



Nanopore pour la détection de peptides vers le séquençage de protéines à l'échelle de la molécule unique.

Hadjer Ouldali

► To cite this version:

Hadjer Ouldali. Nanopore pour la détection de peptides vers le séquençage de protéines à l'échelle de la molécule unique.. Biologie moléculaire. CY Cergy Paris Université, 2021. Français. NNT : 2021CYUN1024 . tel-03635875

HAL Id: tel-03635875

<https://theses.hal.science/tel-03635875>

Submitted on 8 Apr 2022

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.



Thèse de Doctorat

présentée par **Hadjer OULDALI**

pour l'obtention du grade de **Docteur en Sciences de la Vie et de la Santé**

délivré par **CY Cergy Paris Université**

Préparée au sein de l'Ecole doctorale n°417 Sciences et Ingénierie et du Laboratoire
Analyse, Modélisation, et Matériaux pour la Biologie et l'Environnement (UMR 8587)

Nanopore pour la détection de peptides vers le séquençage de protéines à l'échelle de la molécule unique

Soutenue publiquement le 11 Février 2021 devant le jury composé de :

M. Emmanuel PAUTHE, CY Cergy Paris Université	Président du jury
M. Jean-Michel BETTON, Institut Pasteur de Paris	Rapporteur
M. Mathias WINTERHALTER, Jacobs University Bremen	Rapporteur
M. Yves GAUDIN, Institut de Biologie Intégrative de la Cellule	Examineur
Mme. Patricia BASSEREAU, Institut Curie de Paris	Examineur
M. Juan PELTA, Université d'Evry Val d'Essonne	Co-directeur de thèse
M. Abdelghani OUKHALED, CY Cergy Paris Université	Directeur de thèse

Remerciements

La thèse que j'ai réalisée porte sur l'utilisation de nanopores biologiques pour le développement d'un séquençage de protéines de nouvelle génération à l'échelle de la molécule unique.

Je tiens à remercier l'ensemble des membres du jury d'avoir accepté d'évaluer mes travaux de thèse : le **professeur Jean-Michel Betton** (Institut Pasteur, Paris), le **professeur Mathias Winterhalter** (Jacobs University, Bremen), le **professeur Yves Gaudin** (Institut de Biologie Intégrative de la Cellule, Paris), le **professeur Patricia Bassereau** (Institut Curie, Paris), et le **professeur Emmanuel Pauthe** (CY Cergy-Paris Université).

Je remercie également les organismes financeurs d'avoir soutenu mon projet et d'avoir permis la réalisation de cette thèse : la **Direction Générale de l'Armement** et plus particulièrement **Rose-Marie Sauvage** qui a été ma tutrice durant ce projet, ainsi que la **région Ile de France** au nom du **Domaine d'Intérêt Majeur « Réseau d'Excellence en Solides Poreux »**.

Ce projet de thèse est né à la suite de ma rencontre avec le **professeur Abdelghani Oukhaled** alors que j'étais étudiante en Licence de Biologie. Il a su me transmettre, avec une très grande pédagogie, sa passion pour la physico-chimie et le comportement des polymères en milieux confinés. Une belle aventure scientifique commençait.

En cela, je tiens tout d'abord à remercier mon directeur de thèse **Abdelghani Oukhaled**.

Ghani, je te remercie de m'avoir ouvert les portes de ton laboratoire et du monde des nanopores que j'affectionne tant. Je te remercie de m'avoir accordé ta confiance sur un sujet de recherche qui, je le sais, est si important pour toi. Ce fut un immense honneur d'avoir travaillé à tes côtés durant ces années. Je te remercie pour ta grande pédagogie et ta passion que tu sais transmettre si naturellement. L'art d'expliquer simplement des phénomènes complexes. Je te remercie d'avoir cru en moi, et de m'avoir rapidement confié des responsabilités qui m'ont fait grandir. Je te remercie également pour ta présence tout au long de ma thèse, ton soutien, et ta patience. Une seule chose pour laquelle je n'aurai pas grandi : je continuerais à cacher tes clés qui traînent partout si c'était à refaire et à rigoler en te regardant courir partout dans le labo, pressé, car comme on t'a beaucoup entendu dire avec Fabien : « je démarre ».

Ce projet m'a fait vivre pleinement l'intensité de la recherche. Les premiers résultats expérimentaux que j'ai savourés et que je n'oublie pas, la réflexion, les discussions menant à la découverte de phénomènes incroyables qui finalement nous permettront de publier de brillants travaux. Merci de m'avoir accompagné dans cette belle aventure du doctorat qui m'a transformé scientifiquement, intellectuellement, humainement.

Je tiens à remercier également mon co-directeur de thèse, le **professeur Juan Pelta**. Juan, je te remercie de m'avoir accordé ta confiance sur ce projet. Je te remercie pour les échanges scientifiques, et plus particulièrement de m'avoir challengé sur la préparation de mon manuscrit de thèse et de ma soutenance avec toujours plus de références bibliographiques.

Je voudrais remercier mon collègue **Fabien Piguet** pour son soutien sans faille depuis mon premier jour au laboratoire. Fabien, je te remercie de m'avoir formé et d'avoir toujours répondu à mes questions. Tu as toujours pris de ton temps pour m'aider à faire avancer ce projet et je t'en suis reconnaissante. Un grand merci également pour ta patience, ton sens de la pédagogie, ta rigueur, et ton organisation qui nous ont été indispensables durant ce projet. Ce fut un grand honneur d'avoir travaillé à tes côtés durant ces années. Je te remercie également pour ta bonne humeur et les discussions que nous avons eu autour d'un nanopore « immortel » (peut-être même un superpore) avec lequel une manip peut durer une éternité, d'un thé, et plus récemment pour toi, d'un café (c'était long mais tu as fini par tomber dedans).

Je voudrais remercier l'ensemble de mes collègues du LAMBE, plus particulièrement **Manuela Pastoriza-Gallego**, **Bénédicte Thiebot**, **Françoise Discala**, **Marie-France Breton**, et **Mazdak Afshar**. Je vous remercie de m'avoir encouragé durant cette aventure et de m'avoir toujours ouvert vos bureaux, pour une question, un coup de main, un soutien.

Manuela, je te remercie d'avoir toujours répondu à mes questions, de m'avoir accompagné tout au long du projet notamment avec la production de l'aérollysine, et plus particulièrement pour la relecture entière de ce manuscrit qui m'a été d'une grande aide.

Bénédicte, je te remercie pour ta bonne humeur et ta générosité. J'ai beaucoup apprécié partager ses petites pauses goûter autour d'une infusion menthe verveine, citronnelle, ou pomme ; des petits chocolats qui arrivaient à temps quand mon cerveau saturait durant l'analyse de données ; ou encore les excellentes figues de ton jardin. C'était un plaisir de t'avoir comme voisine de bureau.

Françoise, je te remercie de m'avoir soutenu dans un des moments les plus difficiles de ma thèse. Tu as su trouver les mots réconfortants me permettant de poursuivre mes efforts. Je te remercie également d'avoir su mettre en avant le côté humain que l'on ne doit pas oublier.

Je remercie le **professeur Jean-Yves Salpin**, directeur du laboratoire LAMBE, de m'avoir accueilli dans sa structure et d'avoir suivi l'évolution de mon projet de thèse.

Je souhaite également remercier chaleureusement **Evely Sachou** que j'ai encadré comme stagiaire au LAMBE et qui m'a apporté son aide sur le projet. Je te remercie pour ton investissement, ton enthousiasme, et ta générosité.

Une pensée et un chaleureux merci à l'équipe de la prépa TP, **Karine**, **Caroline**, **Asma**, **Nicolas** et **Sabine**. Je vous remercie pour vos encouragements et votre bonne humeur qui fait du bien. C'était toujours un plaisir de vous croiser dans les couloirs et de partager la pause déjeuner avec vous.

J'adresse un remerciement particulier aux professeurs **Emmanuel Pauthe** et **Véronique Larreta-Garde** qui ont été mes responsables de formation du Cours de Master en Ingénierie Biomatériaux pour la Santé durant mes cinq premières années d'étude. Emmanuel et

Véronique, je vous remercie très sincèrement de m'avoir permis d'intégrer cette formation qui a été un véritable moteur dans mon parcours et de m'avoir permis de mettre les pieds très tôt dans les laboratoires pour y faire mes preuves.

Je tiens à remercier mes chères amies **Amel, Mylène, et Alexandra**.

Amel, ma complice de toujours, mon binôme, le duo infernal. La première fois que je t'ai vu, jamais je n'aurais pensé que l'on vivrait autant d'étapes de nos parcours ensemble. J'ai retrouvé chez toi toute cette motivation, détermination, patience, et rigueur que je n'ai su trouver chez quelqu'un d'autre. Pas besoin de parler, un regard suffit pour se comprendre. Les projets à tes côtés ont toujours été un succès et de bons souvenirs malgré tout ce stress que l'on sait si bien se mettre mais qui a toujours été moteur. Toujours sur la même longueur d'onde. Je suis fière d'avoir parcouru ce chemin à tes côtés, de t'avoir vu progresser, devenir brillante, et exceller. Je te remercie d'avoir toujours été présente dans les moments difficiles mais aussi dans les moments de joie. Ma sœur, je te souhaite le meilleur pour la suite car tu le mérites.

Mylène, nous nous sommes rencontrées lors de nos premiers jours à l'université. Une belle amitié a commencé et nous avons grandi ensemble durant toutes ces années. Je suis ravie d'avoir fait ce chemin à tes côtés et de t'avoir vu réussir et t'épanouir dans une voie que tu aimes. Je te remercie de m'avoir soutenu, supporté, et de m'avoir toujours encouragé à aller de l'avant.

Alexandra, je te remercie pour ton soutien durant ce parcours, ta bonne humeur, et les bons moments passés ensemble. Je garde un bon souvenir de nos week-ends d'étudiantes !

Je voudrais terminer en remerciant les personnes les plus importantes dans ma vie, celles qui m'ont aidé à me construire, qui m'ont transmis des valeurs fortes, et m'ont apporté un soutien sans faille sur tous les plans. Sans vous, rien de tout cela n'aurait été possible.

Tout d'abord, **Papa & Maman**, je vous remercie pour tout ce que vous m'avez apporté depuis petite, l'exceptionnelle éducation que vous m'avez donnée, et les belles valeurs de travail, de respect, de résilience, et de détermination que vous n'avez cessé de me transmettre depuis toujours. Vous m'avez toujours appris que tout était possible avec du travail, de toujours garder espoir en la vie, et de se battre pour obtenir ce que l'on désire. Je vous remercie pour tous les sacrifices que vous avez fait pour moi, pour m'aider à construire un projet de vie, à m'épanouir, à devenir une belle personne. Je vous remercie de m'avoir donné tous ce dont j'avais besoin pour parvenir à ces réussites, votre amour avant tout. Vous avez su passer au-dessus des moments difficiles que j'ai traversé. Personne n'aura jamais autant compris, accepté, et supporté avec autant de calme et de réconfort, le stress et la fatigue que j'ai vécu,. Tout cela n'aurait jamais été possible sans vous. Je ne vous remercierai jamais assez pour tout.

Merzak & Nesrine, mon frère et ma sœur, mes meilleurs amis. Je vous remercie de m'avoir accompagné, d'avoir toujours su accepter mon caractère particulier, de m'avoir soutenu dans les moments difficiles. Vous savez toujours me faire relativiser et me faire retrouver le sourire.

Une discussion, un moment passé ensemble, et tout va mieux. Mes plus beaux souvenirs d'enfance sont avec vous. Je suis tellement reconnaissante de vous avoir dans ma vie et je ne vous remercierai jamais assez de m'apporter autant de joie.

Je tiens à remercier également ma merveilleuse belle famille, mes beaux-parents **tatie Khawlah & tonton Marwan**, et mes beaux-frères **Rami & Feras**. Je n'aurais pu rêver avoir une aussi magnifique belle famille. Je vous remercie de m'avoir acceptée et accueillie à bras ouverts dans votre merveilleuse famille si aimante, généreuse, et bienveillante. Vous m'avez toujours soutenue, encouragée durant cette aventure. Vous avez pris soin de moi quand j'en avez besoin, je ne l'oublierai jamais. Votre bonne humeur et votre joie de vivre est une bouffée d'air frais.

Je tiens également à remercier les autres membres de la **famille Hasan**, oncles, tantes, cousins, et cousines, qui m'ont encouragé durant cette aventure. J'adresse un remerciement particulier au **docteur Oussama** pour avoir pris le temps de relire l'anglais de ce manuscrit et pour ses conseils pertinents.

Enfin, je tiens à remercier mon époux, **Motaz**.

Si ce dernier paragraphe t'est dédié, ce n'est pas anodin car tu es tout pour moi. Mon cher mari, je ne te remercierai jamais assez pour tout le soutien, l'amour, le courage, le réconfort que tu me donnes depuis notre rencontre. Cette rencontre qui, d'ailleurs, s'est faite alors que j'allais commencer ma thèse. Qui de meilleur, et de plus courageux que toi pour m'accompagner dans cette aventure d'une vie, et tu l'as si bien fait chaque jour. Je pense toujours à comment se serait dérouler cette étape de ma vie sans toi. A vrai dire, je ne parviens pas à l'imaginer tant ta présence a été d'une importance inestimable à chaque instant. Je te remercie de m'avoir toujours soutenue, de m'avoir supportée dans les moments les plus difficiles, et de m'avoir toujours poussée vers le haut, pour faire mieux, pour être quelqu'un de meilleur, et de m'avoir également toujours donné le meilleur. Je te remercie d'avoir toujours su garder ton optimisme que j'aime, de m'avoir appris à relativiser et à retrouver le sourire, et de m'avoir transmis ta joie de vivre. Je te remercie de m'avoir laissé manger autant de gourmandises que le souhaitais notamment durant l'écriture de ce manuscrit, j'en avais besoin. Enfin, je te remercie également de m'avoir accompagné dans la lecture, relecture, et correction de ce manuscrit, à la virgule près. La fierté que je vois dans tes yeux me comble de joie. Je pourrais écrire encore longtemps mais tout le reste tu le sais déjà ...

Résumé

Les protéines sont des macromolécules indispensables au fonctionnement des cellules vivantes, pouvant fournir des informations clés pour comprendre les mécanismes biologiques et les maladies. Le développement de nouveaux outils pour le séquençage des protéines, rapide et à l'échelle de la molécule unique, constituerait un apport dans le diagnostic de nombreuses maladies et dans le développement de la médecine personnalisée.

La technologie de détection électrique par nanopore est une méthode d'analyse à l'échelle de la molécule unique, en temps réel, sans marquage, peu coûteuse, qui se présente comme un outil prometteur pour le développement de cette application. Le principe de cette technique consiste à utiliser une membrane isolant deux compartiments remplis par une solution ionique, et percée d'un trou nanométrique, un nanopore. Sous l'action d'une différence de potentiel, un flux d'ions circule à travers le nanopore. L'entrée d'une molécule individuelle dans le nanopore entraîne une chute du courant ionique dont les caractéristiques permettent d'obtenir des informations sur la nature chimique de la molécule analysée. Sur ce principe, le séquençage de biomolécules peut être réalisé par la lecture séquentielle de l'information d'une chaîne dépliée à travers un nanopore.

Les avancées spectaculaires du séquençage de l'ADN par nanopore ont mis en lumière le haut potentiel de cette technologie et ont ouvert la voie vers le développement du séquençage des protéines par nanopore. Ce nouveau challenge est néanmoins bien plus difficile que pour les acides nucléiques en raison de nombreuses contraintes principalement liées à la nature des protéines qui présentent une structure complexe, qui peuvent être composées de 20 acides aminés, et qui présentent des modifications chimiques de certains de leurs acides aminés. A notre connaissance, au commencement de mes travaux de thèse, aucun nanopore, qu'il soit biologique non modifié, solide, ou artificiel, n'avait permis d'obtenir une résolution d'un acide aminé suffisante à la lecture de la séquence d'une protéine.

L'ambition de mon projet de thèse est de développer une stratégie de séquençage des protéines basée sur l'utilisation d'un nanopore biologique, l'aérolysine, une toxine produite par la bactérie *Aeromonas hydrophila*, capable de former des pores nanométriques dans les membranes biologiques. L'aérolysine, a récemment démontré un haut pouvoir de résolution, dans des conditions physico-chimiques optimales, pour l'analyse de différents types de molécules, comme des polymères neutres et des oligonucléotides. L'aérolysine se présente ainsi comme un candidat de choix pour tenter d'atteindre une résolution de détection d'un acide aminé.

