unlike P3 this aggregation is much weaker, as on average the most probable is that 2 chains are aggregated together.



Figure 64: Dynamic light scattering of B1 at a concentration of 20 μ M (0.5%) in 300mM KCI. Blue and red plots represent respectively the correlation functions of the intensity and electromagnetic field. Black plot is the distribution of relaxation times, which can be directly correlated to the hydrodynamic radius, as shown in the top axis. Amplitude of correlation, time and radius axis are in logarithmic scale.

This aggregation effect seems to be happening for polymers that bear side chains containing alkyne groups. These groups are known to be hydrophobic and this characteristic might lead to hydrophobic bonds between these groups in solution, which might explain the fact that the homopolymer shows a high aggregation degree. However, the fact that the copolymer seems to be much less aggregated is surprising as half of the polymer contains the alkyne groups, it could be that the negative charges repulsion between the phosphates is stronger than the hydrophobic effect, which might minimize the aggregation for B1. Nevertheless, it is clear that one has to carefully take into consideration all of these parameters to carry out nanopore experiments, and that the behavior of this polymers in solution is widely varying even within similar structures.



Figure 65: X-Ray scattering measurements for B1 at a concentration of 40 μ M (1%) in KCI 300mM. Both curves are plots of the form factor $f_1(q)$ deduced from the normalized intensity. the top plot corresponds to $f_1(q) \alpha q^2$, used to fit the Debye equation, valid only for $qR_g < 3$. both axis are in a logarithmic scale. Bottom plot also called the Zimm representation is the inverse form factor $1/f_1(q) \alpha q^2$, used to fit the Guinier equation, valid for the range $qR_g < 1$. There is a slight indent on the plots for very low values of q, which usually represent repulsion interactions.

It is however surprising that we are able to observe interaction and shallow enough events (low values of I/Io) knowing that these chains are forming quite big aggregates. This might be a reflection that even at low concentration it is still possible to observe single chains, but it becomes less clear, if the observed distributions correspond to an orientation phenomenon or to the fact that we are observing single but also multiple chains aggregation which are still possible to accommodate within the pore, especially considering that the observed distributions are broader than for P1a for example. The fact that the distributions of the homopolymer P3 get even broader

in 1M KCl and shows multiple peaks within the same distribution, can be once more related to the possible fact that multiple structures are formed in the solution. We are mainly observing the bigger aggregates through characterization however this doesn't mean that smaller structures which could be sized using nanopores, are not formed, as it is for example the case of polydisperse PEG [33], [35]. This shows the power of detection of biological channels to polymers that are present in solution even at low concentrations.

Other analysis techniques to observe the polydispersity of these polymers in solution, include the size exclusion chromatography (SEC), this method allows the separation of fractions having different molecular weights using a column, however, one has to be careful with polyelectrolytes which can interact with the column because of electrostatic effects.

It seems that it is also very important to carefully prepare the polymer solutions beforehand, in our case as these polymers were only dissolved in water and vortexed for some time, it might also be important to use other methods like heating, allowing some time for dissolving or using mixed solvents that might counter act hydrophobic effects and prove useful for such experiments to reduce aggregation.

4. Poly(phosphodiester)s with comonomer (2) and the primer sequence 1-thymidine (P4)

4.1. Interaction with aerolysin : 300mM KCl

The next polymer that was studied, has a similar backbone but side chains in this case contain triazole rings. These structures are expected to be bulkier than the previously described polymers.

This polymer is synthesized using the same phosphoramidite chemistry previously described however it is obtained through post-polymerization modification of the P3 polymer [150]. The interaction is happening at *trans*-positive voltages, coherent with negatively charged polymer. An all point histogram over 4 seconds of current recordings under 120mV voltage displayed on top of the current trace in Figure 66.a, shows a main distribution I₂ and a deeper events distribution I₁ of current. The all-trace analysis, in Figure 66.b shows also two main distributions I₁ and I₂, that have almost equal probability, at peaks of I/Io \approx 0.45, and 0.65 respectively, superimposed dwell times represented by the scatter plots are shorter than characteristic values of dwell times for P1a and P1b, similarly to P3 and B1.



Figure 66: Interaction of P4 with the aerolysin pore in 300mM KCI at +120 mV. The polymer was added to a final concentration of 120 μ M. (a): Current traces shown over 4 s. Superimposed on the traces is an all point histogram (red line), the dashed black line represents zero current. Dashed red lines (I1, I2) represent, respectively, the two main levels (see the main text). (b): Histogram of residual conductance (*II*/0, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots).

The fact that the dwell times are shorter for the P3, P4 and B1 polymers might be due to the fact that they interact less strongly with the nanopore, the ssDNA sequence however short, as previously suggested, promotes the interaction of P1a and P1b. It is also notable that the addition of side chains to the polymer P1c also seems to strengthen the interaction in comparison to the bear backbone.

4.2. Interaction with aerolysin : 1M KCl

The experiments in 1M, are displayed in Figure 67 which shows a histogram and scatter plot of all the trace of one experiment and the interaction of P4 at 1M with aerolysin. The interaction, which happens again at *trans*-positive voltages seems to be more complex in 1M. It is possible to observe different residual conductances levels at a much lower voltage than for 300mM. Here, the applied voltage was only 50mV, which was strong enough to observe interaction and more than 5 peaks within the histogram, the main peak of the histogram has a dwell time of 0.57 ms. The observed multi peaks clearly indicate that we are observing a polydisperse solution. It is quite possible that the sample is not monodisperse in this case, to the contrary of previous polymers. Another possibility is that the polymer forms superstructures, as the aromatic rings have an affinity to interact together [151]. It might be possible that interaction is happening within the chains, or between chains, but it is not clear which hypothesis is the most likely, as these samples lack quantitative analysis, showing again the importance of physical characterization.



Figure 67: Interaction in 1 KCI with the aerolysin pore at +120 mV of P4. The polymer was added to a final concentration of 60μ M. (a): Current traces shown over 4 s. Superimposed on the traces is an all point histogram (red line), the dashed black line represents zero current. (b): Histogram of residual conductance (*III*0, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots).

4.3. Scattering experiments and characterization

Similarly we have performed the analysis of the polymer P4 using scattering techniques. Only X-Ray analysis was successful for this polymer due to the fact that the sample seemed to be contaminated (results not shown), the analysis with DLS of the first sample of polymer P4 seem difficult to interpret. The DLS result gave surprising results, where multiple distributions, corresponding to various hydrodynamic radii were observed, which pointed to a contaminated sample. The estimated dimensions of this polymer with X-Ray extracted from fits shown in Figure 68 seem smaller than previously characterized polymers with $R_g \approx 3.61$ and 4.25 nm and N=38 and 35.7, this is surprising considering the fact it is supposed to be bulkier than the previously described polymers. However, ESI-MS analysis was not performed by the chemists for this sample of polymer P4 so there is no proof that the polymer synthesized was the one expected or that the sample was pure.

First the fact that the value of N is estimated to be low using X-ray scattering might be an indicator that polymer is not the one that is expected, the method of post-polimerization modification could be responsible for less predictable results, even so that variations might happen from reaction to other. For example if the modification results in a polymer with side chains that are not a triazoles but other groups, then the calculated contrast K is different than its real value, this could lead to an overestimation of the calculated contrast, such that we are normalizing the intensity over higher values, which in turn gives a wrong estimate after fitting of the final value of polymerization degree N. So it is not clear if our obtained results, really correspond to the type of polymer that we want to analyze, or another form of it, it seems here that it is especially critical to perform multiple characterizations for this polymer, that is post modified, and where it seems the resulting molecule is less well controlled.



Figure 68: X-Ray scattering measurements for P4 at a concentration of 30 μ M (1%) in KCI 300mM. Both curves are plots of the form factor $f_1(q)$ deduced from the normalized intensity. the top plot corresponds to $f_1(q) \alpha q^2$, used to fit the Debye equation, valid only for $qR_g < 3$. both axis are in a logarithmic scale. Bottom plot also called the Zimm representation is the inverse form factor $1/f_1(q) \alpha q^2$, used to fit the Guinier equation, valid for the range $qR_g < 1$. For low q values we don't observe the expected plateau regime, pointing to the fact that the polymer is not in a single chain conformation and there might be some residual interaction between chains.

5. Interaction with MspA

One of the last experiments we performed was with the *mycobacterium smegmatis* porin A (MspA). As explained in previous chapters, MspA is a nanopore that has a particular geometry, with a large vestibule (~48 A inner diameter at channel's entrance) and very narrow constriction site towards the exit of the pore (~10 A). This means that it can accommodate bigger molecules,

but also that the construction site can be helpful for sequencing smaller part of the polymer as it has been shown for DNA. Indeed, at the constriction site only 4 bases of DNA participate at in the current blockade[46]. The inconvenient with the use of MspA is that it is a little more difficult to handle experimentally which is why it was only tried last and only for a few polymers. The insertion of this pore in the lipid membrane happens via fusion of micelles to the artificial planar lipid bilayer, to the contrary of aerolysin or α -hemolysin which are added as water soluble monomers to the solution and self-assemble upon binding to the membrane. The MspA pore is also more subject to gating at higher voltages, whereas the two other pores (aerolysin and α hemolysin) can be stable even at very high applied voltages (> 150mV). But because of the interesting properties mentioned earlier above, this pore also shows high potential for sequencing purposes.

We probed the interaction of polymers, P1a, and P3 which are respectively the most simple version of the poly(phosphodiester)s and the corresponding polymer with side chains. Figure 69 represents the interaction of the polymers P1a and P3 with MspA in 300mM KCl, Tris 10mM and pH=7.5 (similar experimental conditions than the other nanopores). For both polymers we only performed single experiments, and the following results represent preliminary data for probing the behavior of this type of polymers with this particular pore.

Figure 69 displays the interaction of P1a at 150mV, this interaction shows similarities to what was observed in the earlier experiments with α -hemolysin, namely, long shallow events (~many seconds) accompanied by some deeper shorter visits within the event. These events much like with α -hemolysin lasted for many seconds to minutes, and the open-pore configuration was retrieved again only after reversal of the potential polarity. On the other hand the interaction with P3 at 120 mV showed reversible events, most of them very shallow, however many distributions can be observed in the histogram of the whole trace analysis (Figure 69.b2), and it is possible to observe many peaks within the shallowest distribution. This can be a reflection of the presence of multiple aggregates (as it has been shown through light scattering), with possibly preferred sizes of aggregate formation (because we can observe preferred distributions values of I/Io, cf. Figure 69.b2) and we observe indeed two main distributions, which seem to represent interaction of various sizes of the polymer. In addition it is possible to observe a distribution of deeper events which might represent the entry of multiple polymers together, or the detection of much bigger aggregates, because of the fact that MspA has a wide entrance, as explained earlier it can accommodate bigger objects.

It is however interesting to see that that the polymers, first are able to interact with MspA, and that second, the interaction with this pore, might give new information about the poly(phosphodiester)s and in this case it is possible to observe the polydispersity of the solution (aggregates of P3) at lower salt concentration than it is possible for aerolysin where only broad distributions (without sub-peaks) were detected, which might suggest that MspA is a more powerful tool for these type of experiments.



Figure 69: Interaction in 300mM KCI with the MspA pore of (a) P1a at +150mV and (b) P3 at +100mV. The polymer was added to a final concentration of (a) 20 μ M and (b) 400 μ M. (a1, b1): Current traces shown over 4 s. Superimposed on the traces is an all point histogram (red line), the dashed black line represents zero current. (a2,b2): Histogram of residual conductance (*III*0, black line) and scatter plot of dwell time (logarithmic axis) versus residual conductance (red dots).

IV. Summary

In this study we have seen that polyphophodiesters which are negatively charged polyelectrolytes interact more favourably with aerolysin in contrast to α -hemolysin. It is clear that the interaction with the pore of the polymers that have no side chains is still quite strong which seemed to be due to the added ssDNA sequence, and seemed to show an effect of orientation-dependent interaction. The highly fact that they are very flexible renders their translocation difficult, making the sequencing process hard to achieve. However, the rest of the analyzed polymers containing longer side groups, seemed to interact for shorter times, in comparison to the simple polyphosphodiesters as shown via dwell time analysis, pointing the possibility that they become stiffer with longer side chains which shortens the duration of the interaction but more data still needs to be acquired for validation.

It has also been shown that it is of high importance to perform physical characterization of these polymers in solution in parallel to other experiments. Understanding the conformation of the chains in solution is of key importance. The DLS and X-Ray scattering showed that some of the polymers studied were more inclined to aggregation to the contrary of others, which makes single chain pore-sequencing very improbable. Because of the difficult task of attributing the observed electrical signals to single chains or aggregates, it becomes not possible to separate aggregation from single chains and much less to distinguish complex patterns encoded within the polymer.

In general it also seemed that the attempt at sequencing these simple fairly short polymers, did not succeed, which might indicate that they might need to be longer, in order for the interaction with each block within the copolymers and the pore to be slowed down and allow their characterization.

More experiments need to be performed with the MspA nanopore, which might be a good candidate for a sequencing purpose as it contains a narrow constriction, important for the reading resolution of polymers, as it has been shown for DNA [45], [46]. Because interaction was observed with this pore and the poly(phosphodiester)s, an optimization of the whole system of polymer-pore is still needed: for example using a polymerase coupled with the nanopore reading as it has been done with DNA [94], would help stretch the polymer and make it less flexible inside the pore and also slow down the translocation. One could for example add a DNA primer sequence at the beginning of the polymer, which can be recognized by the polymerase and help the directed threading through the pore. Solvents, are also important to take into account for the

conformation of the chain especially for these purposes, and characterization coupled to pore experiments in various solvents, is necessary to have a complete description of the system.

Chapter 6: Dynamics of ssDNA interaction with aerolysin: polyadenines A3, A4 and A5

DNA is the major molecule for life; it contains all the genetic information necessary for the survival of the cells, and the perpetuation of each species. This molecule which is a polymer exists in different forms, of single and double stranded DNA, and one of the most interesting aspects of this polymer, is the physics behind its complex structure. In the cell the DNA doesn't exist in elongated conformation and has a length of a few centimeters, however it packs in a nanometer scale nucleus through a folding process around histones, making the chromatin structure, in order to fit in the cell nucleus [152], [153]. All these interesting aspects, have gave rise to a wide area of research on these subjects, due to the high complexity of the DNA molecule and all its derivatives as well as the mechanism related to it. One of the major challenges of DNA research is fast and cheap sequencing of DNA and genomes. Different methods have been used, and biological and solid state nanopores have rose into one of the most promising techniques for this purpose [4], [5], [71]. Understanding the translocation mechanism of RNA through the nuclear pore [57], [58] or the virus' DNA packaging and ejection through its capsid [62], [63] is also of major importance. Another field of study focuses on DNA and also RNA, binding mechanisms to proteins, like the polymerase for example and is also a wide research area [154], [155].

Since the first experiments of detection of DNA molecules, with the α -hemolysin nanopore [3], the number of experiments has increased very rapidly. And as explained in the first chapter of this thesis, various studies ranging from simple detection, complex dynamics of DNA molecules in confined channels or sequencing purposes, saw an increasing interest [156].

In the following chapter we will present the dynamics of interaction of oligonucleotides with the aerolysin nanopore. Characterizing the statistics related to the interaction of small adenines with sizes, A3, A4 and A5, gave rise to events, which have specific and complex patterns, which we will try to analyze, interpret and quantify. We will present throughout this chapter the findings related to the experimental analysis of this study, and all the quantifications, probabilities and dynamics of DNA-induced events and we will relate the experimental findings to a theoretical model which aims at describing these results.

I. The ssDNA molecule

Single stranded DNA is highly flexible, it has been estimated by fluorescence spectroscopy to have a persistence length decreasing with the concentration of salt [157], [158] reaching 1.5 nm in 2M NaCl [158] suggesting an even smaller value for higher concentrations. This makes the DNA much more flexible than its double stranded counterpart which in contrast to the double stranded DNA, is more rigid, and has a persistence length of ~50 nm. This means that it can still be flexible and bend inside nanometric pores as it has also been shown by refs [131], [149]. In most research with nanopore, it is the form of the DNA molecule which is most used for characterization, because of the fact that it can generally be accommodated more easily inside the nanopore (in comparison to ssDNA, which has a bigger diameter).

II. Results

1. Pattern of events in 4M KCl

We performed the experiments, in a 4M KCl +10mM Tris at pH7.5. The interaction of DNA with aerolysin has been less investigated than with other pores, to our knowledge, only two other groups have worked on this type of experiments, and their work was published in refs [88], [115], [36], [47].

In ref. [36] Cao and collaborators observed the interaction of monodisperse solutions of polyadenines from 3 to 10 bases with the aerolysin nanopore in 1M KCl. They showed that the residual current was discrete and dependent on the size of the molecules, proving the ability of aerolysin in discrimination of very short DNA sequences. They later showed the fact that it is possible to perform the same kind of experiments, with the other bases types (CTG) [47].

In the light of these studies we reproduced similar experiments, performed in the higher salt concentration of 4M KCl. We demonstrate the ability of discrimination of these molecules by aerolysin in higher salt concentration which is shown in Figure 70, and proves again the fact that these molecules are blocking at discrete discernable values of I/Io for a given voltage. In 4M KCl at 120mV, A3, has a value of residual conductance of I/Io \approx 0.63, A4 has a value of I/Io \approx 0.43, and A5 I/Io \approx 0.27. The adenine-induced events show an increase of current blocked respective to the size of the molecule which seems to be consistent with the molecule's length increase. These values are systematic and reproducible for all the experiments, which gives a current-block

signature for each molecule similarly to what has been shown in 1M KCl [36]. This fact is important because it allows studying individual dynamics of the ssDNA molecules and separate them related to their residual conductance values.



Figure 70: Discrimination of adenine oligonucleotides by aerolysin in 4M KCI. Normalized histograms for the detection of A3, A4 and A5. Molecules were added to a final concentration of ~0.5 μ M.

In addition to these results, we observed additional features in the interaction of these molecules with the aerolysin pore, and the most striking effect was the complexity of the pattern of the A_n -induced events. As displayed in Figure 71, which shows as an example a library of events for A3, the event contains mainly a most probable level (labeled in red in Figure 71); this is the level that is detected in Figure 70 and is the level that allows discriminating between the various sizes of adenine molecules. In addition, each event seems to contain other sublevels, which are shorter and deeper level blocks of current, more specifically, one event contains:

- 1) A quasi-systematic deeper sub-level at its beginning
- 2) A deeper sub-level at its end
- 3) Brief visits to this deep level during the course of the A_n -induced events.

A3 40mV

120 mV



Figure 71: Library of A3-induced events. a) 40mV, rows respectively from top to bottom: deep-level-only events, size-dependent events with no pre or post blocks, size dependent with pre and post blocks and in-between-deep events, size-dependent event of longer duration similar to previously described. b) 120mV, rows respectively from top to bottom: size-dependent events with no pre or post blocks, size-dependent events with pre blocks, post blocks, size-dependent events with pre blocks, post blocks, size-dependent events of short duration with only pre blocks, similar events of longer duration. Dashed black line is zero current, red dashed-line represent the main level value

Through first qualitative observations, it seemed that the presence of the first deep-induced block, which we will call "pre-block", is quasi systematic. In contrast, the presence of the last deep-induced block, which will be named "post-block", seemed to be dependent on the magnitude of the applied voltage; the higher the magnitude the less probable it was to observe this last induced block.

Another class of events that were observed were deep-only induced levels, which are of short duration in comparison to the previous class of described events and show more blocked current (close to 0). These events seemed to be dependent on the voltage as well, and to happen most often for lower voltages.

Following these first qualitative observations we attempted at a quantification of these blocks, namely the frequency of these quantities, their duration and their dependency as a function of the voltage.

2. Analysis of the A_n-induced events

We analyzed the fraction of deep-only level events and of pre-block and post blocks and transition blocks (or returns) and worked-out their dependency on the voltage. The analysis procedure for pre- and post-block is detailed in Figure 73.

2.1. Deep-only-level events

The deep-only level events were quantified for each pore through their proportion, through the division of their count by the total count of events for each voltage. These events which have a similar amplitude than the pre and post-blocks, seemed to have similar durations too. We also notice that the proportion of the deep-only level events evolve in relationship to the magnitude of the voltage.

These events, as shown in Figure 71 represent a non-negligible proportion of the total number of these DNA-induced blocks. The analysis of their probability (fraction) as function of the voltage which is shown in Figure 72 displays a decreasing effect with the applied voltage. For higher voltages their probability decreases to reach a steady value of ~ 0.2 (20% of the total observed events belong to this category).

The next class of events analyzed is the one which contains the pre- and post-blocks which were described previously and which are quantified in the following section.



Figure 72: Probability (fraction) of events with deep-only level *vs.* voltage. Caption in figure shows respectively the plots belonging to A3, A4 and A5. In all cases the fraction reaches a steady value of 0.2 for higher voltages. The fits correspond to equation (65) (f_1), (see model in text for further explanation). Errors bars are standard deviation.

2.2. Pre- and post-block

For the analysis of the fraction of pre- and post-block, we have proceeded as follows: We extracted the value of the first and last level value of each event and quantified it in histograms for each applied voltage. This gave the proportions of events that started with a pre-block and post-block in respect to total events including the ones starting or ending with other level values. As observed for this analysis (see Figure 73) there are mainly two distributions possible: a distribution corresponding to the main level value and another broader distribution corresponding to the main level value and either start with the deeper level or with the main level, and similarly it can end with the main level or with a deeper level.

It is possible to quantify the total pre- or post-block fraction by a count which is divided by the total number of events with a main level, or through an integration method. The different steps of quantifying the probability of finding a pre- or post-block with the integration method, for a given adenine at a given voltage, are enumerated below:

 First, calculating the whole surface of the histogram by integrating the curve in order to quantify the total number of events.

- 2) Setting the maximum value of the integral corresponding to all values of I/Io inferior to a value which corresponds to the peak value of main distribution of the adenine analyzed plus 3 standard-deviations of the main-level distribution. This corresponds to a value a little above the last point of the distribution of the main level (red arrow in Figure 73). This step allows taking into the count all the events that correspond to the specific molecule and discard all other events (all shallower current-blocks).
- 3) Normalizing the integral curve by this maximum.
- 4) Selecting for the pre- or post-block fraction the value of the normalized integral corresponding to all the events belonging to the deeper distribution of the histogram (black arrow in Figure 73), which finally gives the probability of pre- or post-blocks. The procedure is applied separately for the quantification of pre- and post-blocks.



Figure 73: Method of analysis for determination of the probability of pre-(left) and post-block (right) for two voltages. First level (left) and last level (right): histogram (black) of the event and corresponding normalized histogram-integral (red). Each peak in the histogram is linked to the corresponding level in the example top events. Red arrows correspond to all the events taken for the analysis, black arrows represent the value of pre-block and post-block fraction extracted from the normalized integral and used in Figure 74. a) 30mV. Fraction of pre- and post-block is similar. b) 120mV. Fraction of post-block is close to zero.

The final obtained curve (see Figure 74.a) shows for all the adenines, first, for voltages from 10 mV to 30 mV, an increasing value for the pre-block, varying from 0.6 to 0.8 of the total number of events followed by a constant value for voltages above 30 mV to 120 mV which can be fitted at ~0.85. The post-block shows a sigmoid-type of transition, it shows first a similar probability than the pre-block for voltages that are inferior to 30mV reaches a steady value of 0.8, then it is followed by a decreasing probability, as the voltage increases, reaching an almost zero probability of finding a post-block for voltages above 100mV. These effects are observed for all the studied oligonucleotides. Nevertheless, the curves show some discrepancy between the different molecules especially for the post-block which displays a sharper transition of the from 0.8 to lower values for A4 and A5 and a slower decrease for A3. This shows a dependence of the probability of post-block on the size of the molecule. We can also notice the fact that the A4 and A5 curves are quite close to each other in comparison to A3.

2.3. Transition blocks

A third class of observed blocks correspond to deep brief visits to the same deep level than the pre- or post-block during the course of the A_n -induced events, that we will also call transition blocks or returns.

A return per-event corresponds to at least one-deep level visit or more after the pre-block. We divide then the total counts of returns over the total count of events which contain a main level. These can also be quantified using the same integration method mentioned in section 2.2.

Figure 74.c displays the average of this fraction of returns over the pre-block (normalization explained below) and similarly to the post-block analysis described in 2.2 we observe a transition starting from higher values decreases, with a similar transition to reach almost 0 probability. Here, the difference with the post-block resides in the slower transitions (less steep) of return-curve in comparison.

2.4. ssDNA and binding sites hypothesis

These observations and quantifications led to the following proposed hypothesis: the DNA molecule is passing in the pore by multiple binding sites, producing the various blocks that are observed.

1) It seems that the first block is a weak binding at a first site where most of the ionic flow is blocked, pointing to the fact that there is possibly a geometrical constraint as well.

- The molecule binds to a second site which is responsible for the main level value; this second binding seems to be very strong and long lasting.
- 3) There is a non-zero probability for the molecule to go back to the first binding site, which could be responsible for the presence of other sub-deep-levels, the fact that a final post block is present depends on the stength of voltage applied, as it is seen in Figure 74.

These multiple visits during one event to the deep blocks, might point to the fact that the molecule is oscillating between two binding sites, and has the possibility to exit the pore from the *cis* side or to move forward by translocating through the entire pore to the *trans* side. This claim is supported by the fact that the last level of these events is either the main level (corresponding to the second binding site) or a deep level block, which sees its probability decrease with increasing voltage. A second point that corroborates this hypothesis is that we also observe short lived only-deep-level blocks with a higher probability of happening for lower voltages, and which points to the probability that the molecule is binding to the first site but is not moving forward inside the channel but exiting the pore from the *cis* again. In all case the electrical field seems to be a major factor in this dynamics, and that translocation is helped by the electrical force. This whole process is summarized in the schematics of Figure 75

Following this hypothesis, one would expect the pre-block probability to be close to 1 for the analyzed main events namely the events which contain the main level, as the molecule should always pass by the first binding site and consequently produce a systematic pre-block. In addition, the observation of the post-block, which has similar probabilities than the pre-block for voltages lower than 30 mV, points to the fact that some of the pre- and post-blocks are missed during the recording or after detection due to the resolution of the experiment. This effect is more prominent for very low voltages (10 mV to 30mV) where the probabilities are close to 0.6 (instead the 0.8 steady value for higher voltages), and for which it is known that the signal to noise ratio is much lower.