Ce manuscrit a pour objectif de présenter les résultats que j'ai obtenus durant ma thèse. La première contribution majeure de mes travaux est de montrer, pour la première fois, que l'aérolysine permet d'obtenir la discrimination de courts peptides (de 5 à 10 acides aminés), analysés dans un mélange ou individuellement, avec une résolution de séparation d'un acide aminé individuel. La deuxième contribution majeure de mes travaux est l'établissement de la première cartographie des 20 acides aminés naturels par nanopore, sur la base de leurs signatures électriques mesurées à travers l'aérolysine. Nous avons également montré qu'il est possible d'identifier des modifications chimiques portées par certains acides aminés.

Abstract

Proteins are large molecules that play essential roles in living cells including structure, regulation and functions. They provide key information for understanding biological mechanisms and diseases. The development of new tools for rapid protein sequencing at the single molecule scale, would contribute to the diagnosis of numerous diseases and to the development of personalized medical assessment.

Nanopore-based single-molecule electrical detection technology is a low-cost, real-time, and label-free analysis method, that stands out as a promising tool for the development of this application. The principle of this technique consists of using a nanometric pore (nanopore) created in a membrane separating two compartments filled with electrolyte solution. Due to the applied voltage, the electrolyte ions flow through the nanopore. The entry of individual molecules in the nanopore causes a drop in ionic current whose characteristics provide information on the chemical nature of the molecule being analyzed. Based on this principle, biomolecule sequencing can be achieved by sequentially reading the information of an unfolded protein through a nanopore.

The significant advances in nanopore DNA sequencing have enhanced the potential of this technology and expanded the development of nanopore protein sequencing. This latter is, nonetheless, more challenging compared to nucleic acids because of many constraints mainly related to the nature of proteins which have a complex structure. It can be composed of 20 amino acids which may also involve chemical modifications of some of them. To our knowledge, when my thesis work has been initiated, no nanopore, whether biological unmodified, solid, or artificial, has so far made it possible to obtain sufficient amino acid resolution for reading the sequence of a protein.

The foreseen scope of my thesis work is to develop a protein sequencing strategy based on the use of a biological nanopore called aerolysin, a toxin produced by the bacteria *Aeromonas hydrophila* which forms nanometric pores in biological membranes. Aerolysin has recently demonstrated a high resolving power, under optimal physicochemical conditions, for the analysis of different types of molecules such as neutral polymers, and oligonucleotides. Aerolysin thus represents an ideal candidate for attempting to achieve a high separation resolution permitting the detection of an amino acid.

This manuscript explores the results that I obtained during my thesis work. My first major contribution is to demonstrate, for the first time, that aerolysin allows the discrimination of short peptides (from 5 to 10 amino acids) analyzed in a mixture or individually, with a separation resolution equivalent to that of an individual amino acid. My second major contribution is to demonstrate for the first time a mapping of 20 natural amino acids per nanopore, based on their electrical signatures measured through aerolysin. We have also shown that it is possible to identify chemical modifications carried by certain amino acids.

Table des matières

Introduction générale.....	1
Chapitre 1 : Revue bibliographique	7
I. La détection électrique à l'échelle de la molécule unique par nanopore	7
1.1 Principe et origine de la technique.....	7
1.2 Catégories de nanopores	9
1.2.1 Les nanopores biologiques	10
a) Les canaux ioniques	10
b) Les peptides antimicrobiens.....	11
c) Les toxines qui forment des pores.....	12
d) Les porines	15
e) Viral portal protein.....	16
1.2.2 Autres catégories de nanopores	19
a) Les nanopores solides	19
b) Les nanopores hybrides.....	20
1.3 Expériences pionnières et applications.....	22
1.3.1 Caractériser des pores avec des polymères	23
1.3.2 Caractériser des polymères avec des pores	23
a) Spectrométrie de taille par nanopore.....	23
b) Séquençage des acides nucléiques par nanopore	26
1.4 Résumé	38
II. Progrès vers le séquençage de protéines par nanopore	41
2.1 Séquençage des protéines par nanopore	42
2.1.1 Contrôle du transport unidirectionnel	43
2.1.2 Contrôle du dépliement et de la translocation dans un nanopore.....	45
2.1.3 Détection et discrimination des acides aminés.....	46
2.1.4 Détection des modifications post-traductionnelles	52
2.2 Autres approches de séquençage de protéines à l'échelle de la molécule unique...53	
2.2.1 Séquençage des protéines par courant tunnel.....	53
2.2.2 Séquençage des protéines par fluorescence.....	55
a) FRET fingerprinting.....	55

b) Fluoroséquençage	57
2.3 Résumé	59
III. Le nanopore biologique d'aérolysine	61
3.1 Structure de la pro-aérolysine.....	62
3.2 Mécanisme d'insertion du pore.....	65
3.3 Structure et propriétés électriques du canal.....	67
3.4 Détection de molécules par le nanopore d'aérolysine	69
3.5 Résumé	77
Chapitre 2 : Méthode expérimentale.....	79
I. Dispositifs expérimentaux de mesure électrique.....	79
1.1 Dispositif expérimental classique.....	79
1.1.1 Cellule de mesure	79
1.1.2 Electrodes de mesure et chaîne d'acquisition du signal.....	80
1.2 Dispositif expérimental haut-débit (Orbit 16, MECA chips)	83
II. Membranes de lipides	86
2.1 Composition chimique et organisation des bicouches de lipides.....	86
2.2 Méthodes de reconstitution des bicouches de lipides	88
2.3 Caractérisation électrique des bicouches de lipides	90
III. Aérolysine recombinante et insertion d'un nanopore dans une membrane	91
3.1 Production et purification de l'aérolysine.....	91
3.2 Insertion et propriétés électriques du canal.....	91
IV. Détection électrique	93
4.1 Résistance électrique d'un pore cylindrique en présence d'une particule sphérique	93
4.2 Grandeurs mesurables.....	96
4.3 Analyse de données.....	97
4.3.1 Histogramme tout point, méthode des seuils, et détection	97

4.3.2 Caractérisation de l'amplitude ou du courant relatif résiduel	99
4.3.3 Caractérisation des durées de blocage du courant.....	102
4.3.4 Caractérisation des fréquences de blocage du courant.....	102
4.4 Expériences	103
4.4.1 Conditions expérimentales	103
4.4.2 Acquisition du signal et analyse de données	104
4.4.3 Analytes.....	104
Chapitre 3 : Discrimination d'homopeptides avec une résolution d'un acide aminé par le nanopore d'aérolysine.....	107
I. Introduction.....	107
II. Résultats & Discussion.....	108
2.1 Analyse d'une courte chaîne peptidique issue d'un échantillon monodisperse....	108
2.2 Détection et identification d'impuretés peptidiques	113
2.3 Détection et discrimination en taille des peptides avec une résolution d'un acide aminé	117
2.4 Effet de la tension électrique sur la discrimination en taille des peptides	120
2.5 Extension de la limite de détection à des peptides d'une longueur de 4 acides aminés.....	123
2.6 Application à un autre type d'acide aminé	126
2.7 Discrimination de peptides selon la nature chimique des acides aminés	129
III. Discussion & Conclusion.....	132
Chapitre 4 : Signature électrique des 20 acides aminés protéinogènes par le nanopore d'aérolysine	135
I. Introduction.....	135
II. Résultats & discussion	136
2.1 Détection des 20 acides aminés protéinogènes.....	136
2.2 Discrimination de modifications chimiques subtiles.....	142
2.3 Discrimination des 20 acides aminés protéinogènes dans un mélange.....	142

2.4 Dépendance en masse et en volume	147
2.5 Simulation	149
2.6 Pistes d'améliorations pour une stratégie de séquençage.....	154
2.6.1 Augmentation des temps de résidence	154
2.6.2 Introduction de modifications chimiques.....	155
2.6.3 Variation des conditions expérimentales et d'enregistrement électriques.....	157
III. Discussion & Conclusion.....	159
Conclusion générale	163
Références	167
Glossaire des abréviations	193
Table des figures.....	197
Annexes	203

Introduction générale

Les cellules, qui sont les plus petites unités structurales et fonctionnelles des organismes vivants, sont structurées en sous-compartiments et sont séparées de leur environnement par des membranes lipidiques. Le transport des molécules ou des macromolécules à l'échelle cellulaire, de part et d'autre des membranes lipidiques, se fait à travers des canaux protéiques qui sont des pores de diamètre nanométrique (nanopores) enchâssés dans ces membranes. On peut citer la synthèse des biopolymères (ADN, ARN, protéines, et polysaccharides), l'export d'ARN à travers des pores nucléaires, l'infection de cellules par des bactéries et des virus, ou encore la translocation des protéines. L'apparition à la fin des années 80 d'une méthode de détection électrique couplée à un canal protéique nanométrique inséré dans une bicouche lipidique séparant deux compartiments chacun rempli d'une solution ionique [1], a rendu possible l'étude de ces processus biologiques complexes à l'échelle moléculaire. Cette technologie est connue sous le nom de détection électrique par nanopore à l'échelle de la molécule unique.

Le principe général de cette technique repose sur la mesure de chutes du courant ionique, engendrées par la présence d'une molécule individuelle entraînée sous l'action d'un champ électrique dans un nanopore. Les caractéristiques du blocage du courant ionique induit par la présence de la molécule dans le nanopore dépendent, en particulier, de la conformation, de la taille, de la composition ou de la séquence de la molécule, de la nature du canal, et de l'interaction potentielle entre la molécule et le canal.

Elle a permis depuis son apparition la détection électrique d'un large éventail d'analytes [2] comme des ions, des polymères synthétiques, des sucres, des acides nucléiques, ou encore des peptides et des protéines ; et a mené au développement de plusieurs applications dont la plus connue d'entre elles est le séquençage de l'ADN par nanopore, qui est aujourd'hui acquis et commercialisé par Oxford Nanopore Technologies [2–4].

Le succès du séquençage de l'ADN par nanopore a ouvert la voie au développement du séquençage de protéines par nanopore, qui est devenu le nouveau challenge [5–10].

Les protéines sont des macromolécules indispensables au fonctionnement des cellules vivantes. Des milliers de protéines assurent l'ensemble des fonctions cellulaires, du contrôle des mouvements de la cellule, à la catalyse du métabolisme, en passant par la copie de l'ADN. Certaines protéines sont également impliquées dans des processus pathologiques. Les protéines

peuvent ainsi fournir des informations clés pour comprendre les mécanismes biologiques et les maladies. Un des objectifs importants de la protéomique consiste à découvrir et à identifier des biomarqueurs, comme des protéines ou des peptides, et/ou des modifications des protéines ou des peptides, dont la présence, l'absence, ou l'altération, est liée à un état physiologique ou pathologique particulier. Le séquençage des protéines, qui consiste à caractériser la structure primaire d'une protéine, c'est-à-dire à déterminer le nombre, la nature chimique, et l'ordre de l'ensemble des acides aminés la composant, peut permettre de découvrir et d'identifier ces biomarqueurs.

La spectrométrie de masse [11], qui est actuellement la technique de référence pour le séquençage des protéines [12,13], est un outil puissant mais qui demeure principalement limité par sa sensibilité, par les étapes complexes de préparation des échantillons, et par l'utilisation de bases de données incomplètes [9]. En l'occurrence, la spectrométrie de masse n'est pas encore suffisamment sensible pour permettre la détection de très faibles quantités de protéines présentes dans une cellule unique, une vésicule extracellulaire (exosome), ou encore moins, à l'état de quelques copies dans un fluide biologique ; ce qui peut être d'une importance cruciale dans un contexte de diagnostic médical.

Le défi serait de parvenir à séquencer des protéines entières par lecture directe de leur composition en acides aminés, sans étapes préalables complexes de préparation d'échantillon, avec un outil disposant d'un coût abordable, d'une haute sensibilité, d'un haut débit, et d'une précision suffisamment importante pour déterminer à la fois la structure primaire et les modifications post-traductionnelles, sans faire appel à l'utilisation de bases de données [9,14].

La détection électrique par nanopore est une technique d'analyse en temps réel, à l'échelle de la molécule unique, sans marquage, et bon marché, qui pourrait répondre à ces exigences.

L'intérêt de l'analyse d'une molécule unique réside dans la précision de la mesure. En effet, les techniques en volume permettent d'accéder à des informations moléculaires impliquant un nombre macroscopique de molécules et donnant des informations moyennées, donc relativement limitées. La possibilité de détecter des molécules individuelles une par une permet de bénéficier d'une haute sensibilité, -sélectivité, et -spécificité. De plus, la possibilité de détecter des molécules à partir d'un signal électrique permet, en comparaison à d'autres approches d'analyse à l'échelle de la molécule unique [6,15] de s'affranchir d'étapes complexes, de préparation de l'échantillon par marquage fluorescent [16–19], ou de fonctionnalisation de l'outil de mesure [20].

Le séquençage d'une protéine par un nanopore peut être réalisé par la lecture séquentielle des différents acides aminés d'une protéine dépliée, traversant de façon unidirectionnelle un nanopore qui joue alors le rôle d'une tête de lecture. Le défi est néanmoins de taille en raison des nombreuses contraintes principalement liées à la structure des protéines, mais également au manque de résolution des nanopores utilisés jusque-là. D'une part, la distribution de charge au sein des différents acides aminés, au contraire de l'ADN qui est uniformément chargé, rend plus difficile le transport unidirectionnel d'une protéine native [21,22] à travers le nanopore. D'autre part, la réalisation du séquençage de protéines nécessite de pouvoir distinguer au moins les 20 acides aminés naturels pouvant composer une protéine, et de détecter la présence de modifications post-traductionnelles. La première étape vers le séquençage de protéines par nanopore est d'atteindre une précision de lecture d'un acide aminé. Jusque-là, aucun nanopore, qu'il soit biologique non modifié, ou artificiel, ni aucune autre méthode à l'échelle de la molécule unique, n'a permis d'obtenir une résolution suffisante à la lecture de la séquence d'une protéine avec une précision d'un acide aminé.

C'est dans ce contexte que s'inscrivent mes travaux de thèse dont l'objectif est de définir une stratégie de séquençage des protéines basée sur l'utilisation d'un nanopore stable et facilement modulable permettant d'atteindre une résolution à l'échelle d'un acide aminé. Pour relever ce défi, notre choix s'est porté sur l'utilisation d'un nanopore biologique particulier, l'aérolysine, une toxine utilisée par la bactérie *Aeromonas hydrophila* pour former des pores dans les membranes biologiques de ses cibles [23], et qui est produite dans notre laboratoire [24].

L'aérolysine a préalablement été utilisée pour l'étude, du dépliement de protéines sauvages et mutantes, de la dynamique du transport de polymères, de protéines dépliées, de peptides, de sucres, d'oligonucléotides, du suivi en temps-réel d'une réaction de dégradation enzymatique, et de la caractérisation de structures secondaires [25]. En revanche, l'aérolysine n'avait été que peu utilisée pour la détermination de la séquence de protéines à l'échelle de la molécule unique au commencement de ma thèse.

Récemment, il a plus particulièrement été démontré que l'aérolysine, en comparaison à d'autres nanopores biologiques, possède un pouvoir de résolution supérieur pour l'analyse de différents types de molécules, notamment pour la détection et la discrimination en taille avec une résolution de séparation d'une seule unité monomérique au sein d'un mélange polydisperse de polymères synthétiques [26], et au sein d'un mélange polydisperse d'oligonucléotides [27]. Ce pouvoir de résolution a été obtenu dans le cas de l'analyse de ces chaînes dans des conditions

physico-chimiques optimales qui permettent de travailler dans un régime d'analyse dans lequel les chaînes adoptent une conformation. Cette formation permet aux chaînes de se retrouver entièrement logées à l'intérieur du nanopore, avec lequel elles interagissent durant des temps suffisamment longs pour permettre la mesure de la contribution spécifique de chaque espèce d'une longueur donnée au blocage du courant.

En regard de cela, nous avons choisi d'utiliser le nanopore d'aérolysine pour explorer son potentiel afin de nous permettre d'atteindre une résolution à l'échelle de l'acide aminé dans le cas l'analyse de courts peptides ; dans l'optique de proposer une nouvelle stratégie originale de séquençage de protéines à l'échelle de la molécule unique basée sur l'utilisation de ce nanopore, différente de l'approche initiale visant à lire la séquence d'une protéine lors de son transport unidirectionnel à travers un nanopore.

Dans le **chapitre n°1**, je présente une **synthèse bibliographique** qui est organisée en trois parties.

La **première partie** a pour objectif : (i) d'introduire la technique en présentant son principe et l'intérêt de son utilisation, (ii) de présenter les différentes catégories de nanopores et les avantages et inconvénients de leur utilisation, avec une présentation plus détaillée des nanopores biologiques en raison de l'utilisation de ce type de nanopore dans cette thèse, et enfin ; (iii) de présenter quelques expériences pionnières ayant montré le potentiel de cette technique à ses débuts pour la détection et la caractérisation de polymères synthétiques et d'acides nucléiques. Je décrirai notamment dans cette partie comment ces études ont mené au développement de plusieurs applications comme la discrimination en taille de polymères synthétiques et le séquençage d'ADN par nanopore.

La **deuxième partie** a pour objectif : (i) d'introduire les techniques traditionnelles permettant de réaliser le séquençage des protéines ; (ii) de décrire les stratégies proposées pour le développement du séquençage des protéines par nanopore ; et enfin, (iii) de décrire d'autres approches de séquençage des protéines à l'échelle de la molécule unique.

La **troisième partie** a pour objectif de présenter le nanopore d'aérolysine avec une description : (i) de la structure du monomère, (ii) du mécanisme de formation d'un pore et d'insertion dans une membrane de lipides, (iii) de la structure et des propriétés électriques du pore, et (iv) de son utilisation comme nanopore pour la détection de molécules.

Dans le **chapitre n°2**, je présente la **méthode expérimentale** ce qui inclut la description des dispositifs expérimentaux de mesure utilisés, la méthode de reconstitution d'une membrane lipidique et d'insertion d'un nanopore biologique dans une membrane de lipides, des paramètres du signal auxquels nous nous intéressons, et des analyses de données permettant d'extraire de l'information à partir des mesures réalisées.

Dans les **chapitres n°3 et n°4**, je présente les résultats obtenus durant mes travaux de thèse.

Le **chapitre n°3** décrit les résultats qui montrent, pour la première fois, la capacité du nanopore d'aérolysine à séparer et identifier en taille des homopeptides d'une longueur de 5 à 10 acides aminés avec une résolution de discrimination d'un acide aminé. Nous avons également démontré la capacité de ce nanopore à détecter et identifier des impuretés peptiques de quelques centaines de molécules qui ne parviennent pas à être détectées par d'autres techniques existantes.

Le **chapitre n°4** décrit l'approche mise en place nous ayant permis de cartographier électriquement entre 13 et 16 des 20 acides aminés naturels par le nanopore d'aérolysine. Nous avons également montré qu'il est possible de détecter et d'identifier des modifications chimiques sur différents acides aminés. Je décrirai comment ces résultats nous ont permis de proposer une stratégie de séquençage des protéines basée sur l'utilisation du nanopore d'aérolysine.

Enfin, je terminerai ce manuscrit par une **conclusion générale** qui résume les travaux de cette thèse.

Chapitre 1 : Revue bibliographique

I. La détection électrique à l'échelle de la molécule unique par nanopore

1.1 Principe et origine de la technique

Le principe de la détection électrique d'une molécule unique par un nanopore est directement inspiré de la méthode du compteur Coulter (**Fig. 1**) [28]. Le compteur Coulter est une méthode de détection d'impulsions résistives (Resistive Pulse Sensing, RPS) [29] apparue dans les années 50, qui permet de déterminer la taille et le nombre de cellules présentes en suspension dans une solution électrolytique.

Le dispositif développé par Wallace Coulter se présente sous la forme d'une cuve constituée de deux compartiments séparés par une paroi solide dans laquelle est percée un trou (pore, ouverture). Chaque compartiment est rempli d'une solution ionique et d'objets d'intérêt dispersés dans cette solution. Un flux d'ions s'établit entre les deux compartiments sous l'action d'une tension électrique ou par simple diffusion. Deux électrodes placées dans chacun des compartiments permettent de mesurer la quantité d'ions circulant à travers le pore percé, ce qui se traduit par la mesure d'un courant ionique. Lorsque le pore est vide, les ions circulent librement à travers celui-ci et le courant ionique mesuré est maximal et dépend, de la dimension et de la géométrie du trou, et de la concentration ionique de l'électrolyte. Sous l'action d'une tension électrique, de l'agitation thermique, d'un gradient de concentration, d'un flux électro-osmotique, ou d'une pression, un objet peut entrer dans le pore. La présence d'un objet qui entre dans le pore et/ou interagit avec le pore, gêne partiellement le passage des ions et induit une chute brusque et temporaire du courant ionique (impulsion ou blocage du courant). Les caractéristiques des événements de chute du courant ionique (profondeur, durée, fréquence) sont liées aux propriétés de l'objet (longueur, composition chimique, conformation, et structure). Le courant ionique initialement mesuré est rétabli lorsque l'objet quitte le pore et regagne l'un des deux compartiments. L'objet ne traverse pas systématiquement le pore et peut regagner le compartiment initial dans lequel il résidait.

Lorsque Coulter propose son dispositif dans les années 50, celui-ci a pour objectif de dénombrer des cellules de plusieurs dizaines de micromètres en solution en associant le nombre de chutes du courant ionique mesurées au nombre de cellules ayant visité le pore [28]. L'automatisation du procédé a permis l'utilisation en routine de cette technique, notamment en hématologie pour l'analyse des cellules sanguines [30–32].

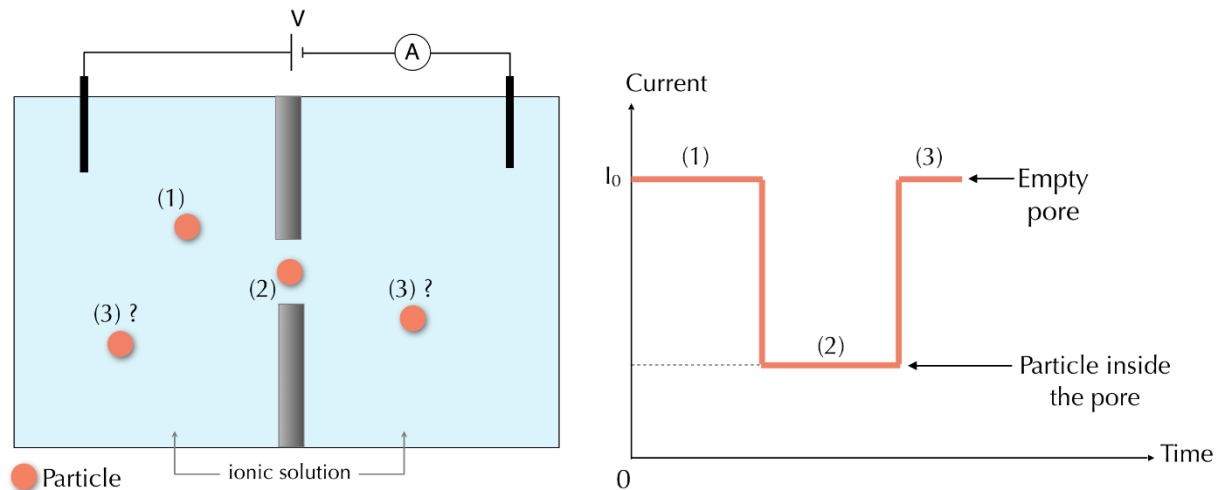


Figure 1 - Schéma représentatif du principe du compteur Coulter. Une cuve est séparée en deux compartiments par une paroi percée d'une ouverture micrométrique qui relie les deux compartiments remplis d'une solution ionique. Deux électrodes plongées dans chacun des compartiments permettent d'appliquer une tension électrique et de mesurer un courant ionique. Lorsque l'ouverture est vide, un courant ionique I_0 est mesuré. Une particule ajoutée en solution (1) peut entrer dans l'ouverture sous l'action d'une tension électrique ou d'une autre force motrice et induire une chute du courant ionique qui est mesurable (2). Le courant ionique I_0 est restauré lorsque la particule quitte l'ouverture, d'un côté ou de l'autre (3).

De telles mesures nécessitent que les particules soient d'une part, suffisamment séparées spatialement (diluées) et petites (ou déformables) pour pénétrer dans le canal une à une, et d'autre part, suffisamment grandes par rapport au diamètre de l'ouverture pour induire un blocage du courant détectable. Il convient de rappeler que dans ce cas de figure, la conductivité des particules est négligeable par rapport à l'électrolyte environnant. Dans le cas contraire, la particule devient moins détectable à mesure que sa conductivité s'approche de celle de l'électrolyte [33], ou dans de rares cas, peut même induire une augmentation détectable du courant [34,35] si la conductivité de la particule est supérieure à celle de l'électrolyte.

Compte tenu des dimensions micrométriques de l'ouverture du dispositif de Coulter (environ 100 μm de diamètre et plusieurs centaines de micromètres de longueur), ce dernier ne permet de dimensionner que des objets de taille micrométrique ou submillimétrique. La détection et le dimensionnement d'objets submicrométriques avec le dispositif de Coulter nécessite de diminuer les dimensions de l'ouverture. Dans ce contexte, Herbert E. Kubitschek a remplacé l'ouverture du compteur Coulter par une petite ouverture de 10 μm de diamètre percée dans une feuille de verre de 50 μm d'épaisseur pour compter et dimensionner des bactéries (1-10 μm) [36]. Dans le même sens, en 1970, Charles Bean et Ralph de Blois ont encore réduit l'ouverture du compteur Coulter pour atteindre un trou submicrométrique de 500 nm de diamètre percé

dans une feuille de plastique, pour analyser des objets submicrométriques tels que des particules sphériques de polystyrène (90 nm) [37], ou des virus (60 nm) [38,39].

Les expériences de détection d'objets nanométriques comme des molécules, nécessitent donc des pores de largeur nanométrique (nanopores). La première démonstration d'un dispositif permettant de détecter, de caractériser, et de compter des molécules individuelles, c'est-à-dire de façon analogue au compteur Coulter mais à l'échelle moléculaire, a été décrite par Zimmerberg et Parsegian [40], puis a été réalisée par John Kasianowicz et ses collègues en 1996 [1]. L'ouverture du dispositif de Coulter percée dans une paroi isolante solide a été substituée par une toxine bactérienne, l' α -hémolysine, formant un pore de diamètre nanométrique reconstituée dans une bicouche lipidique isolante.

Il est important de souligner que dans le contexte biologique de l'étude des membranes protéiques, la reconstitution de bicouches lipidiques et de membranes protéolipidiques a été découverte par Paul Mueller et ses collègues dans les années 1960 [41–43]. À la fin des années 1970, Erwin Neher et Bert Sakmann ont mis au point la technique du patch-clamp, dans laquelle les courants ioniques ont été enregistrés pour la première fois dans des cellules vivantes individuelles [44]. Il convient également de rappeler que le voltage-clamp a été développé au milieu du siècle dernier pour des travaux sur la membrane excitable de l'axone géant du calmar, afin de permettre l'élucidation du mécanisme ionique du potentiel d'action [45–47].

L'intérêt de réaliser un compteur Coulter moléculaire est de réaliser des mesures de grande précision sur des molécules uniques, au contraire des techniques en volume donnant accès à des mesures moyennées ; et cela sans marquage des molécules, ce qui peut s'avérer être une étape de préparation complexe.

1.2 Catégories de nanopores

Nous présentons dans cette section les différentes catégories de nanopores, ainsi que leurs avantages et leurs inconvénients, en donnant quelques exemples de pores ayant été utilisés comme biocapteurs électriques.

Un nanopore se définit comme un trou de diamètre nanométrique (environ 1 à 100 nm). On distingue au moins deux grandes catégories de pores : les nanopores biologiques qui sont formés par des canaux protéiques enchâssés dans des membranes de lipides reconstituées *in vitro*, et les nanopores solides qui sont formés par des trous percés dans des membranes synthétiques. Une troisième catégorie, plus récente, correspond aux nanopores hybrides qui

permettent de tirer profit à la fois des avantages des nanopores biologiques et solides. Enfin, d'autres types de nanopores plus particuliers peuvent également être formés par des nanotubes de cyclodextrine ou de carbone [48,49], par des assemblages d'ADN (DNA origami) [50], ou par des réseaux métalliques organiques (Metal Organic Framework, MOF) [51,52], mais ne seront pas développés ici. Nous ne développerons également pas les translocons, qui sont les canaux permettant le transport des protéines, et les protéasomes, qui sont les canaux participant à la dégradation des protéines pour leur recyclage. Le choix de l'utilisation d'un type de nanopore plutôt qu'un autre dépend de l'application recherchée et des avantages et inconvénients liés à chacun d'eux qui sont présentés à la suite. Durant cette thèse, nous avons utilisé un nanopore biologique, l'aérolysine, qui est une toxine bactérienne formant des pores dans les membranes de lipides.

1.2.1 Les nanopores biologiques

Les nanopores biologiques [53,54] correspondent principalement à des canaux ioniques de différents types qui parviennent à s'assembler sous la forme de pores et à s'insérer dans les membranes de lipides. Les canaux ioniques sont des pores protéiques enchâssés dans les membranes biologiques [55] qui assurent le transport des ions, de nutriments, de molécules d'eau, ou encore de macromolécules (acides nucléiques, protéines, sucres), de part et d'autre des cellules [56], et des organites [57]. On distingue les canaux qui ne permettent que le transport d'ions à proprement dit ; et les autres types de canaux qui permettent également le transport de molécules.

a) Les canaux ioniques

Les canaux ioniques les plus connus sont ceux qui sont exclusivement dédiés au transport d'ions et qui sont répartis en trois grandes catégories : les canaux dont l'ouverture est déclenchée par la liaison d'un ligand (chimio-dépendants) ; les canaux dont l'ouverture est déclenchée par un potentiel de membrane (voltage-dépendants) ; les canaux ioniques dont l'ouverture est déclenchée par une variation de la concentration d'un messager intracellulaire comme l'ATP ; ou encore les canaux mécano-sensibles. L'ouverture des canaux ioniques peut également être activée par d'autres stimuli comme une action mécanique ou une variation de température. Certains canaux ioniques présentent une grande sélectivité en ne permettant le passage que d'un seul type d'ion, alors que d'autres permettent le passage de plusieurs types d'ions [55,58,59]. Ces canaux protéiques sont très connus parce qu'ils sont à la base moléculaire de nombreux

mécanismes cellulaires comme l'activité nerveuse par la transmission des potentiels d'action électriques dans les cellules [60], la sensibilité aux neurotransmetteurs et aux hormones [61], et l'activité de contraction musculaire [55]. Les canaux ioniques jouent également un rôle essentiel dans la régulation de la vie de la cellule, notamment en limitant le nombre de cycles cellulaires par l'enclenchement de processus comme la mort cellulaire programmée ou apoptose [62]. Ainsi, les canaux ioniques sont omniprésents, enchâssés dans les membranes lipidiques, et leur bon fonctionnement est essentiel à tous les aspects de la vie [58,59]. Le dysfonctionnement de ces canaux peut entraîner le développement de certaines maladies comme c'est le cas de la mucoviscidose qui a pour origine une mutation du canal échangeur d'ions CFTR [63].

Il existe d'autres catégories de canaux protéiques qui participent au transport d'acides nucléiques, de protéines, de sucres, et de molécules d'eau, à travers un large éventail de membranes.