Consequently, we have normalized the post-block probability by the pre-block probability in order to obtain what one would expect as the actual probability of post-block in the system. Indeed, if we consider that we make the same error for the detection of the pre-block than the post-block and that the pre-block is close to a probability of 1 for all voltages, then by this normalization we reduce the error due to the detection, and this should give a probability of post-block close to the real one. This normalized probability is shown in Figure 74.b, and one can

clearly see that the normalized probability of post-block is showing a transition from probability \sim 1 to zero as expected, and seems to validate the emitted hypothesis (namely the error due to the detection). We use the same reasoning of fraction of events missing due to the detection's resolution for the faction of returns (Figure 74.c), and we normalize it by the fraction of pre-blocks and perform analysis on the final curve.



Figure 74: Pre- post-block and returns probability *vs.* voltage for A3 (blue), A4 (red) and A5 (black). a) Pre- (triangles) and post-block (squares) average over at least three experiments, lines are guide to the eye. b) Relative probability of post-block (post block probability over pre-block probability, see text for further explanation), representing the actual probability of post block. Error bars are standard deviation. c) Relative probability of returns (returns probability over pre-block probability, see text for further explanation), representing the actual probability, see text for further explanation), representing the actual probability of returns. Lines correspond to fits of equation (69) (f_2). Bars are standard deviation.



Figure 75: Schematics of the states of the nucleotide in aerolysin. 1st row: Deep only level: the nucleotide binds to the first site and exits the pore from the *cis* side. 2nd row: The nucleotide binds to the first site, the second site and back to the first site then exits the pore from the *cis* side. 3rd row: The nucleotide binds to the first and second site and is able to translocate through the pore, exits to the *trans* side. Each state is attached to the corresponding event pattern (left).

We were next interested in the dwell times of these molecules in the aerolysin pore, which also gives information on the dynamics and is described in the following section.

2.5. Dwell time analysis

We have analyzed the dwell times of these molecules as a function of the voltage. Figure 76 shows the dwell time of the second binding site (m_2) (see model in section 3 for further explanation) and the total dwell time of the whole events. The evolution of this parameter in function of the voltage shows a biphasic dependence for both types of analysis, with an exponential increase and decrease of the dwell time for all the molecules. The dwell time of the whole event has a bigger value than the m_2 -dwell time which is logical. The position of the maximum increases with the molecular size and scales for m_2 from 1.18 ± 0.03 ms, 15.14 ± 0.38 ms to 37.64 ± 1.65 ms and for the total duration 2.02 ± 0.15 ms 21.27 ± 1.68 ms 61.29 ± 8.13 ms, for A3, A4 and A5 respectively. It increases for m_2 from A3 to A4 for about 13 fold and from A4 to A5 for about 2.5 fold. For the total duration it increases from A3 to A4 for about 10.5 fold and from A4 to A5 for about 3 fold. This increase is more pronounced from A3 to A4-A5 than

between A4 and A5 which shows again that the dynamics of A4 and A5 are closer to each other than that of A3 similarly to what was observed for the post-block evolution. However this whole analysis shows a reproducibility of the behavior of the dwell time for all the adenine molecules and demonstrates that there is a specificity of interaction of the adenine single stranded DNA with aerolysin.

The observations of this biphasic dependence seem to point to the fact that there are two competing effects, due to a probable high energy barrier preventing the molecules from translocation at lower voltages. It seems that the voltage is high enough to drive the oligonucleotides into the pore and overcomes the first entropic barrier of confinement, driving it into the first binding site, the second barrier moving the molecule to the secondary binding site also seems to be overcome at lower voltages, however the barrier for translocation seems to be proportionally very high, this is supported by the evolution of the dwell time *vs* voltage. The first stage consists in the increase of dwell time where the molecule is trying to overcome the barrier but mostly fails, and the second stage consists in the success of the translocation and therefore the decrease of the time that the molecule spends within the pore.



Figure 76: Dwell time *vs* voltage. m2 (second binding site, see text) dwell time (circles) and total event dwell time (triangles) for A3 (blue), A4 (red) and A5 (black). Dwell time axis is in a logarithmic scale. Fits of m2 dwell time correspond to equation (72) (τ_2). Error bars are standard deviation.

Following the quantification of all the previously described effects, we developed a model which attempts at describing the whole dynamic of the adenine nucleotide within aerolysin through a binding-unbinding reaction, ruled by the presence of multiple free energy barriers. The model is described below.

3. Model for the description of the dynamics of A₃₋₅ with aerolysin

Following these observations, in collaboration with Murugappan Muthukumar we postulate a minimal three state kinetic model based on the previous results namely:

- 1) The decrease with voltage of the probability of deep-level-only
- 2) The decrease of probability of returns (deep level visits during the course of the event) with voltage
- 3) The biphasic behavior of the dwell time

We try to describe more rigorously these dynamics using this physical model that aims at fitting the previously presented results. This model takes into account all the previously described effects through specific energy barriers for the adenine inside aerolysin. We assume in this model:

- Two binding sites: a first binding site near the *cis* entrance, where the molecule binds and blocks most of the current (deep-block-only events, pre and post blocks), and a second binding site where the particle binds for longer durations, and where the DNA blocks the current only partly; based on the molecular structure of aerolysin of [84] and the fact that DNA is negatively charged, it is most likely that the first binding site might be situated at the ARG220 ring (paired with the glutamate GLU221 ring) and the second binding site to be at the unpaired positively charged ring of LYS238 (cf chapter 2 for the aerolysin structure), similar suggestions for the binding sites have been made in ref.[47].
- Three free energy barriers: the first barrier is of entropic nature due to confinement and allows binding in the first site, the second barrier needs to be overcome in order to move to the second binding site and the third barrier is for the translocation and complete exit of the molecule from the pore.
- The particle is able to move in either direction of the pore: forward and backward.

As mentionned before, the presence of the electrical field can lower the height of these energy barriers because it confers an additional energy to the charged DNA molecule. This can be calculated by deriving the value of the electrical energy of the DNA particle due to the applied electrical field.

The electrostatic energy of charged particle of charge q is written as

$$U = qV \tag{53}$$

where q can also be written as q = |ze| and V is the electrical potential of the particle.

Considering the electrical field *E* as constant through the pore (constant potential difference *V* applied across the pore) and therefore equal to V/L we can write the energy of a particle at a position *x* in the pore as

$$U = qV(x) = \frac{zexV}{L}$$
(54)

We assume based on the obtained experimental results that the particle has to overcome three energy barriers in order to completely translocate through the pore, and that it is possible for it to go in either directions inside the pore. These energetic barriers exist in the absence of the electrical field, and are lowered by the fact that the particle acquires an additional energy (of electrical nature). All the above features of the model are summarized in Figure 77 and Figure 78 and reaction equation (61).



Figure 77: Cartoon and schematics of the binding sites of adenine in aerolysin and the differents states of binding reaction of equation (61) and corresponding length scale.



Figure 78: Schematics of energy landscape of adenine along aerolysin pore with and without electrical field; black curve is original three state energy landscape (three barriers). Red curve is landscape in presence of electrical field, corresponding to original landscape lowered by $-\frac{qVx}{L}$. Distances and energies corresponding to first (m_1) and second binding (m_2) sites are shown in the scheme. Arrows correspond to barrier heights in absence (black) and presence (red) of electrical field.

The energy barriers can be as written as a function of the original barriers F_1^0 , F_2^0 , F_3^0 (in the absence of electrical field), which are shown in Figure 78, and gives the following equations of free energy for the forward direction:

$$F_1^{\dagger} = F_1^0 - \frac{qVx_1}{2L}$$
(55)

$$F_2^{\dagger} = F_2^0 - \frac{qV(x_2 - x_1)}{2L}$$
(56)

$$F_3^{\dagger} = F_3^0 - \frac{qV(L - x_2)}{2L}$$
(57)

Where x_1 , x_2 are respectively the position of the first and second binding site and *L* is total length of the channel, F_1^0 , F_2^0 , F_3^0 are respectively the original energy barriers.

And the corresponding backward energy barriers are derived as below:

$$F_{-1}^{\dagger} = F_1^0 + \varepsilon_1 + \frac{qVx_1}{2L}$$
(58)

$$F_{-2}^{\dagger} = F_2^0 + \varepsilon_2 - \varepsilon_1 + \frac{qV(x_2 - x_1)}{2L}$$
(59)

$$F_{-3}^{\dagger} = F_3^0 - \varepsilon_2 + \frac{qV(L - x_2)}{2L}$$
(60)

Where ε_1 and ε_2 are the absolute energy values of the first and second binding sites in absence of electrical field. We write the kinetics of binding/unbinding reactions which describe Figure 77 as below

Where α_j and β_j are the reaction rates of forward and backward reactions respectively. *i*, *f*, *m*₁, and *m*₂ are respectively the initial state (before entering the pore), final state (complete exit from the pore), first and second binding site.

The rate constants are dependent on the free energy barriers and are described by the following equations:

$$\alpha_j = k_o e^{-\frac{F_j^{\dagger}}{k_B T}} \qquad (j = 1, 2, 3)$$
(62)

$$\beta_j = k_o e^{-\frac{F_{-j}^{\dagger}}{k_B T}} \qquad (j = 1, 2, 3)$$
(63)

Where k_BT is the Boltzmann constant time the absolute temperature and k_o the collision parameter, and is assumed to be constant for all the reactions.

The fraction of events that are bound to m_1 and return to the initial state *i* without ever binding to m_2 are described by the following equation:

$$f_1 = \frac{\beta_1}{\alpha_2 + \beta_1} \tag{64}$$

A mathematical development of this equation using the expressions of α_2 and β_1 from equations (62) and (63) yields the following equation:

$$f_1 = \frac{1}{1 + e^{-\eta(V_0 - V)}} \tag{65}$$

Where

$$\eta = \frac{zex_2}{2k_BTL} \qquad \text{and} \qquad V_o = \frac{F_2^0 - F_1^0 - \varepsilon_1}{k_B T \eta} \tag{66}$$

There appears to be a characteristic voltage V_o , which describes the approximate voltage above which most of the molecules transition to go to the second binding site. V_o is controlled by the difference in first and second energy heights, the binding energy of the first binding site which is close to *cis* and by the second binding site position. This equation needs to be rewritten to account for the previous described effects of background, namely that the fraction of events that return to the *cis* side without ever binding to m_2 reaches the steady value of 0.2 which gives:

$$f_1 = 0.2 + \frac{0.8}{1 + e^{-\eta(V_0 - V)}} \tag{67}$$

The probability of returning to the first binding site as opposed to exiting the pore is described by the following equation, this is the fraction of returns:

$$f_2 = \frac{\beta_2}{\alpha_3 + \beta_2} \tag{68}$$

Substituting with the expressions of α_3 and β_2 from equations (62) and (63) we obtain:

$$f_2 = \frac{1}{1 + e^{-\eta'(v_0' - V)}} \tag{69}$$

Where

$$\eta' = \frac{ze}{2k_BTL}(L - x_1)$$
 and $V'_o = \frac{(F_3^0 - F_2^0) - (\varepsilon_2 - \varepsilon_1)}{k_BT\eta'}$ (70)

Again there is a characteristic voltage V'_o , which represents the transition voltage above which most of the molecules don't return to the first binding site and it depends on the difference of the two last energy barrier heights, on the difference between the two binding energies and on the difference between the total length of the pore and the position of the first binding site. And in this case there is an additional constraint on η' which needs to be smaller than η because x_1 is very close to the *cis* entrance.

And finally the dwell time at the second binding site is described by the following equation:

$$\tau_2 = \frac{1}{\alpha_3 + \beta_2} \tag{71}$$

And again after substitution of α_3 and β_2 becomes:

$$\tau_2 = \frac{1}{k_o} \frac{1}{e^{-\eta''(v_o''+V)}} \frac{1}{1 + e^{-\eta'(v_o'-V)}}$$
(72)

Where

$$\eta'' = \frac{ze}{2k_BTL}(x_2 - x_1)$$
 and $V_0'' = \frac{F_2^0 + (\varepsilon_2 - \varepsilon_1)}{k_BT\eta''}$ (73)

This last equation depends on the previously described parameters V'_o and η' of the fraction f_2 , and k_o the collision parameter for the fitting of the dwell times results (k_o is fixed arbitrarily for all molecules and kept the same). We fix V'_o and η' after extraction from the previous fitting of the returns curve with equation f_2 . τ_2 depends on two additional parameters V''_o and η'' . V''_o depends on the value of the second barrier height, on the difference of binding energies and on the first and second binding sites positions, V''_o describes the threshold voltage above which the molecule is able to overcome the second energy barrier and to bind to the second binding site once it is bound to the first binding site, however it is still possible for some of the molecules to return to the initial

state (*cis*) after binding to the first binding site, it is a value which describe the possibility for the molecule to overcome the barrier but where the backward barrier for return to the *cis* compartment is still low enough that it is possible to overcome it as well. τ_2 describes the competition between two effects namely: the force that pushes the molecule for translocation and the the opposition due to the barrier heights for its translocation.