Une grande variété de canaux provenant d'organismes multiples, sont parvenus à être utilisés comme des nanopores biologiques insérés dans des membranes de lipides, pour la détection de molécules uniques. Parmi les canaux protéiques les plus utilisés et les plus connus pour leur utilisation pour la détection électrique d'analytes, on peut citer les catégories des peptides antimicrobiens ; des porines ; des canaux protéiques d'origine virale ; et des toxines qui forment des pores.

b) Les peptides antimicrobiens

Les peptides antimicrobiens [64] sont des peptides multifonctionnels dont le rôle biologique fondamental *in vivo* est d'éliminer des microorganismes pathogènes, y compris les bactéries Gram-positif et Gram-négatif, des champignons, et des virus. Certains de ces peptides microbiens sont capables de former des nanopores dans les membranes lipidiques comme c'est le cas de la gramicidine et de l'alaméthicine qui sont deux exemples très populaires de peptides microbiens ayant été largement utilisés comme canaux modèles pour l'étude de l'insertion d'un nanopore dans une membrane lipidique reconstituée *in vitro* [65]. Cependant, principalement en raison de leur diamètre étroit (environ 0,4 nm pour la gramicidine et 1,1 nm pour l'alaméthicine), et de leur cinétique rapide d'ouverture et de fermeture, leur utilisation la plus connue s'est limitée à la compréhension des processus de transport des ions, ou de petites molécules dans le cas de l'alaméthicine, en étudiant par analyse spectrale les fluctuations du bruit générées par les ions ou les molécules [66]. Ainsi, l'intérêt s'est tourné vers l'utilisation de

pores de plus larges diamètres comme les porines, les toxines qui forment des pores, et les canaux protéiques d'origine virale (viral portal protein).

c) Les toxines qui forment des pores

Les toxines qui forment des pores (Pore-Forming Toxin, PFT) [67–70] sont des protéines produites et sécrétées par une grande variété d'organismes dont certaines bactéries, certaines plantes, ou certains insectes, pour agir comme facteur de virulence en ciblant les membranes lipidiques des cellules hôtes, dans le contexte du processus d'infection de cellules cibles [71]. Les hôtes des PFTs sont la majorité du temps des cellules eucaryotes mais certaines toxines sont produites par certaines bactéries pour être dirigées contre d'autres bactéries comme c'est le cas de la colicine produite par *Escherichia coli*. En permettant la formation d'un pore dans la membrane lipidique de la cellule hôte, l'agent infectieux peut pénétrer dans la cellule pour dérégler son fonctionnement ou provoquer directement sa lyse. Cela rend les PFTs responsables d'un grand nombre de maladies infectieuses [72,73], comme l'infection des plaies profondes, la septicémie, ou les épidémies de gastro-entérite d'origine hydrique et alimentaire, pouvant aller jusqu'à la mort de l'organisme infecté de par leur capacité à perméabiliser les membranes cellulaires de leur hôte.

Les PFTs sont classées en deux familles structurales : les α -PFTs et les β -PFTs qui forment des pores respectivement sous la forme de structures en hélices α , ou en feuillets β . Bien que chacune des familles présente des caractéristiques structurales et des mécanismes de formation des pores spécifiques, un mécanisme général de leur mode d'action qui est similaire a été défini (**Fig. 2**) [70]. Les toxines sont d'abord produites sous la forme de monomères solubles qui peuvent nécessiter une activation par protéolyse enzymatique pour leur assemblage sous la forme d'un pore. Les toxines diffusent ensuite vers la membrane de la cellule cible, au niveau de laquelle elles se fixent par le biais d'une interaction avec un récepteur qui peut être une protéine, un peptide, ou un carbohydrate ; ou avec un lipide en absence de récepteurs. La concentration des monomères à la surface de la membrane cellulaire, favorise l'oligomérisation des unités monomériques de la toxine en une structure cylindrique qui parvient à s'insérer dans la membrane pour y former un pore.

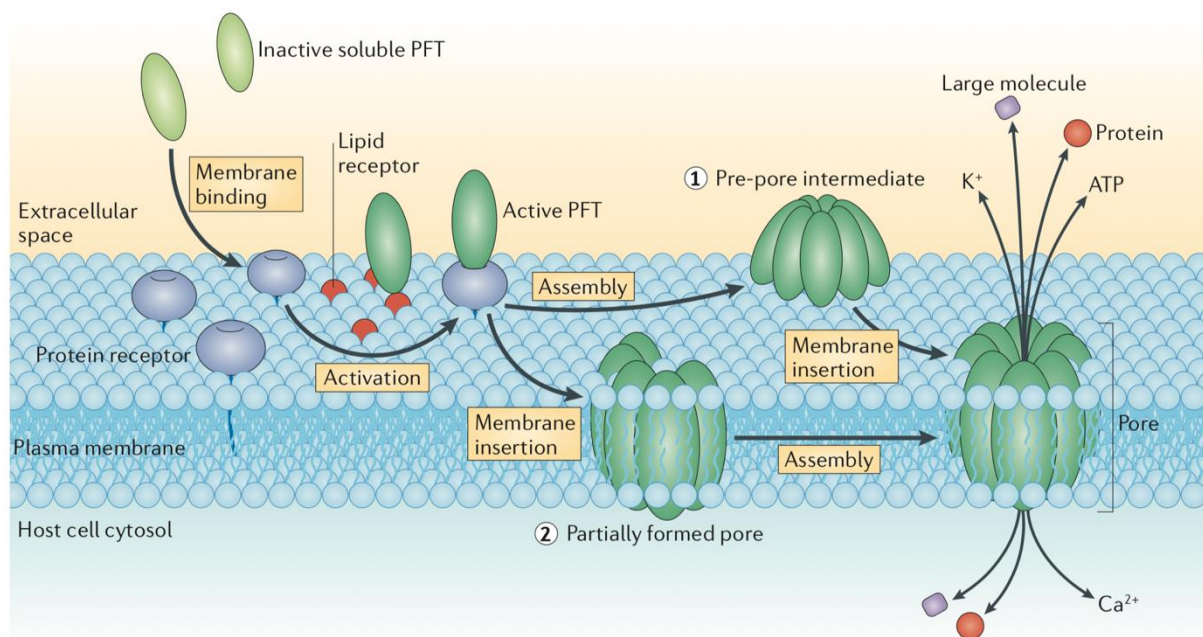


Figure 2 - Schéma représentatif des étapes de formation d'un pore par une PFT. Les monomères sont sécrétés par les bactéries sous une forme soluble inactive. Les monomères se lient ensuite à la membrane, ce qui entraîne son activation. Cela est suivi de la pré-formation de pores (1) ou de l'insertion de pores dans la membrane (2). Enfin, l'assemblage final conduit à l'existence d'un pore stable inséré dans la membrane de lipides. (Adaptée de [70]).

De nombreuses PFTs ont été utilisées comme des nanopores pour la détection de divers analytes.

Parmi les α -PFTs, on peut citer l'actinoporine Fragaceatoxin (FraC) produite par certaines anémones de mer [74], ou la cytolysine A (ClyA) produite par *Escherichia coli* et certaines souches de *Salmonella enterica* [75]. Parmi les β -PFTs, on peut citer l' α -hémolysine produite par *Staphylococcus aureus* [76], l'aérolysine produite par *Aeromonas hydrophila* [23,77], ou la lysénine produite par *Eisenia fetida* [78,79] et qui est un membre de la famille des aérolysines.

L' α -hémolysine est à ce jour la toxine qui a été la plus utilisée comme nanopore biologique pour la détection électrique de molécules uniques. Elle possède une structure très proche du nanopore d'aérolysine que nous avons utilisé. Pour ces raisons, la structure de l' α -hémolysine est détaillée ci-dessous. Les informations concernant la structure et les propriétés du pore d'aérolysine seront présentées dans la partie n°3 de ce chapitre.

La toxine α -hémolysine est synthétisée sous la forme de monomères solubles de 33,2 kDa qui s'oligomérisent sous la forme d'un heptamère pour former un pore de 232,4 kDa [80]. La structure tridimensionnelle de l' α -hémolysine a été déterminée par cristallographie aux rayons X en 1996 avec une résolution de 1,9 nm [76]. D'après ces travaux, le pore mesure un total de

10 nm de largeur et de 10 nm de longueur (**Fig. 3**). Le canal adopte une forme semblable à celle d'un champignon où deux parties sont distinguées : un chapeau constitué d'un vestibule d'un diamètre relativement large de 3,6 à 4,6 nm qui facilite la capture des macromolécules ; et un grand segment transmembranaire correspondant à une structure en tonneau bêta d'une longueur de 5,2 nm et d'un diamètre de 2,6 nm (**Fig. 3**). Une zone de constriction présente entre ces deux segments présente un diamètre de 1,4 nm, et constitue une barrière pour la translocation d'analytes qui peut être surmontée en appliquant une différence de potentiel suffisamment grande [81–83]. Le chapeau du pore est hautement chargé alors que l'intérieur du tonneau bêta est constitué de résidus polaires non chargés, à l'exception de la constriction au niveau de laquelle sont retrouvés des anneaux de résidus d'acide glutamique et de lysine, et à la sortie du canal où sont retrouvés des anneaux de résidus de lysine et d'acide aspartique [76,84]. Le canal présente une légère sélectivité aux anions et une rectification du courant dans le sens des anions, c'est-à-dire correspondant à une plus grande quantité de courant conduite sous l'action d'une tension électrique positive [76,84].

L'analyse de molécules uniques par nanopore a beaucoup été étudiée avec l' α -hémolysine depuis les deux articles pionniers [1,85] sur la détection de molécules d'ADN [1], et d'ions métalliques [85], qui ont ouvert la voie vers l'étude d'autres analytes avec ce nanopore tels que de petites molécules, des peptides, des polymères organiques, des protéines, des enzymes.

Des caractéristiques spécifiques de ce nanopore expliquent son succès pour l'analyse de molécules uniques. Premièrement, le pore présente une structure stable qui ne contient aucune partie mobile, le rendant idéalement adapté pour détecter et caractériser des molécules avec précision. En effet, la structure du pore permet d'avoir une conductance unitaire qui correspond à un niveau de courant constant au cours du temps, ce qui facilite la détection des blocages engendrés par l'analyte ; bien que dans certaines conditions (*e.g* faibles concentrations ioniques), le courant peut cependant fluctuer [86]. Deuxièmement, l' α -hémolysine ne présente aucune capacité intrinsèque à se lier spécifiquement à des analytes. Cette spécificité peut néanmoins être apportée par l'introduction de modifications génétiques ou chimiques [54]. Enfin, le pore présente une bonne stabilité à des températures proches de 100°C et dans une large gamme de pH (pH 2-12) [87], mais également en présence d'agents chaotropiques [88].

Bien que le nanopore d' α -hémolysine soit répandu pour l'analyse de nombreuses molécules, l'utilisation de celui-ci s'est trouvée limitée par son diamètre interne (environ 1,4 nm).

Pour prendre l'exemple de l'analyse de l'ADN, l' α -hémolysine permet de transporter une chaîne d'ADN simple brin, alors qu'une chaîne double brin est trop large pour entrer dans le pore. De plus, la longueur du tonneau bêta du pore n'est pas adaptée pour atteindre une précision de détection d'un nucléotide individuel pour la lecture d'une séquence d'ADN par un nanopore. Si l'on considère également l'exemple de l'analyse des protéines, le diamètre interne de l' α -hémolysine n'est pas toujours adapté à l'entrée de certaines protéines natives.

Pour contourner les limitations rencontrées avec l' α -hémolysine, d'autres PFTs ont été utilisées, comme la toxine ClyA qui présente un diamètre plus large que le nanopore d' α -hémolysine et qui s'est révélée être plus adaptée à l'analyse de larges protéines [89] ; ou la toxine FraC, qui au contraire de l' α -hémolysine, présente une région transmembranaire structurée en hélices alpha avec un large vestibule, suivi d'une zone de constriction dont le diamètre est modulable par mutation génétique [90,91], ce qui a permis de transporter des doubles brins d'ADN avec ce nanopore [92].

D'autres types de nanopores biologiques présentés à la suite ont également été utilisés pour contourner certaines limitations rencontrées avec le nanopore d' α -hémolysine.

d) Les porines

Les porines [93] sont des canaux ioniques de diamètre nanométrique qui représentent les composants majoritaires de la membrane extérieure (Outer membrane protein, Omp) des bactéries à Gram-négatif à double membrane [94–96], et qui sont également retrouvés chez les mitochondries et les plastes. Les porines sont classées selon leurs propriétés de transport, ou bien selon leur assemblage structural. Concernant les propriétés de transport [94,97] on distingue : (i) les porines aux propriétés générales de diffusion (*e.g* OmpF, OmpC, PhoE) qui assurent la diffusion non spécifique de molécules hydrophiles (< 600 Da), et de certains ions avec une légère sélectivité ; et (ii) les porines aux propriétés de diffusion spécifique par exemple de sucres comme la maltodextrine (*e.g* LamB), de nucléosides (*e.g* Tsx), ou de métaux (*e.g* FhuA, BtuB). Concernant leur assemblage, les porines sont généralement constituées de 8 à 24 brins bêta qui sont structurés sous la forme d'un tonneau bêta. Les tonneaux bêta formés s'assemblent sous la forme de monomères (*e.g* MspA, OmpG, FhuA), ou sous la forme de trimères (*e.g* OmpF).

Parmi les porines les plus célèbres pour leur utilisation comme nanopores pour la détection d'analytes, on peut citer OmpF [98–102] et OmpC [101–104] qui ont plus particulièrement été

utilisées comme nanopores pour l'étude du transport d'antibiotiques [96,98,99,105], OmpG [106,107], FhuA [108], LamB [109], NfpA et NfpB [110], la MspA [111], ou encore le VDAC [112–114] d'origine mitochondriale [115].

La porine MspA produite par *Mycobacterium segmatis* a suscité un très grand intérêt dans le développement du séquençage de l'ADN par nanopore [111]. La MspA se présente sous la forme d'un octamère organisé en tonneau bêta dont la zone de sensibilité présente des dimensions plus adaptées que celle de l' α -hémolysine (**Fig. 3**) dans l'optique d'atteindre une précision de lecture d'une base nucléotidique, comme présenté plus loin dans la section consacrée au séquençage d'ADN par nanopore (chapitre n°1, 1.3.2, b).

On peut également citer le canal CsgG responsable de la sécrétion de protéines amyloïdes chez la bactérie *Escherichia coli* [116–118]. Le pore CsgG forme un canal structuré en un tonneau bêta à 36 brins, de 4 nm de large, et qui présente une constriction d'environ 1,2 nm de large et 2 nm de longueur située en son centre [119] (**Fig. 3**) qui aurait notamment permis à l'entreprise Oxford Nanopore Technologies de l'utiliser pour le séquençage d'ADN par nanopore [120,121].

Une limitation majeure connue à l'utilisation des porines concerne la lumière de ces pores qui peut contenir des boucles peptidiques flexibles qui sont à l'origine de fluctuations du courant mesurées en absence d'analytes (gating) [122], qui peut être surmontée par l'introduction de mutations génétiques [123,124].

e) Viral portal protein

Enfin la dernière catégorie de nanopores biologiques que l'on décrira ici correspond à celle des canaux protéiques d'origine virale. On prendra plus particulièrement ici l'exemple du canal du phi29.

Le bactériophage phi29 est un virus à ADN double brin qui infecte les bactéries et dont le génome est contenu dans une procapside préformée [125–128]. Le virus utilise un moteur à ATP pour encapsuler son génome dans sa capsid nanométrique, et assemble des canaux nanométriques pour transporter son génome pendant la réplication. Le moteur d'empaquetage de l'ADN du bactériophage phi29, qui permet à l'ADN double brin d'entrer dans le virus pendant la maturation, et de sortir pendant une infection, contient une protéine de connexion (connecteur) qui forme un canal d'une largeur comprise entre 3,6 et 6 nm et qui est un composant essentiel du moteur. Le connecteur est dodécamérique, c'est-à-dire qu'il est composé

de douze sous-unités de la protéine gp10, qui s'assemblent en un canal avec une section transversale servant de chemin pour la translocation de l'ADN double brin [129–131].

Le connecteur est parvenu à être inséré avec succès dans une bicouche de lipides et a montré la possibilité d'être utilisé comme un nanopore biologique pour la détection à l'échelle de la molécule unique [132,133]. L'utilisation du moteur phi29 comme nanopore pour la détection d'ADN a été montrée en 2009 [132]. Le diamètre interne du canal est de 3,6 nm pour la partie la plus étroite et de 6 nm pour la partie la plus large (**Fig. 3**) ce qui fait de ce nanopore l'un des pores biologiques les plus larges [129,130].

En comparaison à d'autres nanopores biologiques comme l' α -hémolysine et la MspA, le pore phi29 présente un plus large diamètre qui permet l'entrée de molécules d'ADN simple brin et double brin [132–137]. Par ailleurs, au contraire de nombreux autres nanopores biologiques, la structure du pore formée par le phi29 ne correspond pas à un tonneau bêta.

Plus récemment, un autre canal d'origine viral, le connecteur du thermophage G20c est parvenu à être utilisé comme un nanopore biologique inséré dans une membrane de lipides pour la détection de molécules de cyclodextrine [138].

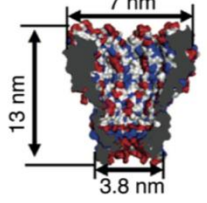
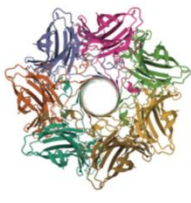
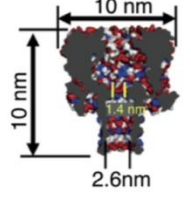
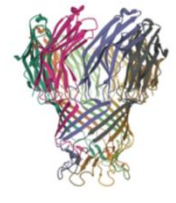
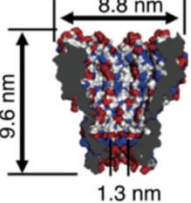
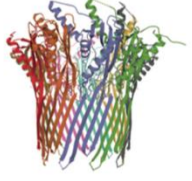
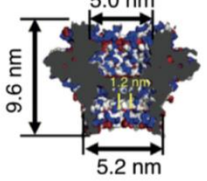
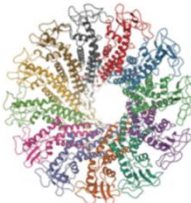
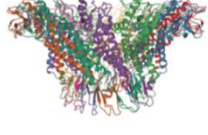
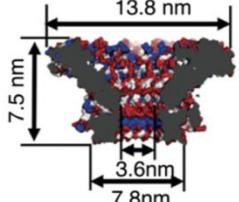
Pore family	Nanopore	Top view	Front view	Dimensions
α -PFT	ClyA			
β -PFT	α -hemolysin			
Porins	MspA			
	CsgG			
Portal Protein	Phi29 motor			

Figure 3 - Représentation de quelques nanopores biologiques communément utilisés. Structure et dimensions des pores ClyA, α -hémolysine, MspA, CsgG, et Phi29 (PDB : respectivement 2WCD, 7AHL, 1UUN, 4Q79, 1FOU). De gauche à droite : la troisième et la quatrième colonne montrent respectivement une vue de haut et une vue de face de chaque pore représenté selon sa composition en structures secondaires, et avec chaque monomère constitutive d'un pore représenté par une couleur donnée.. La cinquième colonne montre une coupe frontale des différents pores qui sont représentés selon leur distribution de charge, avec une indication de leurs dimensions. (Adaptée de [139]).

1.2.2 Autres catégories de nanopores

Les nanopores biologiques sont historiquement les premiers à avoir été utilisés pour la détection électrique de molécules. Un de leurs plus grands avantages est une très grande reproductibilité en termes d'assemblage ce qui permet d'obtenir des nanopores de dimensions, de géométrie, et de structure chimique contrôlées. Ce paramètre est important pour obtenir des mesures précises et reproductibles. Les nanopores peuvent également être modifiés génétiquement, chimiquement, ou être combinés à des moteurs enzymatiques. L'utilisation des nanopores biologiques est plutôt facile à mettre en œuvre et peu coûteuse. Une fois insérés dans la membrane, certains nanopores présentent une grande stabilité et peuvent y rester plusieurs heures. Les membranes de lipides reconstituées *in vitro* sont également résistantes à l'insertion du pore et à des tensions électriques pouvant aller jusqu'à une centaine de millivolt. Néanmoins, elles peuvent se rompre spontanément au cours de l'expérience. De plus, les pores protéiques de par leur structure chimique sont connus pour être sensibles aux paramètres physico-chimiques tels que la température, le pH, la force ionique, la tension électrique, ou encore la présence d'un agent chimique dénaturant. Enfin, les pores qui existent dans la nature présentent des diamètres compris entre moins de 1 nm à environ 5-7 nm pour les pores protéiques les plus larges, ce qui peut être limitant pour l'analyse d'objets plus volumineux (> 10 nm) [140]. Une solution proposée à ce problème est l'utilisation de molécules, comme un polymère synthétique [141,142], un disaccharide [143], une protéine kinase inhibitrice de peptides [144], ou d'un aptamère [145], qui, attachées à l'entrée d'un nanopore, ont permis de détecter des protéines trop volumineuses pour atteindre la lumière d'un nanopore comme celui de l' α -hémolysine, dans certains cas à des concentrations de l'ordre du nanomolaire.

a) Les nanopores solides

Les nanopores solides (**Fig. 4**) ont été développés par la suite en alternative à certains inconvénients rencontrés avec les nanopores biologiques. Les nanopores solides sont des trous de diamètre nanométrique obtenus à partir du perçage de membranes artificielles, par un faisceau d'ions [146], par un faisceau d'électrons [147], ou par claquage diélectrique [148], dans des matériaux de nature organique ou inorganique [149–151], et d'épaisseur pouvant varier de quelques nanomètres à plusieurs micromètres. La nature du matériau utilisé dépend des dimensions, de la géométrie, et des propriétés du nanopore souhaité, qui sont directement en lien avec l'application recherchée [151]. Parmi ceux utilisés, on retrouve des matériaux à base de silicium (SiN, SiO₂) [146,147] ; d'aluminium (Al₂O₃) [152] ; de bore (BN) [49] ; de

molybdène (MoS_2) [153] (**Fig. 4**) ; de graphène [154,155] ; de polymères (*e.g* polyéthylène tétraphthalate, polyfluorure de vinyldène, polycarbonate) [156] ; ou encore des matériaux composites [157–159] afin de combiner différentes propriétés. Le processus de fabrication du pore dépend également du matériau utilisé et du nanopore recherché. D’une part, les membranes solides présentent l’avantage d’être robustes et résistantes aux variations des paramètres physico-chimiques. D’autre part, les nanopores solides sont fabriqués artificiellement et leurs dimensions peuvent ainsi être modulées selon l’analyse envisagée. Les nanopores solides connus dans la littérature présentent des diamètres compris entre 0,5 nm à plusieurs centaines de nanomètres. Cependant, comme les nanopores solides sont fabriqués artificiellement, ils présentent au contraire des nanopores biologiques, l’inconvénient majeur de ne pas présenter des propriétés de dimensions, de géométrie, et de structures chimiques finement reproductibles et contrôlées [151]. Cet aspect est limitant pour pouvoir comparer les résultats d’expériences indépendantes. Il est également difficile de contrôler l’état de surface [160]. Une alternative à cette limitation est de fonctionnaliser les nanopores solides (**Fig. 4**) avec des protéines [149], des épingles d’ADN [161], des lipides [162,163], ou des polymères synthétiques [164].

b) Les nanopores hybrides

Les limites rencontrées avec les nanopores biologiques et solides ont mené au développement de nanopores hybrides qui combinent à la fois les avantages des nanopores biologiques et solides. On peut citer l'exemple d'un nanopore d' α -hémolysine inséré dans un nanopore solide de polyéthylène tétraphthalate (PET) recouvert d’une alternance de films minces d’oxide de zinc (ZnO) et d’oxide d’aluminium (Al_2O_3) (**Fig. 4**) [159,165] ; d'un nanopore de gramicidine A inséré dans une membrane de polycarbonate [166,167] ; ou encore d’un canal protéique viral G20c inséré dans un nanopore solide de nitrure de silicium (SiN) (**Fig. 1.4**) [168].

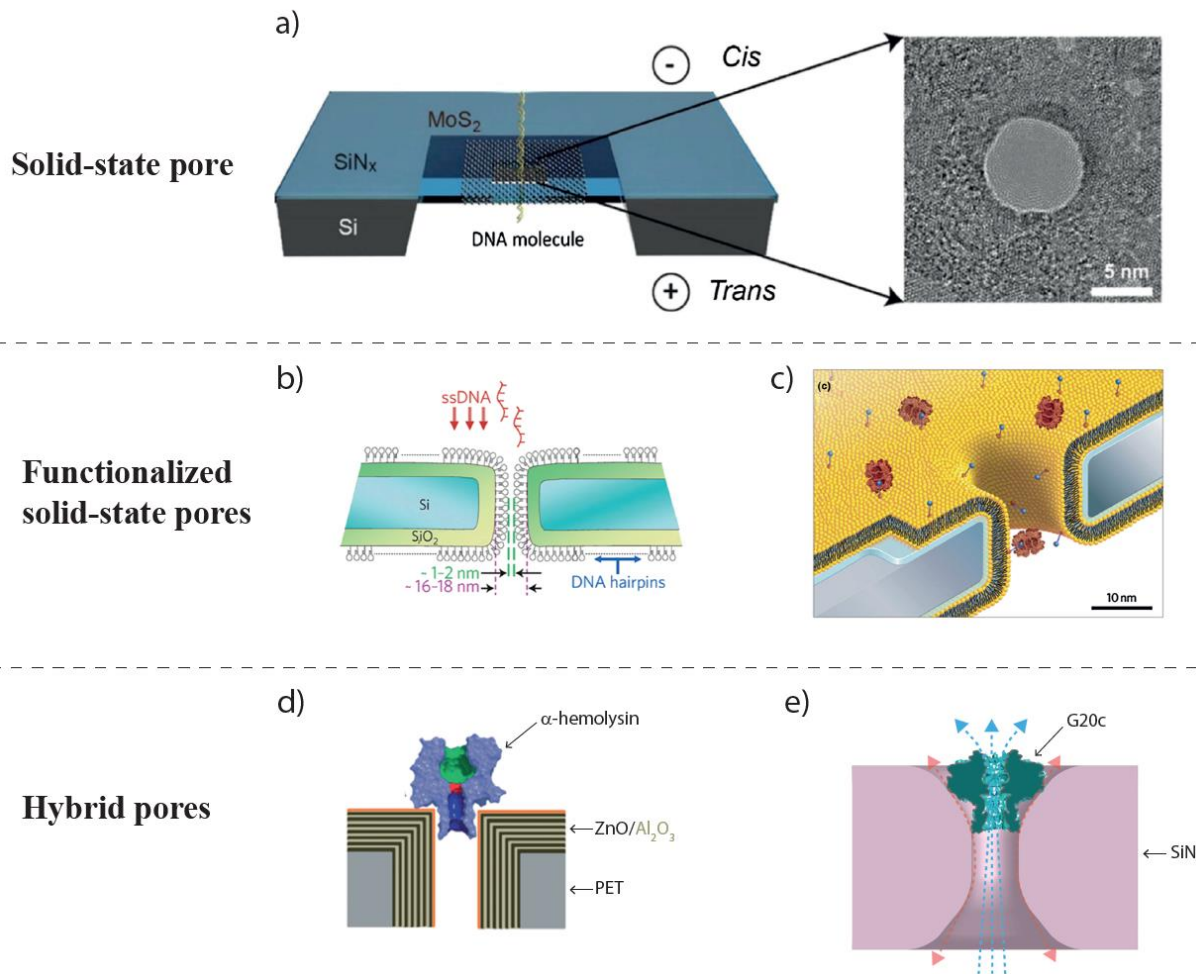


Figure 4 - Représentation de quelques exemples de nanopores solides et hybrides. (a) Schéma d'un nanopore solide percé dans une membrane de MoS_2 dans lequel une molécule d'ADN est transportée (gauche), et cliché de microscopie électronique à transmission du nanopore formé dans la membrane (droite) [153]. (b) Schéma d'un nanopore solide fonctionnalisé par le recouvrement de sa surface par des épingles d'ADN [161]. (c) Schéma d'un nanopore solide fonctionnalisé par le recouvrement de sa surface par une bicouche de lipides permettant l'incorporation de ligands pour le transport sélectif de molécules [162]. (d) Schéma d'un nanopore hybride correspondant à un pore d' α -hémolysine inséré dans un nanopore solide [159]. (e) Schéma d'un moteur moléculaire G20c inséré dans un nanopore solide [168].

1.3 Expériences pionnières et applications

Les premières expériences à l'échelle de la molécule unique associées à un nanopore biologique, couplé à un système de détection électrique apparaissent au début des années 80.

Comme pour le compteur Coulter, le principe de ces expériences consiste à mesurer les fluctuations ou impulsions du courant, engendrées par le transport et/ou l'interaction d'une molécule individuelle dans un canal protéique de diamètre nanométrique, inséré dans une bicouche de lipides reconstituée manuellement, aussi connu sous le nom de bicouche lipidique noire (Black Lipid Membrane, BLM) (**Fig. 5**) [169]. Les caractéristiques des impulsions du courant (amplitude, durée, et fréquence) dépendent en particulier de la molécule transportée (nature chimique, conformation, dimensions) et du type d'interaction entre la molécule et le pore.

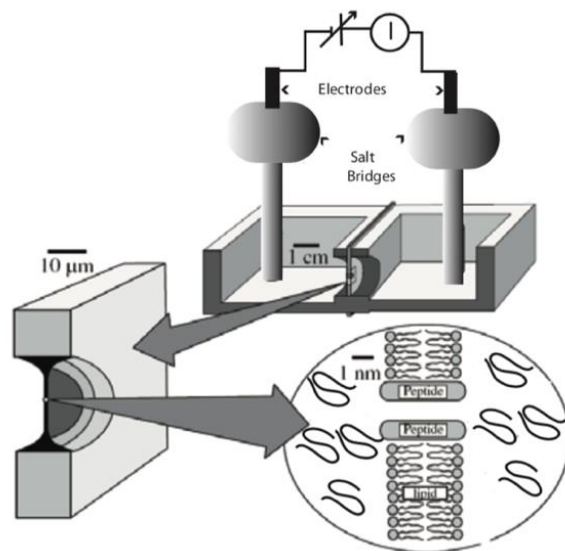


Figure 5 - Schéma représentatif d'un montage expérimental de détection à l'échelle moléculaire par un nanopore biologique inséré dans une membrane de lipides. Une cuve est séparée en deux compartiments par une paroi solide. La paroi solide est percée d'un trou de l'ordre d'une centaine de micromètres reliant les deux compartiments. Deux électrodes sont plongées dans chacun des compartiments et permettent d'appliquer une tension électrique entre les deux compartiments et de mesurer le courant résultant (haut). Une bicouche de lipide est reconstituée manuellement par un dépôt de lipides au niveau du trou micrométrique (milieu). Un pore de diamètre nanométrique, le nanopore, est inséré dans la bicouche de lipides préformée. Des molécules ajoutées en solution peuvent être transportées dans le nanopore (bas). (Adaptée de [170]).

1.3.1 Caractériser des pores avec des polymères

A la fin des années 80, peu d'informations structurales étaient connues sur les canaux protéiques. Pour résoudre partiellement ce problème, les premières expériences de transport d'une molécule unique ont été dédiées à la caractérisation des dimensions (diamètre, géométrie) des pores protéiques. Le principe de cette caractérisation repose sur l'utilisation de certains sucres (*e.g* dextrane), ou de polymères synthétiques non ioniques comme le polyéthylène glycol (PEG) qui est disponible dans une large gamme de masses moléculaires. L'utilisation de molécules de PEG de rayon de giration inférieur au rayon du pore étudié, et qui sont injectées dans le pore, entraînerait une diminution de la conductance moyenne du canal. La taille du pore est ensuite estimée en déterminant le plus gros polymère pouvant pénétrer dans le pore. Sur ce principe, plusieurs études ont proposé de caractériser les dimensions de différents nanopores, comme l' α -hémolysine, en y injectant des chaînes de polymères de tailles connues et croissantes [40,171–174].

1.3.2 Caractériser des polymères avec des pores

a) Spectrométrie de taille par nanopore

Après avoir montré que les polymères peuvent être utilisés pour sonder la taille des nanopores, les groupes de Krasilnikov et de Kasianowicz démontrent en 2007 [175], pour la première fois, que les nanopores peuvent également être utilisés pour déterminer avec précision et avec une haute résolution la taille des polymères.

Avant cette expérience fondatrice, on peut citer la première expérience de détection de petites chaînes de PEG [170,173] transportées à travers un pore d'alaméthicine inséré dans une bicouche de lipides reconstituée *in vitro*. Cependant, au contraire du compteur Coulter, l'entrée des chaînes de PEG dans le pore provoque des changements de conductance et du bruit électrique traduisant la présence des chaînes de polymères dans le pore mais n'induisent pas, par manque de résolution, des impulsions électriques du courant directement mesurables.