4. Extraction of fit values from experimental results

We have extracted fit values from the previously described equation. The equation used to fit the data, were the fraction of deep only level, returns and dwell time at the second binding site m_2 corresponding respectively to equations (65), (69) and (72). The results of the fits are summarized in the following table:

Fit parameters	A3	A4	A5
η	0.07	0.06	0.07
Vo	48.88	54.35	59.06
η'	0.11	0.13	0.12
V _o '	67.64	48.25	55.71
η"	0.075	0.09	0.06
<i>V</i> _o "	15.70	43.24	86.74
k _o	100	100	100

Table 6: Fit parameters of equations (65), (69) and (72) corresponding to deep-only-levels, returns and dwell times.

This model seems to correctly fit the above results as shown in Figure 72, Figure 74 and Figure 76. The fit of the first figure yields a voltage values increasing with the size of the molecule, and points to the fact that this value correspond to a threshold value increasing with the size of the molecule, and thus this transition value seems to not be dependent on the charge (otherwise V_o would decrease) but dependent on a geometrical factor, therefore dependent on the

confinement of the molecule, thus validates that the first fraction is mostly dependent on entropy. This can also be related to the fact that V_o is partly dependent on F_o^1 which represents the barrier for confinement. Another effect is that the transitions of f_1 reach a steady value and that these events can always be observed even at high voltages, which suggests that the first energy barrier of confinement can never fully be overcome leading to collisions of the molecules with the pore without ever going past the first binding site.

The fit of the second fraction f_2 which describes the returns to the first binding site, shows good fits for A3 and A4, and is in good accordance with the experimental results. However for A5 the position of the curve doesn't seem to fall into the expected evolution of the whole phenomenon, as it is positioned between the A3 and A4 plots. This could be due to the fact that the noise of events for A5 increases greatly in comparison to the two other molecules, and prevents us from detecting all the returns.

The fit of the dwell time also shows to be in good accordance with the experimental results for the global fitting of the biphasic observed behavior of the dwell time evolution. As explained previously we fix the η' , V'_o from the previous fit of f_2 and also fix k_o for all the molecules and extract the rest of the parameter, namely η'' and V_o'' . In this case though, the fit seems to fail for the A5 molecule which may be due to the fact that the detection of returns is of lower quality as explained earlier.

Despite some limitations, this model yields a good estimate of the rest of the fit values, and describes well the different experimental observed phenomenons. It is even able to describe the biphasic behavior of dwell time which to our knowledge was never explained theoretically.

III. Summary

Through this study we reported for the first time a particular dynamic for the adenine nucleotides with aerolysin in high salt concentration. We have studied the complex patterns of events obtained through the electrical recordings and tried to describe quantitatively the different blocks that were produced by the adenine nucleotides when interacting with the aerolysin nucleotides. We developed a minimalistic three state model which took into account all our observations and quantifications based on the hypothesis that the molecule has different states in this dynamic and 2 different binding sites within the aerolysin, and is able to move in either direction inside the pore. This model was able to describe quite well the different effects, namely the decrease of only-deep-level events, the decrease of returns to the first binding site and the dwell-time evolution with voltage; however we see some limitations to this model including for the dwell times of A5, probably due to a default in detection of the returns for that molecule.

These results give more insight into the dynamics of small nucleotides inside pores in general, and help in the understanding of the dynamics of bigger nucleotides of the same type (A6-A10) which will be presented in the following and final experimental chapter.

Chapter 7: Orientation dependent interaction of ssDNA with aerolysin: A6-10

In the previous chapter we have described the interaction of adenine nucleotides of very small size (A3-5) with the aerolysin nanopore, exhibiting a specific dynamic of interaction with this pore. In this chapter, we complete the study by the observation of the interaction of bigger adenines nucleotides namely, and try to probe the interactions of each of them and observe the differences in comparison to what has been observed in chapter 6.

In this chapter, the interaction mode of five adenine nucleotides A6, A7, A8, A9, A10 was probed. In the following discussion, the interaction mode of these nucleotides is sequentially explored based on current traces extracted from nucleotide-aerolysin interactions. This analysis yielded a particular dynamic and behavior for these polyadenines, which will be explained below.

I. Results

1. Dynamics of poly-A6

Figure 79 shows 250 ms of the current trace of the interaction of a monodisperse polynucleotide solution of A6 (6- adenine oligonucleotides) with aerolysin. Observation of Figure 79 give a first insight into the trace, which shows that the detection (obtained with the DetectEvent software) seems to be pointing at 120mV to one level of events, on average (labeled I_1 , blue dashed line, Figure 79).

In order to investigate this even more we observe Figure 80 which displays a complete analysis of these events. Similarly, to the trace of Figure 79, the histogram shows only one main distribution for all the voltages, namely 80mV, 100mV 120mV, and even at the high voltage of 180mV. The multi-peak Gaussian fit of the Igor software produces only one peak. This reveals that the molecule, A6 produces only one type of blockade on average.


Figure 79: Current trace of A6-aerolysin interaction over 250 ms (black) at 120mV. I_0 is the average value of the baseline. Red plot, represents the reconstruction of the DetectEvent detection of onset and end of transitions. Blue dashed line (I_1) represents the average value of the distribution of A6. The average value for the detection (red plot) correspond to I_1 . Black dashed line is zero current.

The observation of the dwell time scatter plot (left side of Figure 80) however shows that as the voltage increases, the distribution is skewed to the right, in the direction of higher residual conductance (higher I/Io values), with a tendency of decrease of the dwell time for the protruded part of the main distribution, in contrast, the distribution can still only be fitted with a single Gaussian. This raises the question for the evolution of this distribution for bigger sized molecule (see the following section).

The respective positions of the peak of the fitted Gaussians and their corresponding width, for voltages 80mV, 100mV, 120mV and 180mV are shown in Table 7. No significant increase of the peak's position is shown and only a slight increase of the width (σ) with voltage increase is detectable. This might be explained by increased fluctuations of current with the voltage.

Voltage (mV)	80	100	120	180
Peak position	0.23	0.23	0.23	0.24
Width	0.01	0.01	0.02	0.02

Table 7: Current distribution properties for A6



Figure 80: Analysis of the current histogram peak distributions for A6 for voltages of 80 mV, 100 mV, 120 mV, 180mV (left side). Fitting using the multiple-peak fitting tool yields only one distribution, which seems to broaden with the increase of voltage. Scatter plots (right side) of the dwell time, as a function of the voltage for each voltage, note the broadening of the distribution time as the voltage increases with decreasing of dwell time.

The analysis of the polynucleotide A7 is shown in Figure 81 and Figure 82. First, Figure 81 shows the current trace of A7 in the aerolysin pore. First observations reveal a noise increase for this nucleotide when interacting with the pore which might point to more fluctuations of the molecule inside the pore in comparison to A6. Secondly, it seems in this case that the events average levels (red plot), are more broadly distributed compared to A6. The two levels I_1 and I_2 correspond to the two main distributions revealed by the complete analysis of the traces and seem to correspond to the amplitude of events.

Figure 82 which shows the analysis of the level of the events in the histograms, for voltages 80mV, 100mV, and 120mV and 180mV, validates these preliminary observations. Using the multi-peaking fitting tool, with the Gaussian function it is possible to observe two peaks which add up to produce the distribution. The presence of a second peak appears at 100mV, even though it is difficult to discern between the two fits and how they play out to produce the separate distributions which add up, a small protrusion from the main distribution starts to appear at this voltage. At 120mV, and 180mV the two peaks become more discernable and it is possible to observe the protrusion due to the second peak (peak 2 in Figure 82) more clearly, especially at 180mV, because of the reduction of the total resistance at that voltage which induces a higher currents and allows a better separation of the different type of events. The first distribution is located at similar positions (see Table 8) with no significant difference as the voltage increases, a slight increase of the peak's position can be seen for the distribution (2) but the increase is not significant either.



Figure 81: Current trace of A7-aerolysin interaction over 250 ms (black) at 120mV. I_0 is the average value of the baseline. Red plot, represents the reconstruction of the DetectEvent detection of onset and end of transitions. Blue and green dashed line (I_1 and I_2) represents the average values of the current distributions of A7. Notice that the average values for the detection (red plot) corresponds on average either to I_1 or I_2 . Black dashed line is zero current.

Voltage (mV)	80	100	120	180
1 st peak position	0.20	0.20	0.20	0.21
Width	0.02	0.02	0.02	0.03
Height	18.78	44.69	46.23	42.37
2 nd peak position	/	0.23	0.24	0.25
Width	/	0.01	0.01	0.02
Height	/	5.34	4.11	6.57

Table 8: Current distribution properties for A7

It is also possible to observe that the relative height (heights shown in Table 8) of the 2^{nd} peak (labeled 2 in the graphs of Figure 82) over the first peak (labeled 1) is similar for 100 mV and 120 mV with a ratio of 0.12 and increases for 180mV with a ratio of 0.16 where it starts to be possible to see the protrusion from the distribution more clearly. This, in addition to the fact that the two distributions grow further apart with voltage, permits to discern more between them. For A7 the dwell time scatter plots shows a more visible skewedness which corresponds to lower values of dwell times in comparison to value of dwell time for the first peak labeled (1).



Figure 82: Analysis of the current histogram peak distributions for A7 for voltages of 80 mV, 100 mV, 120 mV and 180mV (left side). Fitting using the multiple-peak fitting tool yields two distributions except for 80mV, labeled (1) and (2) (corresponding to I_1 and I_2 respectively). Scatter plots (right side) of the dwell time, as a function of the voltage for each voltage. Similarly to A6 we can observe the broadening of the distribution time as the voltage increases with shallower amplitude of blockades and decreasing of dwell time for the events corresponding to distribution (2).

Figure 83 and Figure 84 show the interaction of A8 with aerolysin respectively through a current trace observation and statistical analysis of the events at 120mV, and all-events analysis shown for the voltages 80mV, 100mV and 120mV.

In Figure 83 it is possible to see the different contributions of the events to the two different distributions at 120mV. First, similarly to A7, events are distributed around two sets of average current values (I_1 and I_2). Secondly, we also notice from qualitative observations that the noise increases in comparison to A7.

In comparison to A6-A7 an overall decrease in relative current conductance is seen (Table 9), which is in accordance with the increase in size of the molecule. In contrast to A6-A7 the multi-peak fitting shows that two distributions add up to produce the skewed distribution already at 80mV. As demonstrated in Figure 84 the whole distribution gets broader at 100mV in comparison to 80mV and the contribution of the second peak becomes clearer at 120mV.



Figure 83: Current trace of A8-aerolysin interaction over 250 ms (black) at 120mV. I_0 is the average value of the baseline. Red plot, represents the reconstruction of the DetectEvent detection of onset and end of transitions. Blue and green dashed line (I_1 and I_2) represents the average values of the current distributions of A8. One can notice that the average values for the detection (red plot) correspond on average either to I_1 or I_2 . Black dashed line is zero current.

The relative height ratio (height of second peak over the height of the first peak which are shown in Table 9) increases at 120mV in comparison to 80 and 100mV, with the respective values of 0.15, 0.17 and 0.29 for 80mV, 100mV and 120mV. This contributes at the formation of a clearer separation between the two distributions, and reflects the fact that the proportion of the second type of events (shallower and shorter) increases.

It is also possible to observe through the dwell time scatter plot that the values of the distribution decrease with the increase of voltage similarly to A6-A7. In the same way the dwell times corresponding to the distribution labeled (2) are shorter.

Voltage (mV)	80	100	120
1 st peak position	0.17	0.16	0.16
Width	0.03	0.03	0.03
Height	27.22	88.91	38.35
2 nd peak position	0.20	0.20	0.20
Width	0.02	0.02	0.01
Height	4.16	14.84	11.09

Table 9: Current distribution properties for A8



Figure 84: Analysis of the current histogram peak distributions for A8 for voltages of 80 mV, 100 mV and 120 mV (left side). Fitting using the multiple-peak fitting tool yields two distributions, labeled (1) and (2) (corresponding to I_1 and I_2 respectively). Scatter plots (right side) of the dwell time, as a function of the voltage for each voltage, note the broadening of the distribution of time as the voltage increases, with shallower amplitude of blockades and decreasing of dwell time for the distribution (2).

The trace of Figure 85 displays a similar effect than A7-A8, namely events distributed around two average values of current. In comparison to A7 and A8, for A9, the two distributions separate even more, and the separation becomes clearer with the increase of the voltage (see Figure 86). This difference becomes most significant at 120mV (the two peaks are located further apart at 120mV) and the peaks become more apparent in the total distribution. This separation is also possible to observe from the dwell time scatter plot in Figure 86 at 120mV, and it appears clearly that the events corresponding to distribution (2) are shorter in comparison to distribution (1).



Figure 85: Current trace of A9-aerolysin interaction over 400 ms (black) at 120mV. I_0 is the average value of the baseline. Red plot, represents the reconstruction of the DetectEvent detection of onset and end of transitions. Blue and green dashed line (I_1 and I_2) represents the average values of the current distributions of A9. One can notice that the average values for the detection (red plot) correspond on average either to I_1 or I_2 . Black dashed line is zero current.