La preuve de l'existence du transport et le comptage individuel de molécules à travers un nanopore n'a été apportée qu'à la fin des années 90 dans le contexte des premières expériences d'étude de chaînes d'ADN qui sont présentées dans la section suivante. Dans le cas des PEGs, les premières impulsions résistives visibles et résolues avec des temps de séjour caractéristiques de l'ordre d'une dizaine de millisecondes ont été montrées par O. Krasilnikov, C. Rodrigues, et S. Bezrukov [176] en utilisant une concentration quasi saturante de chlorure de potassium (KCl)

à une concentration de 4 mol/L (4 M). Immédiatement après, J. Kasianowicz, O. Krasilnikov, et leurs collègues, ont démontré que les amplitudes des impulsions résistives pouvaient être utilisées pour construire des "spectres de masse" avec une résolution de monomères uniques (Fig. 6), permettant l'identification d'espèces moléculaires d'une taille donnée dans un mélange [175]. Cette expérience fondatrice, initialement appelée « spectrométrie de masse » à l'échelle de la molécule unique par nanopore (Nanopore-based Single Molecule Mass Spectrometry, Np-SMMS) a plus tard été renommée « spectrométrie de taille » à l'échelle de la molécule unique par nanopore (Nanopore-based Single-Molecule Size-Discrimination, Np-SMSD) [177], car la profondeur de l'impulsion résistive fournit des informations plutôt sur la taille que sur la masse de la molécule.

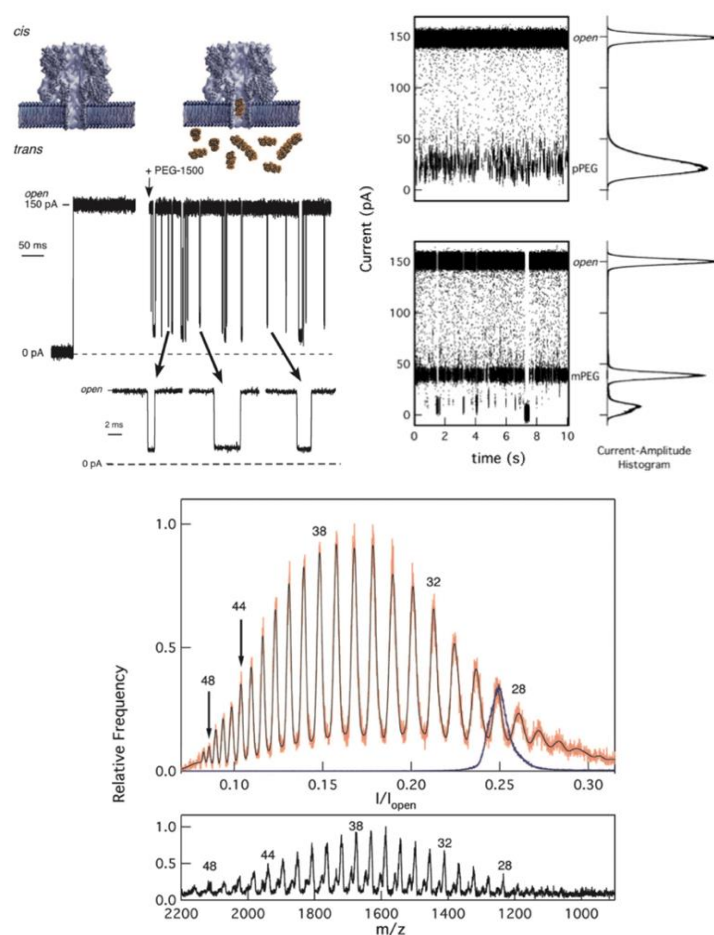


Figure 6 - Discrimination en taille de polymères synthétiques avec le nanopore d' α -hémolysine. (a) Schéma d'un montage expérimental montrant un nanopore d' α -hémolysine inséré dans une bicouche de lipides, avant et après l'ajout d'une solution polydisperse de PEG de masse molaire 1500 g/mol. L'ajout des molécules de PEG du côté cis (chapeau) du nanopore dans une solution ionique de KCl 4M, Tris 5mM, pH = 7,5, entraîne l'apparition de blocages du courant ionique, de durées de l'ordre d'une à plusieurs millisecondes, sous l'action d'une tension électrique de +40 mV. (b) Portion d'une trace du courant en fonction du temps (gauche)

et histogramme du nombre de points en fonction du courant ionique (gauche) dans le cas de l'analyse d'une solution polydisperse (haut) ou monodisperse (bas) de PEG (1500 g/mol) à travers le nanopore d' α -hémolysine. Les histogrammes mettent en évidence la nature monodisperse ou polydisperse de l'échantillon à chaque fois en reflétant la présence, respectivement d'une population étroite. Cette population reflète l'existence de blocages du courant d'une amplitude caractéristique qui est reproduite un grand nombre de fois (une population additionnelle est observée correspondant à des événements causés par des impuretés présentes dans la solution de PEG), et d'une population de base plus large representative d'événements de blocage du courant aux amplitudes multiples correspondant aux différentes longueurs de polymères présentes en solution. Cela indique que le nanopore d' α -hémolysine permet la discrimination de polymères de masses ou de tailles différentes. (c) Distribution du courant relatif résiduel (I/I_{open}) (haut) en comparaison avec une spectre de masse MALDI-TOF (bas) obtenue dans le cas de l'analyse d'une solution polydisperse de PEG (1500 g/mol) ajoutée du côté Cis du nanopore d' α -hémolysine. Les valeurs les plus grandes de I/I_{open} correspondent aux polymères de plus petites masses et vice versa. La superposition de l'histogramme de I/I_{open} dans le cas de l'analyse d'une solution monodisperse de PEG 1294 g/mol (29 monomères) et la comparaison au spectre de masse MALDI-TOF permettent d'attribuer une masse à chaque polymère et de voir que l'histogramme permet de discriminer les espèces de la solution polydisperse avec une différence d'un monomère. (Adaptée de [175]).

En 2001, Vercoutere *et al.* propose l'utilisation du nanopore d' α -hémolysine dans une configuration classique pour l'analyse d'épingles d'ADN constituées de trois à huit nucléotides (**Fig. 7**) [178]. Les données obtenues ont été analysées par un programme qui leur a permis de mettre en évidence l'existence de niveaux de courant moyen spécifiques de chacune des longueurs d'épingles d'ADN analysées qui sont distinguables sur les histogrammes construits, et de conclure ainsi à l'obtention d'une résolution de discrimination d'un nucléotide.

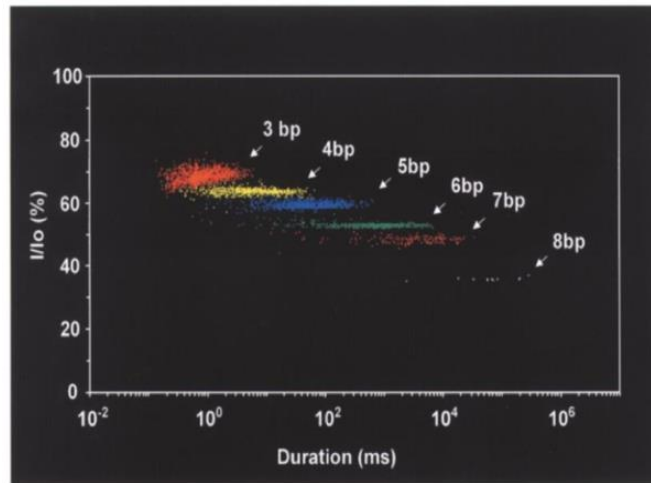


Figure 7 - Discrimination de molécules d'ADN en épingles à cheveux (DNA hairpin) avec une résolution d'une paire de base, obtenue avec le nanopore d' α -hémolysine. Nuages de points représentatifs de la distribution du courant relatif I/I_0 en fonction de la durée des blocages du courant dans le cas de l'analyse d'épingles d'ADN de 3, 4, 5, 6, 7, et 8 paires de bases ajoutées du côté cis (chapeau) de l' α -hémolysine, sous l'action d'une tension électrique de +120 mV appliquée en trans, et dans une solution de KCl 1M, et à température ambiante. Les épingles d'ADN de différentes longueurs sont discriminables par leur niveau de courant résiduel I/I_0 , et présentent des durées de blocage du courant de plus en plus grandes avec l'augmentation de la longueur de la molécule. (Adaptée de [178]).

b) Séquençage des acides nucléiques par nanopore

Après avoir observé que des canaux protéiques comme celui d' α -hémolysine peuvent être maintenus ouverts et stables dans des bicouches de lipides, et qu'il est possible de suivre en temps réel le transport de molécules comme des polymères de PEG dans un nanopore ; les études se sont tournées vers l'utilisation du nanopore d' α -hémolysine pour l'analyse des acides nucléiques.

Dans cette expérience fondatrice, l'objectif des auteurs était d'analyser des homopolymères simple brin d'ADN (poly-A, acide poly-Adénylique) et d'ARN (poly-U, acide poly-Uridylique) en présence d'un nanopore d' α -hémolysine inséré dans une bicouche lipidique [1]. L'addition des brins monocaténares d'ADN et d'ARN montrent l'apparition immédiate de nombreux blocages transitoires du courant d'une durée moyenne de l'ordre de la milliseconde survenant sous l'action d'une tension électrique négative (-120 mV) (**Fig. 8**). Au contraire, l'addition de doubles brins des mêmes homopolymères d'ADN et d'ARN n'a pas donné lieu à l'observation de blocages du courant, indiquant que les molécules bicaténares ne sont pas transportées dans le pore, probablement en raison du diamètre de la constriction de l' α -hémolysine [76].

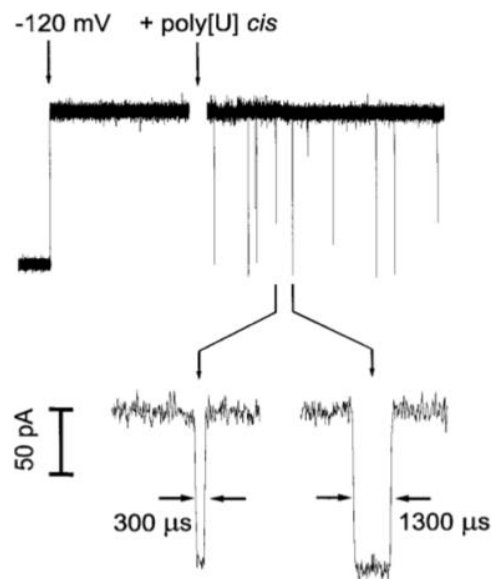


Figure 8 - Détection du transport d'homopolymères simple brin d'ADN avec le nanopore d' α -hémolysine. Evolution du courant ionique mesuré à travers un nanopore d' α -hémolysine inséré dans une bicouche lipidique sous l'action d'une tension électrique de -120 mV en cis, avant et après l'addition d'homopolymères d'uracile (poly[U]) d'une longueur de 210 bases nucléotidiques du côté cis du nanopore et insert montrant deux événements de blocage du courant typiques. (Adaptée de [1]).

Les auteurs ont démontré, par amplification PCR (Polymerase Chain Reaction) que les impulsions du courant observées lors de l'analyse des simples brins d'ADN et d'ARN correspondent à la translocation des molécules, c'est-à-dire à leur passage du compartiment initial dans lequel elles ont été ajoutées, vers le compartiment opposé, *via* leur passage par le nanopore. D'une part, ils ont observé que la durée moyenne de séjour des ARN (poly-U) dans le pore augmente de façon proportionnelle à la longueur de la chaîne, et de façon inversement proportionnelle à la tension électrique appliquée ; ce qu'ils ont interprété comme le transport des homopolymère d'ARN dans le pore dans une conformation linéaire. D'autre part, ils ont déterminé par PCR que le nombre de blocages du courant ionique compté correspond au nombre de chaînes simple brin d'ADN ayant transloqué et ont confirmé que seuls les ADN simple brin étaient transportés dans le nanopore. Ces résultats ont fourni la première preuve convaincante que les homopolymères d'acides nucléiques monocaténares sont transportés par électrophorèse sous l'action d'un champ électrique dans un canal protéique nanométrique.

La démonstration expérimentale de la détection de la translocation d'homopolymères individuels d'ADN et d'ARN linéaires était la première étape suggérant que les nanopores pourraient être utilisés pour réaliser le séquençage des acides nucléiques.

L'intérêt de développer un séquençage des acides nucléiques par nanopore, en plus d'apporter la précision et la sensibilité liées à la mesure en molécule unique, permettrait de limiter les étapes de préparation d'échantillon comme l'utilisation d'enzymes durant les différentes étapes d'amplification et de clonage. De plus, le principe de la technique permet de s'affranchir de l'utilisation de marqueurs fluorescents pour détecter la molécule. Au moins pour ces raisons, l'utilisation de la technologie des nanopores pour le séquençage des acides nucléiques permettrait de réduire le coût de l'analyse en comparaison à des techniques plus traditionnelles comme la méthode de Sanger.

Néanmoins, de nombreuses autres expériences étaient nécessaires avant de parvenir à prouver la faisabilité du séquençage des acides nucléiques par nanopore.

Ces travaux pionniers ont toutefois permis d'établir les premières conditions indispensables à la réalisation du séquençage des acides nucléiques par nanopore, dont l'une d'elle consiste en ce que chacune des bases nucléotidiques (A, T, C, G, et U) puisse être détectée individuellement et fournisse une signature électrique spécifique en termes d'amplitude de blocage du courant. Naturellement, la prochaine question importante était donc de savoir si le nanopore d' α -hémolysine permettrait d'identifier les différentes bases nucléotidiques.

Pour répondre à cette question, des auteurs ont réalisés dans un premier temps des expériences de comparaison d'homopolymères d'adénine (poly-A), de cytosine (poly-C), ou d'uracile (poly-U), de mêmes longueurs [179]. Les résultats ont montré que les brins de poly-A et de poly-C de mêmes longueurs, sont distinguables à la fois par l'amplitude et la durée moyenne des blocages du courant qu'ils provoquent (**Fig. 9**). Les brins de poly-U ne sont eux distinguables des brins de poly-A et de poly-C que par le critère du niveau d'amplitude. Les auteurs ont conclu que l' α -hémolysine peut être utilisé pour distinguer les homopolymères d'ARN soit uniquement sur la base de leur amplitude du blocage du courant (poly-C contre poly-A ou poly-U), soit sur la base de leur durée et de leur amplitude de blocage du courant (poly-A contre poly-U). Il a également été observé que bien que l'adénine soit plus volumineuse que la cytosine, celle-ci engendre des blocages du courant dont l'amplitude est moins profonde. En termes de temps de résidence dans le pore, le poly-C passe moins de temps dans le pore que le poly-A qui est pourtant de la même longueur ce qui suggère que le pore d' α -hémolysine détecte plusieurs bases en même temps et renseignerait sur les structures différentes que peuvent adopter les polynucléotides.

Dans un second temps, une autre expérience a consisté à synthétiser un oligomère de 100 nucléotides composé de 70 cytosines suivies de 30 adénines [179]. L'analyse de l'oligomère avec le nanopore d' α -hémolysine a révélé l'apparition de blocages du courant composés systématiquement de deux niveaux de courants distincts. Il s'est cependant avéré que les variations des niveaux de courant observées au sein des blocages n'étaient pas liées aux propriétés des différentes bases nucléotidiques individuelles, mais plutôt au fait que la partie de l'oligomère contenant les nucléotides de cytosine, forme une hélice simple brin de diamètre étroit qui peut passer à travers l' α -hémolysine sous cette forme, alors que la partie de l'oligomère contenant les nucléotides d'adénine, forme une hélice de plus grand diamètre qui doit être au préalable déroulée pour que le brin unique puisse se déplacer dans le pore.

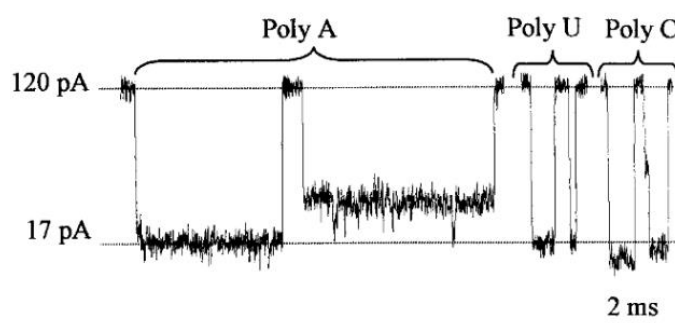


Figure 9 - Discrimination d'homopolymères d'ADN et d'ARN avec le nanopore d' α -hémolysine. Portion d'une trace de courant dans le cas de l'addition d'homopolymères d'adénine (poly-A), d'uracile (poly-U), ou de cytosine (poly-C), ajoutés dans une solution ionique de KCl 1 M, HEPES 5mM, pH 7,5, en présence d'un pore d' α -hémolysine inséré dans une bicouche de lipides, et sous l'action d'une tension électrique de +120 mV. Le poly-A et le poly-C sont chacun caractérisés par des blocages du courant de deux niveaux d'amplitude différents, alors que le poly-U ne présente des blocages du courant que d'un niveau d'amplitude caractéristique. (Adaptée de [179]).