The relative heights of the second peak over the first peak are respectively according to the voltage increase: 0.42, 0.53 and 0.64, indicating that the proportion of events corresponding to the second distribution (2) increases. This combined to the fact that the width of the two distributions becomes similar as the voltage increases contribute jointly the amplitude of the events corresponding to either distribution becomes more discernable

Voltage (mV)	80	100	120
1 st peak position	0.16	0.15	0.14
Width	0.02	0.02	0.02
Height	47.28	55.93	34.74
2 nd peak position	0.18	0.18	0.18
Width	0.01	0.02	0.02
Height	19.81	29.65	22.31

Table 10: Current distribution properties for A9



Figure 86: Analysis of the current histogram peak distributions for A9 for voltages of 80 mV, 100 mV and 120 mV (left side). Fitting with the multiple-peak fitting tool yields two distributions, labeled (1) and (2) (corresponding to I_1 and I_2 respectively) with higher amplitude for (2) in this case in comparison to A6-8. Scatter plots (right side) of the dwell time, as a function of the voltage for each voltage, note the broadening of the distribution of time as the voltage increases, with shallower amplitude of blockades and decreasing of dwell time for the distribution (2).

Finally the last molecule analyzed in this study is A10. The analysis corresponding to A10 is shown in Figure 87 and Figure 88. Similarly, in Figure 87 the DetectEvent software yielded a detection which corresponds to either levels (I_1 or I_2). In Figure 88, the two peaks are clearly independently adding up to produce a two peak distribution and these two distributions are clearly discernable at 80 mV, 100m V and 120mV.



Figure 87: Current trace of A10-aerolysin interaction over 400 ms (black) at 120mV. I_0 is the average value of the baseline. Red plot, represents the reconstruction of the DetectEvent detection of onset and end of transitions. Blue and green dashed line (I_1 and I_2) represents the average values of the current distributions of A10. One can notice that the average values for the detection (red plot) correspond on average either to I_1 or I_2 . Black dashed line is zero current.

In this case the relative height (extracted from Table 11) has the respective values for 80mV, 100mV and 120mV of 0.55, 0.62 and 0.58. These values seem to be similar than for A9, and double in comparison to A8 and smaller polyadenines. The analysis of the peaks seems to yield similar patterns for the distribution of current with no significant difference between the two molecules (A9 and A10); this seems to point to the fact that the mechanism of ssDNA orientation dynamics becomes similar between the two last studied molecules.

Voltage (mV)	80	100	120
1 st peak position	0.13	0.12	0.11
Width	0.02	0.02	0.02
Height	72.85	79.23	103.43
2 nd peak position	0.16	0.16	0.15
Width	0.01	0.01	0.02
Height	40.01	49.22	60.10
Tergit	70.01	77.22	00.10

Table 11: Current distribution properties for A10



Figure 88: Analysis of the current histogram peak distributions for A10 for voltages of 80 mV, 100 mV and 120 mV (left side). Fitting using the multiple-peak fitting tool yields two distributions, labeled (1) and (2) (corresponding to I_1 and I_2 respectively). Scatter plots (right side) of the dwell time, as a function of the voltage for each voltage, note the broadening of the distribution of time as the voltage increases, with shallower amplitude of blockades and decreasing of dwell time for the distribution (2).

Figure 89 shows the evolution of dwell time *vs.* voltage for A6-A10. We have analyzed the dwell times of the distribution (I_1) for all molecules, and analyzed the second distribution (I_2) only for A9 and A10 where it was easier to distinguish between the two different scatter plots. The separation between the two distributions appeared for voltages above 70 mV. This might be due to the fact that the signal's amplitude increases with the increase of voltage and permits to discern between small current differences.

For the first distribution, above 40mV, the dwell time decreases exponentially with the voltage for all the analyzed molecules which suggests a translocation at the voltage mentioned and above. It is also clear that the dwell time increases with the molecule's size which might indicate a strong binding of the adenine nucleotide which would lead to a proportional increase of dwell time as the molecule's size increases.

Secondly the analysis of the second peak, clearly validates the previous qualitative observations, that the dwell times are shorter for this distribution. Similarly to the first distribution the dwell times increase with the molecule's size. This seems to demonstrate two different dynamics for the same molecule inside the same nanopore.



Figure 89: Dwell times of the distributions I_1 and I_2 *vs.* voltage. The dwell times for the first distribution (I_1 , left side) correspond to A6-10. The dwell times for the second distribution (I_2 , right side) correspond to A9 and A10. For both graphs the lines are exponential fits. Dwell time axes are in logarithmic scale.

II. Discussion

From the analysis of the distribution of current of events from the interaction of aerolysin with monodisperse solutions of A6 to A10 nucleotides, the observation of a two peak distribution can be related to orientation of the molecule at the pore's entrance which produces different signals accordingly. This observation was already briefly discussed for the A14 molecule and aerolysin in 1M KCl in ref.[47].

The first deduction from these observation is that the molecule A6 doesn't show a significant difference in current, probably due to the fact that the molecule, which measures ~ 2 nm in length, is in the same order of the pore's diameter and smaller than the pore's length. Thus, whatever its orientation, comparable size of the molecule with the pore's dimensions, might prevent it from producing a significant difference in current

In addition, the molecule's flexibility, with less than 1.5 nm persistence length for ssDNA in high concentrated salt conditions of 4M (1.5 nm in 2M NaCl [158]), might allow the molecule's reorientations inside the pore, oscillating between its two orientations while translocating through the pore. Broadening of the total distribution of current in comparison to A5 discussed in the previous chapter (chapter 6), reinforces this observation. Similar observations of reorientation inside the pore have been discussed in ref. [149] for simulations of ssDNA in hemolysin.

Furthermore, the fact that a two peak distribution for A7-A8 appears, especially as the voltage increases, indicates that above the 6 bases length, it becomes more difficult for the molecule to orient freely inside the pore, especially under the influence of the voltage as it may become stiffer under the influence of the electrical field. The flexibility is thus reduced inside the pore and a more prominent two peak distribution at 120mV appears.

The last results clearly show that for 9 adenine bases and more the orientation dependent interaction is prominent even at lower voltages (70mV) and the two peaks are further apart, indicating that depending on its orientation, the oligonucleotides drastically modifies current transmission, depending on its 3' or 5' orientation. As explained in ref [49], [50], experiments and simulations of DNA hairpins escape experiments with α -hemolysin showed two different distributions of current depending on the 3' or 5' orientation. Mathé *et al.* [49] suggest through their simulation that the bases orient in the direction of the 3' end in solution, rendering its translocation slower of faster and producing events of more or less amplitude depending on its

orientation at the pore's opening. A good analogy is taken to explain this behaviour using the Christmas tree passage through the door as an example: the passage of the Christmas tree would be easier when it is oriented leaves first, as opposed to the trunk orientation first. Similarly here, because the bases are oriented towards the 3' direction, if the molecule is in the 3' orientation, at the pore's entrance the translocation would be faster in comparison to the 5' orientation first. Respectively, the first orientation will block less ions through the pore than the second one; this is clearly the case for our previously described results for deeper distributions of current (I_1) which correspond to longer dwell times and shallower blocks.

III. Summary

We have demonstrated through this study that the interaction of polyadenines with the aerolysin nanopore might be orientation dependent. It is possible for the smaller molecules (A6) to orient when they enter the pore, but because the molecule's size is close to or smaller than the pore's diameter (for A6 and below) combined to the fact that ssDNA is very flexible might lead to reorientation of the smaller polyadenines while inside the pore, therefore, they don't show this orientation dependent interaction in our experiments. In comparison, bigger molecules such as A7 and A8, seem to display this orientation-dependent interaction but at a smaller proportion than A9 and A10. This points to the fact that the difference in blockades due to the orientation of these molecules might still not be significant enough in comparison to the pore's length and the amount of total ions blocked. A9 seems to have the limit size at which this effect of orientation becomes more significant. Similar observations than A9 are made for A10.

Conclusion

We have shown through these work different interactions of various polymers and molecules, with mainly two biological nanopores and investigated both the molecules and pore's properties.

We have demonstrated in the study focusing on the interaction of PEGs and cyclodextrins that the electroosmotic flow seems to be stronger in aerolysin in comparison to hemolysin. The use of neutral molecules, namely cyclodextrins, allowed the probing of the flows in both pores and in two different salts (KCl and LiCl). The mechanism of interaction seemed to be similar in both salts but was stronger in LiCl. The interaction of PEGs with aerolysin showed two different dynamics, and was interpreted to be due to the fact that PEG might complex cations in KCl in contrast to LiCl, where it is thought to be neutral. However this interpretation is still debatable, due to the contradiction of some results, especially for the interaction of PEGs with hemolysin, where no dependency of the frequency on the electrical field was observed in KCl for interaction from the cap side of the pore. Even though in both KCl and LiCl the interaction of PEGs seemed to have the same nature due to the fact that the pore is still sensitive to the mass in both conditions, we have also shown that the interaction of PEGs and aerolysin can be promoted by salt. In all cases it seems important to probe the role of the ions in the interaction of these various molecules with the nanopores, and the necessity to perform more experiments on this subject (ion type influence) in order to understand the nature of the bindng between the molecules and the pore at the molecular level.

We were able to detect precision polymers, namely, poly(phosphodiester)s of various types and sizes. The interaction of these polymers seemed to be best with aerolysin in comparison to hemolysin. We investigated the simplest type of these polymers, with no side chains: the interaction of this polymer seemed to be orientation dependent and we attributed its strong interaction with the pore to the binding of the polymer's DNA primer sequence combines to a high flexibility of the polymer which lengthened the duration of blocks. This high flexibility was also correlated to the fact that seemingly, the polymer only translocates above very high voltages. The detection of similar polymers with side chains revealed that they can also interact with the aerolysin nanopore, however in that case, the detected events were more complex than the first studied polymer. This was attributed to the fact that these polymers might aggregate in solution. By coupling the characterization of these molecules using wave-scattering we were able to validate this hypothesis as the characterization showed aggregation of some of the polymers studied. In any case it seemed important to couple nanopore analysis with other types of physical characterization in order to have a better understanding of the behavior of these polymers in solutions. We also performed a few experiments with these polymers and MspA as a proof of principle, but more data needs to be acquired in order the have a validation of the observed interaction. For the ultimate goal of sequencing, it seems interesting to try even longer structures or to couple polymerases to the pores as has been done with DNA, in order to slow down the interaction and attempt at obtaining a block characterization of these molecules.

The study of dynamics of small polyadenines with the aerolysin nanopore showed a complex interaction. Indeed, it was demonstrated that the events induced by the molecules contained various patterns, which after analysis gave information of the kinetics of the A3-4-5 inside the pore. The analysis of the deep blocks and main events pointed to the fact that there exist two binding sites for this molecule inside the aerolysin nanopore which create two types of blocks and between whom the molecule can oscillate. The whole kinetic process in governed by three energy barriers: 1. for entry and binding to the first site, 2. for binding to the second site and 3. for translocation. Through the experimental findings these barriers were possible to overcome under the influence of the electrical field. We developed a model which aimed at describing this system, and succeeded in fitting many of the experimental results. This led to the further investigation of bigger sized polyadenines, which showed a variation of dynamics. The study of A6-7-8-9-10 showed a mechanism of orientation-dependent interaction with the aerolysin nanopore that seemed to start from A7 and A8 sizes. We attributed the non-observation of orientation dependent interaction of smaller molecules to the fact that their size and the flexibility of ssDNA, might still cause the reorientation of the small oligonucleotides inside the nanopore and that this small size doesn't allow for significant changes on the ionic current passing through the pore, therefore cannot be observed. These findings seem to be of importance in uncovering the dynamics and translocation of DNA with nanopores in general, and can help for a better understanding of pore-DNA interactions.

References

- [1] W. H. Coulter, "Means for counting particles suspended in a fluid," US2656508A, 20-Oct-1953.
- [2] B. Sakmann and E. Neher, "Patch Clamp Techniques for Studying Ionic Channels in Excitable Membranes," Annu. Rev. Physiol., vol. 46, no. 1, pp. 455–472, 1984.
- [3] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, "Characterization of individual polynucleotide molecules using a membrane channel," *Proc. Natl. Acad. Sci.*, vol. 93, no. 24, pp. 13770–13773, Nov. 1996.
- [4] D. Branton *et al.*, "The potential and challenges of nanopore sequencing," *Nat. Biotechnol.*, vol. 26, no. 10, pp. 1146–1153, Oct. 2008.
- [5] B. M. Venkatesan and R. Bashir, "Nanopore sensors for nucleic acid analysis," *Nat. Nanotechnol.*, vol. 6, no. 10, pp. 615–624, Oct. 2011.
- [6] J. J. Kasianowicz, J. W. F. Robertson, E. R. Chan, J. E. Reiner, and V. M. Stanford, "Nanoscopic porous sensors," *Annu. Rev. Anal. Chem. Palo Alto Calif*, vol. 1, pp. 737–766, 2008.
- [7] A. Oukhaled, L. Bacri, M. Pastoriza-Gallego, J.-M. Betton, and J. Pelta, "Sensing Proteins through Nanopores: Fundamental to Applications," *ACS Chem. Biol.*, vol. 7, no. 12, pp. 1935–1949, Dec. 2012.
- [8] L. Auvray, B. Duplantier, and C. Sykes, Physique des polymères et membranes biologiques: majeures de physique et chimie du vivant ... : promotion 2004, année 3, majeure 2 ... École polytechnique, 2007.
- [9] M. Daoud and P. G. D. Gennes, "Statistics of macromolecular solutions trapped in small pores," *J. Phys.*, vol. 38, no. 1, pp. 85–93, Jan. 1977.
- [10] M. Muthukumar, Polymer Translocation. CRC Press, 2016.
- [11] M. Muthukumar, "Polymer translocation through a hole," J. Chem. Phys., vol. 111, no. 22, pp. 10371–10374, Nov. 1999.
- [12] W. Sung and P. J. Park, "Polymer Translocation through a Pore in a Membrane," *Phys. Rev. Lett.*, vol. 77, no. 4, pp. 783–786, Jul. 1996.
- [13] E. A. Di Marzio and A. J. Mandell, "Phase transition behavior of a linear macromolecule threading a membrane," J. Chem. Phys., vol. 107, no. 14, pp. 5510–5514, Oct. 1997.
- [14] S. Daoudi and F. Brochard, "Flows of Flexible Polymer Solutions in Pores," *Macromolecules*, vol. 11, no. 4, pp. 751–758, Jul. 1978.
- [15] L. Béguin, B. Grassl, F. Brochard-Wyart, M. Rakib, and H. Duval, "Suction of hydrosoluble polymers into nanopores," *Soft Matter*, vol. 7, no. 1, pp. 96–103, Dec. 2010.
- [16] T. D. Long and J. L. Anderson, "Flow-dependent rejection of polystyrene from microporous membranes," J. Polym. Sci. Polym. Phys. Ed., vol. 22, no. 7, pp. 1261–1281, Jul. 1984.

- [17] P.-G. de Gennes, "Flexible Polymers in Nanopores," in *Polymers in Confined Environments*, Springer, Berlin, Heidelberg, 1999, pp. 91–105.
- [18] T. Auger *et al.*, "Zero-Mode Waveguide Detection of Flow-Driven DNA Translocation through Nanopores," *Phys. Rev. Lett.*, vol. 113, no. 2, p. 028302, juillet 2014.
- [19] L.-Q. Gu, S. Cheley, and H. Bayley, "Electroosmotic enhancement of the binding of a neutral molecule to a transmembrane pore," *Proc. Natl. Acad. Sci.*, vol. 100, no. 26, pp. 15498–15503, Dec. 2003.
- [20] L. Mereuta *et al.*, "Slowing down single-molecule trafficking through a protein nanopore reveals intermediates for peptide translocation," *Sci. Rep.*, vol. 4, Jan. 2014.
- [21] M. Soskine, A. Biesemans, and G. Maglia, "Single-Molecule Analyte Recognition with ClyA Nanopores Equipped with Internal Protein Adaptors," J. Am. Chem. Soc., vol. 137, no. 17, pp. 5793– 5797, mai 2015.
- [22] A. Asandei, I. Schiopu, M. Chinappi, C. H. Seo, Y. Park, and T. Luchian, "Electroosmotic Trap Against the Electrophoretic Force Near a Protein Nanopore Reveals Peptide Dynamics During Capture and Translocation," ACS Appl. Mater. Interfaces, vol. 8, no. 20, pp. 13166–13179, May 2016.
- [23] S. P. Bhamidimarri, J. D. Prajapati, B. van den Berg, M. Winterhalter, and U. Kleinekathöfer, "Role of Electroosmosis in the Permeation of Neutral Molecules: CymA and Cyclodextrin as an Example," *Biophys. J.*, vol. 110, no. 3, pp. 600–611, Feb. 2016.
- [24] G. Huang, K. Willems, M. Soskine, C. Wloka, and G. Maglia, "Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores," *Nat. Commun.*, vol. 8, no. 1, p. 935, Oct. 2017.
- [25] R. F. Probstein, *Physicochemical Hydrodynamics: An Introduction*. John Wiley & Sons, 2005.
- [26] O. V. Krasilnikov, R. Z. Sabirov, V. I. Ternovsky, P. G. Merzliak, and J. N. Muratkhodjaev, "A simple method for the determination of the pore radius of ion channels in planar lipid bilayer membranes," *FEMS Microbiol. Immunol.*, vol. 5, no. 1–3, pp. 93–100, Sep. 1992.
- [27] S. M. Bezrukov, I. Vodyanoy, and V. A. Parsegian, "Counting polymers moving through a single ion channel," *Nature*, vol. 370, no. 6487, pp. 279–281, juillet 1994.
- [28] S. M. Bezrukov, I. Vodyanoy, R. A. Brutyan, and J. J. Kasianowicz, "Dynamics and Free Energy of Polymers Partitioning into a Nanoscale Pore," *Macromolecules*, vol. 29, no. 26, pp. 8517–8522, Jan. 1996.
- [29] P. G. Merzlyak, L. N. Yuldasheva, C. G. Rodrigues, C. M. Carneiro, O. V. Krasilnikov, and S. M. Bezrukov, "Polymeric nonelectrolytes to probe pore geometry: application to the alpha-toxin transmembrane channel.," *Biophys. J.*, vol. 77, no. 6, pp. 3023–3033, Dec. 1999.
- [30] L. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley, and J. E. Gouaux, "Structure of Staphylococcal α-Hemolysin, a Heptameric Transmembrane Pore," *Science*, vol. 274, no. 5294, pp. 1859–1865, Dec. 1996.

- [31] L. Movileanu, S. Cheley, and H. Bayley, "Partitioning of Individual Flexible Polymers into a Nanoscopic Protein Pore," *Biophys. J.*, vol. 85, no. 2, pp. 897–910, Aug. 2003.
- [32] O. V. Krasilnikov, C. G. Rodrigues, and S. M. Bezrukov, "Single Polymer Molecules in a Protein Nanopore in the Limit of a Strong Polymer-Pore Attraction," *Phys. Rev. Lett.*, vol. 97, no. 1, p. 018301, juillet 2006.
- [33] J. W. F. Robertson, C. G. Rodrigues, V. M. Stanford, K. A. Rubinson, O. V. Krasilnikov, and J. J. Kasianowicz, "Single-molecule mass spectrometry in solution using a solitary nanopore," *Proc. Natl. Acad. Sci.*, vol. 104, no. 20, pp. 8207–8211, May 2007.
- [34] C. G. Rodrigues, D. C. Machado, S. F. Chevtchenko, and O. V. Krasilnikov, "Mechanism of KCl Enhancement in Detection of Nonionic Polymers by Nanopore Sensors," *Biophys. J.*, vol. 95, no. 11, pp. 5186–5192, Dec. 2008.
- [35] G. Baaken, I. Halimeh, L. Bacri, J. Pelta, A. Oukhaled, and J. C. Behrends, "High-Resolution Size-Discrimination of Single Nonionic Synthetic Polymers with a Highly Charged Biological Nanopore," ACS Nano, vol. 9, no. 6, pp. 6443–6449, juin 2015.
- [36] C. Cao, Y.-L. Ying, Z.-L. Hu, D.-F. Liao, H. Tian, and Y.-T. Long, "Discrimination of oligonucleotides of different lengths with a wild-type aerolysin nanopore," *Nat Nano*, vol. advance online publication, Apr. 2016.
- [37] M. F. Breton, F. Discala, L. Bacri, D. Foster, J. Pelta, and A. Oukhaled, "Exploration of Neutral Versus Polyelectrolyte Behavior of Poly(ethylene glycol)s in Alkali Ion Solutions using Single-Nanopore Recording," J. Phys. Chem. Lett., vol. 4, no. 13, pp. 2202–2208, juillet 2013.
- [38] F. Piguet, F. Discala, M.-F. Breton, J. Pelta, L. Bacri, and A. Oukhaled, "Electroosmosis through α-Hemolysin That Depends on Alkali Cation Type," J. Phys. Chem. Lett., vol. 5, no. 24, pp. 4362– 4367, décembre 2014.
- [39] S. E. Henrickson, M. Misakian, B. Robertson, and J. J. Kasianowicz, "Driven DNA Transport into an Asymmetric Nanometer-Scale Pore," *Phys. Rev. Lett.*, vol. 85, no. 14, pp. 3057–3060, Oct. 2000.
- [40] L.-Q. Gu, O. Braha, S. Conlan, S. Cheley, and H. Bayley, "Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter," *Nature*, vol. 398, no. 6729, pp. 686–690, avril 1999.
- [41] J. Clarke, H.-C. Wu, L. Jayasinghe, A. Patel, S. Reid, and H. Bayley, "Continuous base identification for single-molecule nanopore DNA sequencing," *Nat. Nanotechnol.*, vol. 4, no. 4, p. 265, Apr. 2009.
- [42] S. Howorka, S. Cheley, and H. Bayley, "Sequence-specific detection of individual DNA strands using engineered nanopores," *Nat. Biotechnol.*, vol. 19, no. 7, pp. 636–639, Jul. 2001.
- [43] D. Stoddart, A. J. Heron, E. Mikhailova, G. Maglia, and H. Bayley, "Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore," *Proc. Natl. Acad. Sci.*, vol. 106, no. 19, pp. 7702–7707, May 2009.

- [44] N. Ashkenasy, J. Sánchez-Quesada, H. Bayley, and M. R. Ghadiri, "Recognizing a Single Base in an Individual DNA Strand: A Step Toward DNA Sequencing in Nanopores," *Angew. Chem. Int. Ed.*, vol. 44, no. 9, pp. 1401–1404, Feb. 2005.
- [45] T. Z. Butler, M. Pavlenok, I. M. Derrington, M. Niederweis, and J. H. Gundlach, "Single-molecule DNA detection with an engineered MspA protein nanopore," *Proc. Natl. Acad. Sci.*, vol. 105, no. 52, pp. 20647–20652, Dec. 2008.
- [46] I. M. Derrington *et al.*, "Nanopore DNA sequencing with MspA," *Proc. Natl. Acad. Sci.*, vol. 107, no. 37, pp. 16060–16065, Sep. 2010.
- [47] C. Cao *et al.*, "Mapping the sensing spots of aerolysin for single oligonucleotides analysis," *Nat. Commun.*, vol. 9, no. 1, p. 2823, Jul. 2018.
- [48] J. Mathé, H. Visram, V. Viasnoff, Y. Rabin, and A. Meller, "Nanopore Unzipping of Individual DNA Hairpin Molecules," *Biophys. J.*, vol. 87, no. 5, pp. 3205–3212, Nov. 2004.
- [49] J. Mathé, A. Aksimentiev, D. R. Nelson, K. Schulten, and A. Meller, "Orientation discrimination of single-stranded DNA inside the α-hemolysin membrane channel," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 35, pp. 12377–12382, Aug. 2005.
- [50] M. Wanunu, B. Chakrabarti, J. Mathé, D. R. Nelson, and A. Meller, "Orientation-dependent interactions of DNA with an \$\ensuremath{\alpha}\$-hemolysin channel," *Phys. Rev. E*, vol. 77, no. 3, p. 031904, Mar. 2008.
- [51] J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko, "Ion-beam sculpting at nanometre length scales," *Nature*, vol. 412, no. 6843, pp. 166–169, Jul. 2001.
- [52] S. Garaj, W. Hubbard, A. Reina, J. Kong, D. Branton, and J. A. Golovchenko, "Graphene as a subnanometre trans-electrode membrane," *Nature*, vol. 467, no. 7312, pp. 190–193, Sep. 2010.
- [53] C. A. Merchant *et al.*, "DNA Translocation through Graphene Nanopores," *Nano Lett.*, vol. 10, no. 8, pp. 2915–2921, Aug. 2010.
- [54] G. F. Schneider *et al.*, "DNA Translocation through Graphene Nanopores," *Nano Lett.*, vol. 10, no. 8, pp. 3163–3167, Aug. 2010.
- [55] A. R. Hall, A. Scott, D. Rotem, K. K. Mehta, H. Bayley, and C. Dekker, "Hybrid pore formation by directed insertion of α-haemolysin into solid-state nanopores," *Nat. Nanotechnol.*, vol. 5, no. 12, pp. 874–877, Dec. 2010.
- [56] E. C. Yusko *et al.*, "Controlling protein translocation through nanopores with bio-inspired fluid walls," *Nat. Nanotechnol.*, vol. 6, no. 4, pp. 253–260, Apr. 2011.
- [57] S. R. Wente and M. P. Rout, "The Nuclear Pore Complex and Nuclear Transport," *Cold Spring Harb. Perspect. Biol.*, vol. 2, no. 10, p. a000562, Oct. 2010.
- [58] C. Strambio-De-Castillia, M. Niepel, and M. P. Rout, "The nuclear pore complex: bridging nuclear transport and gene regulation," *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 7, pp. 490–501, Jul. 2010.
- [59] D. A. Doyle *et al.*, "The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity," *Science*, vol. 280, no. 5360, pp. 69–77, Apr. 1998.