Ces résultats ont été confirmés et améliorés par un autre groupe [180]. Dans ce travail, les auteurs montrent que des homopolymères d'ADN d'adénine (poly-A) et de cytosine (poly-C) de mêmes longueurs (100 nucléotides) peuvent être distingués par trois paramètres statistiques des blocages du courant mesurés qui sont l'amplitude moyenne, la durée moyenne, et la dispersion autour de la durée moyenne. Un critère important identifié pour obtenir de la discrimination concerne la longueur de la chaîne d'ADN analysée, qui lorsqu'elle devient inférieure à la longueur du pore, entraîne une perte de la discrimination.

Ces travaux qui ont montré la possibilité de transporter des homopolymères d'une centaine de nucléotides avec le nanopore d' α -hémolysine, et de les distinguer selon la nature chimique des

nucléotides qui les composent, ont donné lieu par la suite à de très nombreuses études proposant des stratégies de séquençage de l'ADN.

Deux approches ont été proposées pour déterminer la séquence d'un acide nucléique à travers un nanopore : (i) la lecture de la séquence d'un brin d'ADN déplié au cours de sa translocation à travers le nanopore ; ou (ii) la lecture de la séquence d'un brin d'ADN déplié et clivé base par base à l'entrée d'un nanopore.

(i) Lecture de la séquence d'un ADN déplié à travers un nanopore

La première méthode propose d'utiliser le nanopore comme une tête de lecture permettant de lire en temps réel la séquence d'un simple brin d'ADN déplié au cours de son passage à travers le nanopore (**Fig. 10**). Le transport unidirectionnel progressif du brin dans le nanopore permettrait de lire, base par base, la séquence de l'ADN (**Fig. 10**). Bien qu'en théorie l'idée était prometteuse, les chercheurs se sont heurtés à plusieurs difficultés pour obtenir le séquençage de l'ADN avec la technologie des nanopores.

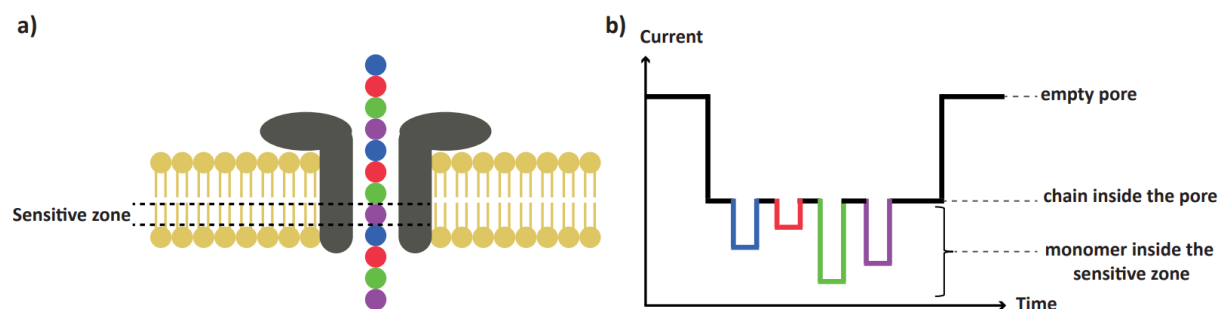


Figure 10 - Représentation schématique du principe de lecture séquentielle d'une chaîne d'ADN dépliée à travers un nanopore. (a) Un nanopore biologique est inséré dans une bicouche de lipides. Une chaîne dépliée traverse le nanopore. (b) Courant électrique en fonction du temps associé au schéma (a) montrant le courant correspondant au pore vide à un premier niveau, le courant correspondant à la chaîne qui entre dans le pore et le traverse, et le courant correspondant à la détection de chaque monomère de la chaîne lors de son passage dans la zone de sensibilité ou de lecture du nanopore.

D'une part, une des premières difficultés rencontrées concerne la vitesse de passage du brin à travers le nanopore qui est trop rapide pour permettre la détection de chaque base nucléotidique avec une résolution temporelle suffisante. Les premières tentatives de lecture d'une séquence d'ADN par nanopore ont montré que la vitesse de translocation d'un polynucléotide à travers un nanopore est de l'ordre de 1 à 10 μ s par base [1], ce qui est bien en-dessous de la résolution temporelle proposée par les dispositifs de détection commercialisés. Le contrôle de la vitesse

de translocation d'un brin d'ADN à travers un nanopore est devenue indispensable à l'obtention d'une résolution de lecture d'un nucléotide pour parvenir à séquencer un brin d'ADN par nanopore.

D'autre part, une autre difficulté concerne la discrimination des bases nucléotidiques pour lesquelles les amplitudes électriques mesurées sont trop proches pour être distinguées et permettre une lecture précise de la séquence.

Pour essayer de ralentir la dynamique du transport d'un brin d'ADN à travers un nanopore, certaines stratégies ont consisté à diminuer la température de l'électrolyte [180], ou à augmenter la viscosité de l'électrolyte [181]. D'autres stratégies ont consisté à faire des modifications des nanopores en y introduisant des modifications génétiques [111,182,183] et en les associant à des adaptateurs moléculaires [184–186], ou à des moteurs moléculaires [187–189].

Les travaux les plus célèbres dans ce sens concernent l'utilisation d'un moteur moléculaire associé à un nanopore biologique pour contrôler la translocation d'un brin d'ADN.

Parmi ceux-là, on peut citer les travaux du groupe de Mark Akeson [187,190–192] dans lesquels les auteurs proposent d'utiliser comme moteur moléculaire, une polymérase (fragment Klenow de l'ADN polymérase I de *Escherichia coli* (KF) [187,190], ADN polymérase du bactériophage T7 [187,190], ADN polymérase du phi29 [191,192]), qui associée à l' α -hémolysine, permettrait de contrôler et de réduire la vitesse de translocation d'un brin d'ADN à travers le nanopore [187]. La polymérase peut se lier à la fois à une amorce d'ADN et au polynucléotide à séquencer. L'extrémité libre du polynucléotide entre dans le pore et entraîne avec lui le complexe formé qui est stabilisé à l'entrée du pore. A chaque nouveau cycle, la polymérase ajoute une base au niveau de l'ADN localisé en dehors du pore, et permet en même temps de faire avancer le polynucléotide à séquencer dans le pore base par base.

Les meilleurs résultats basés sur cette approche ont été obtenus avec l'ADN polymérase du phi29 [191,192]. Une première étude a montré que la polymérase permet de faire avancer le brin d'ADN dans le nanopore à une vitesse d'une centaine de nucléotides par seconde (10 millisecondes/nucléotide) ; ce qui correspond à réduire la vitesse de translocation de l'ADN de quatre ordre de grandeur par rapport à l'ADN qui transloque librement à travers un nanopore d' α -hémolysine [191]. Une deuxième étude est parvenue à ralentir la vitesse de translocation de 2,5 à 40 nucléotides par seconde (25 à 400 millisecondes/nucléotide), ce qui devrait être compatible avec le séquençage [192]. Néanmoins, les auteurs de cette dernière étude n'ont pas

été en mesure d'associer directement les niveaux de courant mesurés à des nucléotides individuels [192].

Un autre paramètre important qui entre en jeu pour atteindre une résolution de lecture d'un nucléotide et son identification, est le nanopore utilisé. Les études précédentes étaient toutes basées sur l'utilisation du nanopore d' α -hémolysine très célèbre en raison de sa stabilité et de sa facilité à être manipulé, notamment pour y introduire des modifications génétiques. Cependant, comme mentionné précédemment, ce nanopore en raison de sa structure, n'est pas le meilleur candidat pour déterminer la contribution spécifique de chacun des nucléotides lors de la lecture d'une séquence. Le nanopore présente une structure en tonneau bêta dont la zone de lecture (ou zone de sensibilité) correspond à une longueur de 5,2 nm (**Fig. 11**) qui permettrait d'accueillir simultanément environ 12 nucléotides [193].

Pour ces raisons, les recherches se sont tournées vers l'utilisation d'un autre nanopore, la porine MpsA de *Mycobacterium smegmatis* qui présente une structure [194] qui convient mieux à la lecture d'une chaîne d'ADN avec un nanopore. En effet, le nanopore dispose d'une zone de lecture d'une longueur d'environ 0,5 nm pour un diamètre d'environ 1,2 nm (**Fig. 11**), plus courte que celle de l' α -hémolysine. Les dimensions de cette zone seraient plus appropriées pour s'approcher de la résolution de lecture d'une base individuelle car cela permettrait d'y accommoder un plus petit nombre de nucléotides (environ 4) et permettrait ainsi de se soustraire des signaux parasites provenant des nucléotides environnants à celui que l'on souhaite identifier.

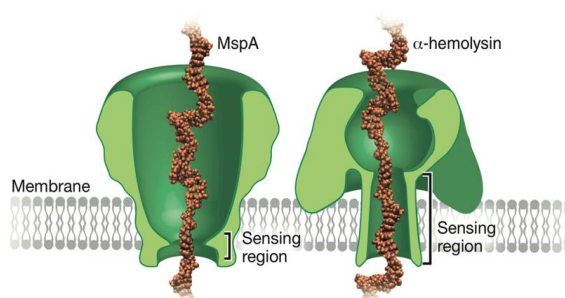


Figure 11 - Schéma représentatif des nanopores biologiques MspA et α -hémolysine et de leurs zones de sensibilité. La zone de sensibilité (ou de lecture) de la MspA est plus courte (0,5 nm) que celle de l' α -hémolysine (5,2 nm). (Adaptée de [195]).

Les travaux de Gundlach *et al.* ont démontré que des mutants de la MspA [111] permettent de distinguer les quatre bases nucléotidiques de simples brins d'ADN et d'obtenir des différences d'amplitudes du courant significativement plus grandes entre les bases en comparaison au

nanopore d' α -hémolysine. Cela a notamment été réalisé grâce à l'utilisation du moteur moléculaire phi29 en association avec la MspA (**Fig. 12**) [188,189,196]. Cette configuration a permis d'obtenir des blocages du courant d'une durée moyenne d'une trentaine de millisecondes par nucléotide, avec des différences d'amplitudes allant d'environ 20 à 50 pA, et a permis d'atteindre une longueur de lecture allant jusqu'à 4,5 kilobases [189]. Malgré l'avantage apporté par la MspA, les niveaux de courant mesurés avec cette méthode ne reflètent toujours par le signal d'un nucléotide individuel. Cette technique qui propose de combiner la polymérase du phi29 et le nanopore MspA a néanmoins donné les meilleurs résultats en termes de résolution et a été exclusivement reprise par Illumina [197].

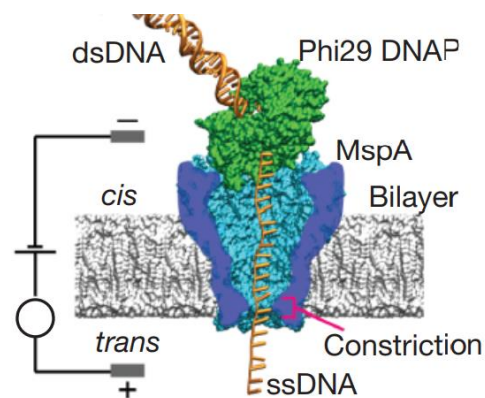


Figure 12 - Représentation schématique du principe de contrôle du transport d'une chaîne d'ADN par nanopore assisté par un moteur moléculaire. Un nanopore biologique, la MspA, est inséré dans une bicouche de lipides. Une molécule d'ADN double brin est capturée par un moteur moléculaire, l'ADN polymérase du Phi29, positionnée à l'entrée du nanopore. Le moteur moléculaire permet de dérouler l'ADN et de contrôler le transport d'un simple brin de cet ADN le long du pore pour la lecture de sa séquence au niveau de la zone de constriction située en partie inférieure du pore. (Adaptée de [189]).

(ii) Lecture de la séquence d'un brin d'ADN séquentiellement clivé à l'entrée d'un pore

En alternative à l'utilisation des moteurs moléculaires, une autre méthode de séquençage des acides nucléiques par nanopore a été proposée. Dans cette configuration, une exonucléase positionnée à l'entrée du nanopore clive un par un les nucléotides du brin d'ADN déplié. Chaque base migre ensuite individuellement vers l'intérieur du nanopore où elle est détectée par la mesure d'une impulsion du courant dont l'amplitude est spécifique de la nature chimique de la base nucléotidique (**Fig. 13**).

Pour cela, une première condition nécessaire à la mise en place de cette approche est de pouvoir détecter la présence d'une particule aussi petite qu'un nucléotide individuel en comparaison au

diamètre du nanopore, c'est-à-dire que sa présence dans le nanopore engendre une impulsion du courant d'amplitude et de durée suffisamment résolues pour que la molécule puisse être détectée. La deuxième condition est de pouvoir distinguer les impulsions du courant engendrées par les différentes bases nucléotidiques. Pour cela, deux approches ont été proposées. La première propose de fixer un élément (adaptateur moléculaire) dans le nanopore auquel pourrait se lier le nucléotide. La deuxième propose d'augmenter le volume du nucléotide en le liant à une chaîne de polymère dont la longueur est spécifique de la nature chimique de la base nucléotidique.

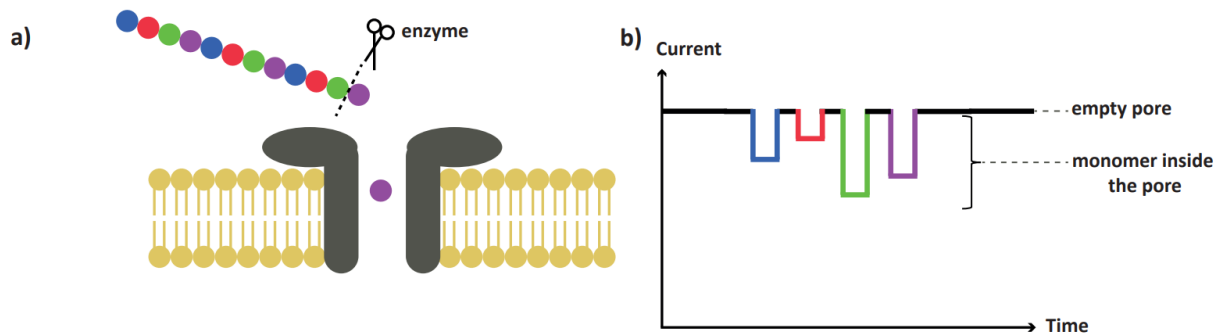


Figure 13 - Représentation schématique du principe de lecture séquentielle d'une chaîne d'ADN clivée base par base à l'entrée d'un nanopore. (a) Un nanopore biologique est inséré dans une bicouche de lipides. Une chaîne dépliée près de l'entrée du nanopore est clivée à son extrémité par une enzyme, base par base. (b) Courant électrique en fonction du temps associé au schéma (a) montrant le courant correspondant au pore vide à un premier niveau, et le courant correspondant aux bases présentes individuellement dans le nanopore après chaque cycle de clivage.

La première étude de Clarke et Bayley *et al.* propose d'utiliser une exonucléase attachée au niveau du chapeau du nanopore d' α -hémolysine. L' α -hémolysine est également modifiée par l'introduction dans la cavité du pore d'une β -cyclodextrine jouant le rôle d'adaptateur moléculaire qui vient réduire le diamètre interne du nanopore (**Fig. 14**) [184,185,198,199]. Les dimensions de la molécule cavitaire que forme la β -cyclodextrine (diamètre externe $\approx 1,53$ nm ; diamètre interne $\approx 0,60$ - $0,64$ nm ; longueur $\approx 0,7$ nm) [200] lui permettent de bien s'accommoder à l'intérieur du canal d' α -hémolysine. De plus, le choix s'est également porté sur la β -cyclodextrine car elle possède une affinité pour se lier à des mononucléotides [201]. Cette approche a permis d'obtenir la détection et la séparation des quatre bases nucléotidiques de l'ADN analysé sur la base des amplitudes de bouchage. Il a été montré que chacun des dNTP (2-désoxyribonucléoside 5'-triphosphates) passant à travers le pore induit un blocage du courant d'amplitude spécifique avec un intervalle de confiance de 99,5%. Les auteurs ont

également montré la possibilité de discriminer la cytosine de la méthyl-cytosine par cette approche. Un problème rencontré avec cette approche est que la probabilité de capture par le nanopore d'une base nucléotidique clivée est trop faible [185].