- [60] S. Choe, "Ion channel structure: Potassium channel structures," *Nat. Rev. Neurosci.*, vol. 3, no. 2, pp. 115–121, Feb. 2002.
- [61] K. Ribbeck and D. Görlich, "Kinetic analysis of translocation through nuclear pore complexes," *EMBO J.*, vol. 20, no. 6, pp. 1320–1330, Mar. 2001.
- [62] W. H. Roos, I. L. Ivanovska, A. Evilevitch, and G. J. L. Wuite, "Viral capsids: Mechanical characteristics, genome packaging and delivery mechanisms," *Cell. Mol. Life Sci.*, vol. 64, no. 12, p. 1484, Mar. 2007.
- [63] P. Guo, Z. Zhao, J. Haak, S. Wang, and T. Weitao, "Common Mechanisms of DNA translocation motors in Bacteria and Viruses Using One-way Revolution Mechanism without Rotation," *Biotechnol. Adv.*, vol. 32, no. 4, pp. 853–872, 2014.
- [64] S. Howorka and Z. Siwy, "Nanopore analytics: sensing of single molecules," *Chem. Soc. Rev.*, vol. 38, no. 8, pp. 2360–2384, Jul. 2009.
- [65] S. Howorka, "Building membrane nanopores," *Nat. Nanotechnol.*, vol. 12, no. 7, pp. 619–630, Jul. 2017.
- [66] M. Lepoitevin, T. Ma, M. Bechelany, J.-M. Janot, and S. Balme, "Functionalization of single solid state nanopores to mimic biological ion channels: A review," *Adv. Colloid Interface Sci.*, vol. 250, pp. 195–213, Dec. 2017.
- [67] B. J. Berube and J. Bubeck Wardenburg, "Staphylococcus aureus α-Toxin: Nearly a Century of Intrigue," *Toxins*, vol. 5, no. 6, pp. 1140–1166, Jun. 2013.
- [68] F. C. O. Los, T. M. Randis, R. V. Aroian, and A. J. Ratner, "Role of Pore-Forming Toxins in Bacterial Infectious Diseases," *Microbiol. Mol. Biol. Rev. MMBR*, vol. 77, no. 2, pp. 173–207, Jun. 2013.
- [69] M. D. Peraro and F. G. van der Goot, "Pore-forming toxins: ancient, but never really out of fashion," *Nat. Rev. Microbiol.*, vol. 14, no. 2, pp. 77–92, Feb. 2016.
- [70] M. R. Gonzalez, M. Bischofberger, L. Pernot, F. G. van der Goot, and B. Frêche, "Bacterial poreforming toxins: The (w)hole story?," *Cell. Mol. Life Sci.*, vol. 65, no. 3, pp. 493–507, Feb. 2008.
- [71] S. Agah, M. Zheng, M. Pasquali, and A. B. Kolomeisky, "DNA sequencing by nanopores: advances and challenges," J. Phys. Appl. Phys., vol. 49, no. 41, p. 413001, 2016.
- [72] J. E. Reiner, J. J. Kasianowicz, B. J. Nablo, and J. W. F. Robertson, "Theory for polymer analysis using nanopore-based single-molecule mass spectrometry," *Proc. Natl. Acad. Sci.*, vol. 107, no. 27, pp. 12080–12085, Jul. 2010.
- [73] A. E. Chavis *et al.*, "Single Molecule Nanopore Spectrometry for Peptide Detection," ACS Sens., vol. 2, no. 9, pp. 1319–1328, Sep. 2017.
- [74] E. Gouaux, "α-Hemolysin fromStaphylococcus aureus:An Archetype of β-Barrel, Channel-Forming Toxins," J. Struct. Biol., vol. 121, no. 2, pp. 110–122, Jan. 1998.

- [75] C. G. Rodrigues, D. C. Machado, A. M. B. da Silva, J. J. S. Júnior, and O. V. Krasilnikov, "Hofmeister Effect in Confined Spaces: Halogen Ions and Single Molecule Detection," *Biophys. J.*, vol. 100, no. 12, pp. 2929–2935, Jun. 2011.
- [76] C. T. A. Wong and M. Muthukumar, "Polymer translocation through α-hemolysin pore with tunable polymer-pore electrostatic interaction," *J. Chem. Phys.*, vol. 133, no. 4, Jul. 2010.
- [77] J. J. Kasianowicz and S. M. Bezrukov, "Protonation dynamics of the alpha-toxin ion channel from spectral analysis of pH-dependent current fluctuations.," *Biophys. J.*, vol. 69, no. 1, pp. 94–105, Jul. 1995.
- [78] F. Piguet *et al.*, "High Temperature Extends the Range of Size Discrimination of Nonionic Polymers by a Biological Nanopore," *Sci. Rep.*, vol. 6, Dec. 2016.
- [79] M. W. Parker *et al.*, "Structure of the Aeromonas toxin proaerolysin in its water-soluble and membrane-channel states," *Nature*, vol. 367, no. 6460, pp. 292–295, Jan. 1994.
- [80] F. G. van der Goot, J. Ausio, K. R. Wong, F. Pattus, and J. T. Buckley, "Dimerization stabilizes the pore-forming toxin aerolysin in solution.," *J. Biol. Chem.*, vol. 268, no. 24, pp. 18272–18279, Aug. 1993.
- [81] K. R. Hardie, A. Schulze, M. W. Parker, and J. T. Buckley, "Vibrio spp. secrete proaerolysin as a folded dimer without the need for disulphide bond formation," *Mol. Microbiol.*, vol. 17, no. 6, pp. 1035–1044.
- [82] M. W. Parker, F. G. V. D. Goot, and J. T. Buckley, "Aerolysin the ins and outs of a model channel-forming toxin," *Mol. Microbiol.*, vol. 19, no. 2, pp. 205–212.
- [83] M. T. Degiacomi *et al.*, "Molecular assembly of the aerolysin pore reveals a swirling membraneinsertion mechanism," *Nat. Chem. Biol.*, vol. 9, no. 10, pp. 623–629, Oct. 2013.
- [84] I. Iacovache, S. D. Carlo, N. Cirauqui, M. D. Peraro, F. G. van der Goot, and B. Zuber, "Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process," *Nat. Commun.*, vol. 7, p. ncomms12062, Jul. 2016.
- [85] M. Pastoriza-Gallego *et al.*, "Dynamics of Unfolded Protein Transport through an Aerolysin Pore," J. Am. Chem. Soc., vol. 133, no. 9, pp. 2923–2931, Mar. 2011.
- [86] B. Cressiot, E. Braselmann, A. Oukhaled, A. H. Elcock, J. Pelta, and P. L. Clark, "Dynamics and Energy Contributions for Transport of Unfolded Pertactin through a Protein Nanopore," ACS Nano, vol. 9, no. 9, pp. 9050–9061, Sep. 2015.
- [87] F. Piguet, H. Ouldali, M. Pastoriza-Gallego, P. Manivet, J. Pelta, and A. Oukhaled, "Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore," *Nat. Commun.*, vol. 9, no. 1, p. 966, Mar. 2018.
- [88] L. Payet *et al.*, "Temperature Effect on Ionic Current and ssDNA Transport through Nanopores," *Biophys. J.*, vol. 109, no. 8, pp. 1600–1607, Oct. 2015.
- [89] M. Niederweis, "Mycobacterial porins new channel proteins in unique outer membranes," Mol. Microbiol., vol. 49, no. 5, pp. 1167–1177.

- [90] P. J. Brennan and H. Nikaido, "The Envelope of Mycobacteria," Annu. Rev. Biochem., vol. 64, no. 1, pp. 29–63, 1995.
- [91] H. Nikaido, "Porins and specific channels of bacterial outer membranes," *Mol. Microbiol.*, vol. 6, no. 4, pp. 435–442.
- [92] H. Nikaido, "Porins and specific diffusion channels in bacterial outer membranes.," J. Biol. Chem., vol. 269, no. 6, pp. 3905–3908, Feb. 1994.
- [93] M. Faller, M. Niederweis, and G. E. Schulz, "The Structure of a Mycobacterial Outer-Membrane Channel," *Science*, vol. 303, no. 5661, pp. 1189–1192, Feb. 2004.
- [94] E. A. Manrao *et al.*, "Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase," *Nat. Biotechnol.*, vol. 30, no. 4, p. 349, Apr. 2012.
- [95] P. Mueller, D. O. Rudin, H. T. Tien, and W. C. Wescott, "Reconstitution of Excitable Cell Membrane Structure in Vitro," *Circulation*, vol. 26, no. 5, pp. 1167–1171, Nov. 1962.
- [96] M. Montal and P. Mueller, "Formation of Bimolecular Membranes from Lipid Monolayers and a Study of Their Electrical Properties," *Proc. Natl. Acad. Sci.*, vol. 69, no. 12, pp. 3561–3566, Dec. 1972.
- [97] K. Simons and J. L. Sampaio, "Membrane Organization and Lipid Rafts," Cold Spring Harb. Perspect. Biol., vol. 3, no. 10, Oct. 2011.
- [98] V. A. Frolov, A. V. Shnyrova, and J. Zimmerberg, "Lipid Polymorphisms and Membrane Shape," *Cold Spring Harb. Perspect. Biol.*, vol. 3, no. 11, p. a004747, Nov. 2011.
- [99] M. F. Brown, "Curvature Forces in Membrane Lipid–Protein Interactions," *Biochemistry (Mosc.)*, vol. 51, no. 49, pp. 9782–9795, Dec. 2012.
- [100] M. S. Bretscher, "Asymmetrical Lipid Bilayer Structure for Biological Membranes," *Nature. New Biol.*, vol. 236, no. 61, pp. 11–12, Mar. 1972.
- [101] J. E. Rothman and J. Lenard, "Membrane asymmetry," *Science*, vol. 195, no. 4280, pp. 743–753, Feb. 1977.
- [102] T. Heimburg, "Lipid ion channels," Biophys. Chem., vol. 150, no. 1, pp. 2–22, Aug. 2010.
- [103] C. G. Siontorou, G.-P. Nikoleli, D. P. Nikolelis, and S. K. Karapetis, "Artificial Lipid Membranes: Past, Present, and Future," *Membranes*, vol. 7, no. 3, Jul. 2017.
- [104] H. Lindsey, N. O. Petersen, and S. I. Chan, "Physicochemical characterization of 1,2-diphytanoyl-snglycero-3-phosphocholine in model membrane systems," *Biochim. Biophys. Acta BBA - Biomembr.*, vol. 555, no. 1, pp. 147–167, Jul. 1979.
- [105] S. H. White, D. C. Petersen, S. Simon, and M. Yafuso, "Formation of planar bilayer membranes from lipid monolayers. A critique," *Biophys. J.*, vol. 16, no. 5, pp. 481–489, May 1976.
- [106] J. M. del Rio Martinez, E. Zaitseva, S. Petersen, G. Baaken, and J. C. Behrends, "Automated Formation of Lipid Membrane Microarrays for Ionic Single-Molecule Sensing with Protein Nanopores," *Small*, vol. 11, no. 1, pp. 119–125, Jan. 2015.

- [107] G. Baaken, M. Sondermann, C. Schlemmer, J. Rühe, and J. C. Behrends, "Planar microelectrodecavity array for high-resolution and parallel electrical recording of membrane ionic currents," *Lab. Chip*, vol. 8, no. 6, pp. 938–944, May 2008.
- [108] G. Baaken, N. Ankri, A.-K. Schuler, J. Rühe, and J. C. Behrends, "Nanopore-Based Single-Molecule Mass Spectrometry on a Lipid Membrane Microarray," ACS Nano, vol. 5, no. 10, pp. 8080–8088, Oct. 2011.
- [109] J. Roman, B. Le Pioufle, L. Auvray, J. Pelta, and L. Bacri, "From current trace to the understanding of confined media," *Eur. Phys. J. E*, vol. 41, no. 9, p. 99, Sep. 2018.
- [110] F. J. Sigworth and S. M. Sine, "Data transformations for improved display and fitting of singlechannel dwell time histograms.," *Biophys. J.*, vol. 52, no. 6, pp. 1047–1054, Dec. 1987.
- [111] D. Colquhoun and F. J. Sigworth, "Fitting and Statistical Analysis of Single-Channel Records," in Single-Channel Recording, B. Sakmann and E. Neher, Eds. Boston, MA: Springer US, 1995, pp. 483–587.
- [112] C. T. A. Wong and M. Muthukumar, "Polymer capture by electro-osmotic flow of oppositely charged nanopores," J. Chem. Phys., vol. 126, no. 16, p. 164903, Apr. 2007.
- [113] M. Firnkes, D. Pedone, J. Knezevic, M. Döblinger, and U. Rant, "Electrically Facilitated Translocations of Proteins through Silicon Nitride Nanopores: Conjoint and Competitive Action of Diffusion, Electrophoresis, and Electroosmosis," *Nano Lett.*, vol. 10, no. 6, pp. 2162–2167, juin 2010.
- [114] V. Van Meervelt *et al.*, "Real-Time Conformational Changes and Controlled Orientation of Native Proteins Inside a Protein Nanoreactor," *J. Am. Chem. Soc.*, vol. 139, no. 51, pp. 18640–18646, Dec. 2017.
- [115] C. Cao, J. Yu, Y.-Q. Wang, Y.-L. Ying, and Y.-T. Long, "Driven Translocation of Polynucleotides Through an Aerolysin Nanopore," *Anal. Chem.*, vol. 88, no. 10, pp. 5046–5049, May 2016.
- [116] K. Knop, R. Hoogenboom, D. Fischer, and U. S. Schubert, "Poly(ethylene glycol) in Drug Delivery: Pros and Cons as Well as Potential Alternatives," *Angew. Chem. Int. Ed.*, vol. 49, no. 36, pp. 6288–6308, Aug. 2010.
- [117] L. L. Y. Lee and J. C. Lee, "Thermal stability of proteins in the presence of poly(ethylene glycols)," *Biochemistry (Mosc.)*, vol. 26, no. 24, pp. 7813–7819, Dec. 1987.
- [118] S. N. Timasheff, "The Control of Protein Stability and Association by Weak Interactions with Water: How Do Solvents Affect These Processes?," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 22, no. 1, pp. 67–97, 1993.
- [119] K. C. Ingham, "Polyethylene glycol in aqueous solution: Solvent perturbation and gel filtration studies," Arch. Biochem. Biophys., vol. 184, no. 1, pp. 59–68, Nov. 1977.
- [120] E. M. M. Del Valle, "Cyclodextrins and their uses: a review," Process Biochem., vol. 39, no. 9, pp. 1033–1046, May 2004.

- [121] L. Q. Gu and H. Bayley, "Interaction of the noncovalent molecular adapter, beta-cyclodextrin, with the staphylococcal alpha-hemolysin pore.," *Biophys. J.*, vol. 79, no. 4, pp. 1967–1975, Oct. 2000.
- [122] "Cyclodextrins." [Online]. Available: http://www1.lsbu.ac.uk/water/cyclodextrin.html. [Accessed: 11-May-2018].
- [123] Y. Marcus, "Ionic radii in aqueous solutions," Chem. Rev., vol. 88, no. 8, pp. 1475–1498, Dec. 1988.
- [124] Y. Marcus, "Thermodynamics of solvation of ions. Part 5.—Gibbs free energy of hydration at 298.15
 K," J. Chem. Soc. Faraday Trans., vol. 87, no. 18, pp. 2995–2999, Jan. 1991.
- [125] S. Varma and S. B. Rempe, "Coordination numbers of alkali metal ions in aqueous solutions," *Biophys. Chem.*, vol. 124, no. 3, pp. 192–199, Dec. 2006.
- [126] J. M\u00e4hler and I. Persson, "A Study of the Hydration of the Alkali Metal Ions in Aqueous Solution," *Inorg. Chem.*, vol. 51, no. 1, pp. 425–438, Jan. 2012.
- [127] J. N. Israelachvili, Intermolecular and Surface Forces. Academic Press, 2015.
- [128] K. Tasaki, "Poly(oxyethylene)-Water Interactions: A Molecular Dynamics Study," J. Am. Chem. Soc., vol. 118, no. 35, pp. 8459–8469, Jan. 1996.
- [129] D. Ray, D. Feller, M. B. More, E. D. Glendening, and P. B. Armentrout, "Cation–Ether Complexes in the Gas Phase: Bond Dissociation Energies and Equilibrium Structures of Li+(1,2dimethoxyethane)x, x = 1 and 2, and Li+(12-crown-4)," *J. Phys. Chem.*, vol. 100, no. 40, pp. 16116– 16125, Jan. 1996.
- [130] A. V. Skolunov, "Solubility of alkali metal chlorides and hydroxides in water," *Fibre Chem.*, vol. 25, no. 6, pp. 463–467, Nov. 1994.
- [131] P. J. Bond, A. T. Guy, A. J. Heron, H. Bayley, and S. Khalid, "Molecular Dynamics Simulations of DNA within a Nanopore: Arginine–Phosphate Tethering and a Binding/Sliding Mechanism for Translocation," *Biochemistry (Mosc.)*, vol. 50, no. 18, pp. 3777–3783, May 2011.
- [132] J.-F. Lutz, "Defining the Field of Sequence-Controlled Polymers," *Macromol. Rapid Commun.*, vol. 38, no. 24, p. 1700582.
- [133] J.-F. Lutz, J.-M. Lehn, E. W. Meijer, and K. Matyjaszewski, "From precision polymers to complex materials and systems," *Nat. Rev. Mater.*, vol. 1, no. 5, p. 16024, May 2016.
- [134] J.-F. Lutz, M. Ouchi, D. R. Liu, and M. Sawamoto, "Sequence-Controlled Polymers," *Science*, vol. 341, no. 6146, p. 1238149, Aug. 2013.
- [135] A. Al Ouahabi, L. Charles, and J.-F. Lutz, "Synthesis of Non-Natural Sequence-Encoded Polymers Using Phosphoramidite Chemistry," J. Am. Chem. Soc., vol. 137, no. 16, pp. 5629–5635, Apr. 2015.
- [136] A. Al Ouahabi, M. Kotera, L. Charles, and J.-F. Lutz, "Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100," ACS Macro Lett., vol. 4, no. 10, pp. 1077–1080, Oct. 2015.
- [137] I. Echavarri Franco, J. Combet, and F. Schosseler, "CNTb, a set of scripts for batch processing and statistical analysis of photon correlation spectroscopy data via CONTIN inversion," Sep-2011.

- [138] C. D. Putnam, M. Hammel, G. L. Hura, and J. A. Tainer, "X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution," *Q. Rev. Biophys.*, vol. 40, no. 3, pp. 191–285, Aug. 2007.
- [139] A. Guinier, "La diffraction des rayons X aux très petits angles : application à l'étude de phénomènes ultramicroscopiques," Ann. Phys., vol. 11, no. 12, pp. 161–237, 1939.
- [140] A. Y. L. Sim, J. Lipfert, D. Herschlag, and S. Doniach, "Salt dependence of the radius of gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray scattering," *Phys. Rev. E*, vol. 86, no. 2, p. 021901, Aug. 2012.
- [141] M. Muthukumar, "Double screening in polyelectrolyte solutions: Limiting laws and crossover formulas," J. Chem. Phys., vol. 105, no. 12, pp. 5183–5199, Sep. 1996.
- [142] H. Schiessel and P. Pincus, "Counterion-Condensation-Induced Collapse of Highly Charged Polyelectrolytes," *Macromolecules*, vol. 31, no. 22, pp. 7953–7959, Nov. 1998.
- [143] R. G. Winkler, M. Gold, and P. Reineker, "Collapse of Polyelectrolyte Macromolecules by Counterion Condensation and Ion Pair Formation: A Molecular Dynamics Simulation Study," *Phys. Rev. Lett.*, vol. 80, no. 17, pp. 3731–3734, Apr. 1998.
- [144] S. Liu, K. Ghosh, and M. Muthukumar, "Polyelectrolyte solutions with added salt: A simulation study," J. Chem. Phys., vol. 119, no. 3, pp. 1813–1823, Jul. 2003.
- [145] P. Loh *et al.*, "Collapse of Linear Polyelectrolyte Chains in a Poor Solvent: When Does a Collapsing Polyelectrolyte Collect its Counterions?," *Macromolecules*, vol. 41, no. 23, pp. 9352–9358, Dec. 2008.
- [146] J.-M. Y. Carrillo and A. V. Dobrynin, "Polyelectrolytes in Salt Solutions: Molecular Dynamics Simulations," *Macromolecules*, vol. 44, no. 14, pp. 5798–5816, Jul. 2011.
- [147] A. H. Laszlo *et al.*, "Decoding long nanopore sequencing reads of natural DNA," *Nat. Biotechnol.*, vol. 32, no. 8, pp. 829–833, Aug. 2014.
- [148] B. Jeon and M. Muthukumar, "Determination of Molecular Weights in Polyelectrolyte Mixtures Using Polymer Translocation through a Protein Nanopore," ACS Macro Lett., vol. 3, no. 9, pp. 911– 915, Sep. 2014.
- [149] R. M. A. Manara, A. T. Guy, E. J. Wallace, and S. Khalid, "Free-Energy Calculations Reveal the Subtle Differences in the Interactions of DNA Bases with α-Hemolysin," J. Chem. Theory Comput., vol. 11, no. 2, pp. 810–816, Feb. 2015.
- [150] N. F. König, A. Al Ouahabi, S. Poyer, L. Charles, and J.-F. Lutz, "A Simple Post-Polymerization Modification Method for Controlling Side-Chain Information in Digital Polymers," *Angew. Chem. Int. Ed.*, vol. 56, no. 25, pp. 7297–7301, May 2017.
- [151] C. R. Martinez and B. L. Iverson, "Rethinking the term 'pi-stacking," Chem. Sci., vol. 3, no. 7, pp. 2191–2201, Jun. 2012.
- [152] V. Ramakrishnan, "Histone Structure and the Organization of the Nucleosome," Annu. Rev. Biophys. Biomol. Struct., vol. 26, no. 1, pp. 83–112, 1997.

- [153] L. Mariño-Ramírez, M. G. Kann, B. A. Shoemaker, and D. Landsman, "Histone structure and nucleosome stability," *Expert Rev. Proteomics*, vol. 2, no. 5, pp. 719–729, Oct. 2005.
- [154] N. M. Luscombe, S. E. Austin, H. M. Berman, and J. M. Thornton, "An overview of the structures of protein-DNA complexes," *Genome Biol.*, vol. 1, no. 1, p. reviews001.1, Jun. 2000.
- [155] W. H. Hudson and E. A. Ortlund, "The structure, function and evolution of proteins that bind DNA and RNA," *Nat. Rev. Mol. Cell Biol.*, vol. 15, no. 11, pp. 749–760, Nov. 2014.
- [156] U. F. Keyser, "Controlling molecular transport through nanopores," J. R. Soc. Interface, vol. 8, no. 63, pp. 1369–1378, Oct. 2011.
- [157] H. Chen, S. P. Meisburger, S. A. Pabit, J. L. Sutton, W. W. Webb, and L. Pollack, "Ionic strengthdependent persistence lengths of single-stranded RNA and DNA," *Proc. Natl. Acad. Sci.*, vol. 109, no. 3, pp. 799–804, Jan. 2012.
- [158] M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, and T. Ha, "Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy," *Biophys. J.*, vol. 86, no. 4, pp. 2530– 2537, Apr. 2004.

List of publications

- M. Boukhet, F. Piguet, H. Ouldali, M. Pastoriza-Gallego, J. Pelta, and A. Oukhaled, "Probing driving forces in aerolysin and α-hemolysin biological nanopores: electrophoresis versus electroosmosis," *Nanoscale*, vol. 8, no. 43, pp. 18352–18359, Nov. 2016.
- [2] M. Boukhet, N. F. König, A. A. Ouahabi, G. Baaken, J.-F. Lutz, and J. C. Behrends, "Translocation of Precision Polymers through Biological Nanopores," *Macromol. Rapid Commun.*, vol. 38, no. 24, p. 1700680, Dec. 2017.

Acknowledgments

I would like to thank all the people which have helped me throughout my work.

I would first like to thank Pr. Jan Behrends and Pr. Juan Pelta for giving me the opportunity to do this PhD and for helping me in my work throughout these 4 years.

During my first year of PhD in France, we had many interesting discussions on my results with Pr. Pelta which were always enlightening for the progression of the project. His helpful comments and ideas always allowed me to progress in my work and for that I greatly thank him.

During the next three years that I spent in Freiburg with Pr. Behrends, I have benefited a lot from his guidance and knowledge. Throughout our discussions I was able to overcome the challenges faced in my research and I have learned a lot from him. His comments were always helpful for the progression of the project and his rigor always pushed me to better myself in my work. For all of that I thank him deeply.

I would also like to thank Dr. Abdelghani Oukhaled who is a co-supervisor of this thesis and has initiated me to the experimentation of lipid bilayers and taught me most of the technical part of this work. He was always there for the questions I had and it was a great pleasure to work with him.

I thank Pr. Günter Reiter for accepting to be my supervisor in the faculty of physics and for all his help.

I would like to thank Dr. Gerhard Baaken from Ionera Technologies, for all the useful discussions we had about work, for his continual support and help throughout my thesis, especially in tougher times, for which he was always present.

I would also like to thank Prof. Jean-Francois Lutz, Prof, Francois Schosseler, Prof. Michel Rawiso, for welcoming me to their groups during my stay in the ICS in Strasbourg and for all their scientific guidance and help on the research carried on the characterization of precision polymers.

I thank all the people from the Eurosequences network and from the IRTG for all the interesting meetings and workshops we had togeteher, and a special thank you to Jana Husse for
all the help, and kindness and Birgitta Zovko, for her kindness, her great help and support and whom constant smile can always brighten your day.

Many thanks to the people in both groups in France and Germany, with whom I shared interesting scientific discussions but also for the nice atmosphere and welcome they offered in the groups. A special thanks to Dr. Manuela Pastoriza-Gallego for teaching me protein purification and for all the useful discussions about nanopores and their structure, but also for all the good moments shared in our office in Cergy. I also thank Noémie Ein-Eli and Mona Talari, for the nice times spent together and all the nice coffee breaks much needed between experiments. I thank Dr. Ekaterina Zaitseva and Dr. Sönke Petersen for the nice atmosphere and nice laughs in the office in Freiburg. I also would like to thank Dr. Daniel Steinbrenner and Tobias Ensslen for their help in the translation of the abstract in German. And in general I thank all the people from LAMBE or the institute of Physiology with whom I shared good times during my PhD.

I would like to thank all my family and friends for their support during my work, especially Emilie Gontran for her many useful comments in first drafts of this thesis and all her moral support and help, Katrina Chan for her help in the English writing and all my friends in general who were there during these four years at all times. I would like to especially thank Joëlle Mercier for all her support during my studies and all her help throughout these years always with her bright smile, and my aunt Fatiha Bendali-Amor, for her warmness towards me, who was always there for the good and difficult times.

I finally deeply thank my parents, which always supported me and have pushed me to be my best-self. I thank them for teaching me so many things, for giving me a curious mind and being part of giving me the interest in science. I wouldn't have been able to reach my goals and overcome all the adversities without their sacrifices. I thank them for always being there and for all their advices which always helped me in my life.