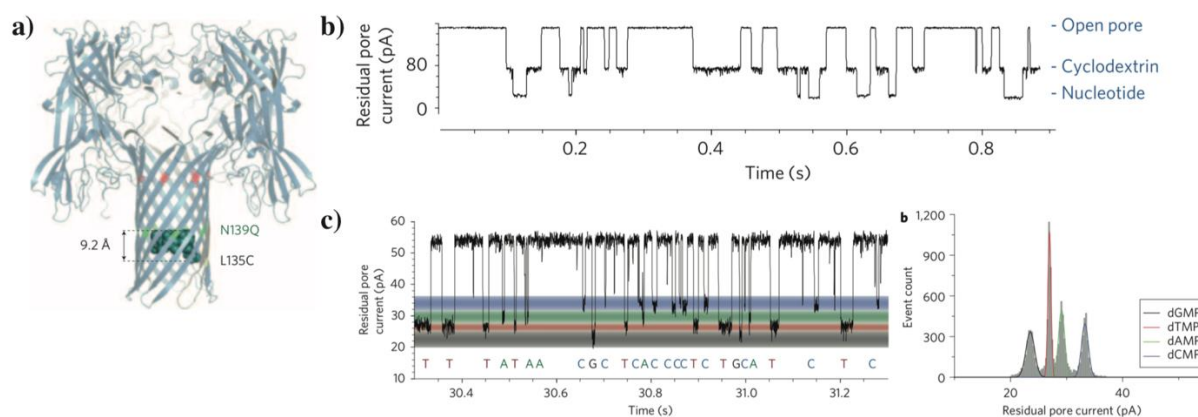


Figure 14 - Approche de discrimination des bases nucléotidiques par l'utilisation d'une β -cyclodextrine comme adaptateur moléculaire associé au nanopore d' α -hémolysine. (a) Représentation d'un nanopore d' α -hémolysine vue de face en présence d'une β -cyclodextrine inséré en position 135. (b) Portion d'une trace de courant électrique en fonction du temps montrant la transition du courant correspondant à un pore vide, au courant correspondant à l'insertion d'une cyclodextrine dans le pore, et enfin, au courant correspondant à l'entrée d'un nucléotide dans le complexe cyclodextrine-pore. (c) Portion d'une trace de courant montrant la détection des 4 nucléotides par un nanopore d' α -hémolysine équipé d'une β -cyclodextrine. Chacune des bases nucléotidiques induit des blocages du courant dont l'amplitude est spécifique et discriminable de celle des autres nucléotides (gauche). Histogramme du courant résiduel montrant 4 populations correspondant aux 4 nucléotides détectés avec le pore équipé de la β -cyclodextrine (droite). (Adaptée de [185]).

La deuxième approche proposée par Kumar *et al.* consiste à utiliser une polymérase qui est immobilisée à l'entrée du pore d' α -hémolysine [202]. Au cours de la synthèse de l'ADN, quatre molécules de dNTP (dATP, dCTP, dGTP, dTTP) chacune marquée respectivement avec des étiquettes correspondant à des complexes formés entre une molécule de coumarin et une molécule de PEG d'une longueur donnée (PEG_n ; n = 36, 24, 20, et 16 monomères) sont incorporés dans le brin en croissance (**Fig. 15**). L'idée d'utiliser des PEG de tailles différentes provient des résultats antérieurs, ayant montré qu'ils sont bien discriminés par le nanopore d' α -hémolysine sauvage, avec une séparation nettement meilleure que celle observée pour l'approche d'interaction des mononucléotides avec une β -cyclodextrine dans le canal d' α -hémolysin précédemment décrite. Les complexes dNTP-PEG sont progressivement libérés et traverseraient séquentiellement le pore, sous l'action d'une tension électrique, pour induire des blocages du courant qui seraient distinguables (**Fig. 15**). Cette méthode (NanoTag-based real-

time Sequencing By Synthesis, NanoTag-SBS) a été notamment reprise pour être développée par Genia Technologies [197]. Le problème rencontré avec cette approche concerne d'une part la fixation de la polymérase à l'entrée du pore qui est extrêmement difficile. Pour faire face à ce problème, les auteurs ont suggéré la possibilité d'utiliser la technologie de fusion des protéines ayant au préalable été utilisée dans une autre méthode de séquençage (de seconde génération) [203]. D'autre part, même si la polymérase parvient à être fixée au nanopore, un autre problème peut concerner la fonctionnalité de la polymérase lorsque celle-ci est attachée au nanopore. Enfin, il est nécessaire de vérifier que le taux de capture du complexe par le nanopore soit suffisamment important pour utiliser cette approche.

Certains de ces travaux ont permis à plusieurs entreprises de développer des stratégies de séquençage basées sur les nanopores [197]. Parmi celles-là, on compte la stratégie d'excision des nucléotides du brin d'ADN et de leur transport un par un à travers un nanopore (NanoTag sequencing, Genia Technologies), ou encore la stratégie de transport d'un brin d'ADN enfilé dans un nanopore dont la séquence est lue progressivement au cours du passage du brin dans le nanopore (MinION, Oxford Nanopore Technologies). C'est cette seconde stratégie qui a eu le plus grand succès et qui a mené Oxford Nanopore Technologies à devenir le leader mondial de la technologie notamment pour le séquençage d'ADN qu'ils ont réussi à développer et à commercialiser [120]. Le séquençage à lecture longue réalisable avec la technologie nanopore présente un avantage technologique important en comparaison au séquençage à lecture courte, car les fragments sont assemblés avec moins d'ambiguïté, ce qui permet de réduire considérablement les possibilités d'erreurs de lecture.

Bien que dans la littérature, le pore ayant eu le plus de succès pour la lecture de la séquence d'une chaîne d'ADN par un nanopore est la MspA, Oxford Nanopore utiliserait un autre nanopore biologique moins connu, le CsgG, en association avec un moteur moléculaire enzymatique. Le dispositif MinION commercialisé par Oxford Nanopore est un séquenceur portable de 90 g composé de cellules de mesures permettant la formation simultanée d'un maximum de 2048 membranes dans lesquelles sont insérées des nanopores [197]. Les canaux peuvent être contrôlés par groupes de 512. Le dispositif utilise des bases de données pour décoder la séquence lue par interprétation des blocages du courant mesurés. La lecture se fait par bloc de six bases et non pas d'une base nucléotidique individuelle [197], et permet d'avoir un débit de 100 bases/seconde et de lire des chaînes d'une longueur allant jusqu'à 50 kilobases [197]. Il existe néanmoins encore un taux d'erreur (30 % en 2016) important [204] lié à des bases de données qui nécessitent d'être améliorées [197], à une lecture trop rapide des

nucléotides (quelques millisecondes) [204], mais également en raison de la présence de plusieurs sites de reconnaissance qui troublent le signal [197]. Ce dernier problème peut être résolu en ne maintenant qu'un seul site de reconnaissance et en supprimant les autres par l'introduction de mutations génétiques dirigées [197,205,206].

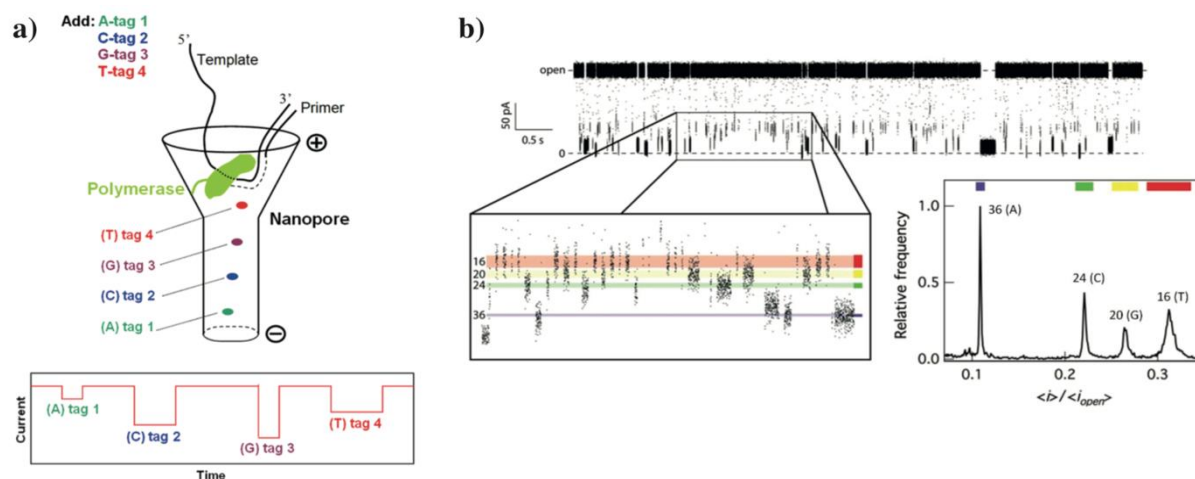


Figure 15 - Approche de séquençage d'ADN par marquage et synthèse en temps réel. (a) Schéma représentatif du principe de séquençage par nanopore d'une chaîne d'ADN dont chacun des nucléotides est marqué par une étiquette spécifique correspondant à une molécule de coumarin liée à une molécule de PEG d'une longueur donnée attachée à l'extrémité 5'-phosphate de chaque nucléotide, et produira dans le pore un blocage du courant d'amplitude spécifique. (b) Discrimination des nucléotides marqués par des étiquettes de coumarin-PEG avec le nanopore d' α -hémolysine. Portion d'une trace de courant et agrandissement d'une partie de cette portion montrant les blocages du courant de différentes amplitudes dont quatre niveaux sont largement reproduits et sont attribués à chacun des nucléotides comme on peut le voir sur l'histogramme du courant relatif résiduel $\langle i \rangle / \langle i_{open} \rangle$ montrant l'apparition de 4 populations discriminables par leur valeur de $\langle i \rangle / \langle i_{open} \rangle$ et correspondant à chaque base nucléotidique. (Adaptée de [202]).

1.4 Résumé

La technologie des nanopores dont le principe repose sur la détection électrique d'une molécule unique traversant un canal de diamètre nanométrique, est un outil fiable, stable, et précis, qui présente un haut potentiel pour la détection et la caractérisation de plusieurs types d'analytes.

La méthode d'impulsion résistive du compteur Coulter d'analyse de particules micrométriques a été dimensionnée avec succès à l'échelle moléculaire, grâce à l'utilisation de nanopores biologiques qui sont parvenus à être insérés dans des bicouches de lipides reconstituées manuellement et à être associés à des dispositifs électriques.

L'existence de différents types de nanopores, aux dimensions variables et aux configurations multiples, permet d'adapter le système pour l'analyse d'une large gamme d'objets. Nous avons exposé dans cette partie le rôle important qu'ont joué les nanopores biologiques dans les premières démonstrations expérimentales de la détection et de la caractérisation d'une molécule par un nanopore.

Les expériences pionnières ont montré que, selon les conditions physico-chimiques choisies, les nanopores peuvent être utilisés pour discriminer des polymères synthétiques selon leur masse ou leur taille, et, pour détecter des chaînes d'ADN et d'ARN et discriminer des nucléotides selon leur nature chimique.

Les travaux de détection et de caractérisation de chaînes d'acides nucléiques ont donné suite au développement de nombreux travaux ayant abouti au succès du séquençage de l'ADN par nanopore. La démonstration du séquençage de l'ADN a été réalisée en utilisant un régime qui consiste à lire la séquence d'un brin d'ADN de longueur supérieure à la longueur du pore, durant sa translocation à travers un nanopore biologique. Cette approche s'est heurtée à de nombreuses difficultés concernant la vitesse de transport d'une chaîne d'ADN à travers un nanopore, la résolution de lecture d'un nucléotide individuel, et la distinction des différents nucléotides. Ces difficultés ont été contournées principalement par l'utilisation de moteurs moléculaires qui ont permis de ralentir la translocation des chaînes d'ADN à travers un nanopore mais pas suffisamment pour atteindre une précision de lecture d'un nucléotide.

Malgré ces difficultés, quelques-uns des très nombreux travaux menés pour démontrer la faisabilité du séquençage de l'ADN par nanopore ont mené à l'acquisition d'un dispositif industriel commercialisé. Ce dispositif, bien qu'il nécessite d'être amélioré pour gagner en fidélité, propose d'utiliser un nanopore biologique associé à un moteur moléculaire qui permet

de suffisamment ralentir le transport d'une chaîne pour déchiffrer la séquence d'une chaîne d'ADN à partir de bases de données.

Le succès du séquençage de l'ADN par nanopore a permis de montrer le potentiel de la technique d'analyse par nanopore à l'échelle de la molécule unique et a mené à s'interroger sur la faisabilité du séquençage de protéines qui est le prochain grand défi de la technologie. Il s'avère être bien plus complexe que celui de l'ADN en raison des propriétés physico-chimiques des acides aminés comme le nombre d'acides aminés, l'hétérogénéité de charge des protéines, ou encore la présence de modifications post-traductionnelles. Les stratégies mises en place doivent intégrer ces complexités et les contourner.

La partie suivante consiste à présenter les progrès en cours vers l'acquisition du séquençage de protéines par nanopore protéique, et les stratégies développées pour y parvenir.

II. Progrès vers le séquençage de protéines par nanopore

La première démonstration du séquençage de protéines a été proposée dans les années 50 par Frederick Sanger qui détermine pour la première fois la structure primaire de l'insuline et reçoit pour cette découverte un prix Nobel en 1958 [207]. Pour arriver à réaliser cela, Sanger et ses collègues utilisent environ 100 mg de protéine et travaillent au moins 10 ans. Par la suite, trois méthodes se sont distinguées pour déterminer la séquence d'une protéine.

Jusqu'au début des années 1990, le séquençage des protéines était principalement réalisé par la dégradation d'Edman [208] qui consiste en une série de réactions chimiques aboutissant à éliminer le résidu présent à l'extrémité N-terminal de la protéine et préalablement marqué, sans rompre les liaisons peptidiques internes [209,210]. Cependant, bien qu'elle ait été automatisée [211,212], la méthode n'est pas généralisable à l'ensemble des protéines, est lente à mettre en œuvre, et est coûteuse car elle nécessite dans certains cas d'être associée avec des dispositifs complémentaires comme par exemple des appareils de chromatographie liquide de haute pression ou en phase gazeuse pour parvenir à identifier les acides aminés.

Par la suite, le succès du séquençage du génome humain a permis le développement d'algorithmes informatiques qui ont permis la prédiction des séquences protéiques à partir des séquences d'ADN et qui ont conduit à la génération de grandes bases de données [213]. Cependant, cette deuxième méthode ne peut pas prédire avec précision la séquence de l'ensemble des protéines car le protéome est une entité bien plus complexe et dynamique que le génome, dans lequel de nombreux mécanismes surviennent directement à l'échelle de la protéine. On peut citer d'une part, le processus d'épissage alternatif à l'origine de l'hétérogénéité de la traduction des ARNm qui aboutit à la synthèse de plusieurs protéines à partir d'un même gène. Il a été montré qu'environ 2% du génome code pour les protéines, ce qui correspond à environ 20 000 à 30 000 gènes alors que les protéines sont de l'ordre du million [214,215]. D'autre part, les protéines peuvent également subir des modifications chimiques essentielles à leur structure et à leur fonction survenant après la traduction (modifications post-traductionnelles) qui ne peuvent pas être détectées à l'échelle du génome.

Actuellement, le séquençage de protéines repose principalement sur l'utilisation de la spectrométrie de masse qui est devenue la technique de référence [12,13] la plus répandue pour ce type d'analyse et qui s'est principalement distinguée par deux approches [216]. La spectrométrie de masse s'est révélée significativement plus sensible et plus rapide que le

séquençage d'Edman ou de Sanger. Sur la base de l'analyse des propriétés de masse et de charge des molécules, une large gamme de protéines peut être analysée, et il devient également possible de détecter certaines modifications post-traductionnelles qui produisent des variations de masse faciles à identifier (*e.g* phosphorylation, hydroxylation). Ces modifications sont en particulier difficiles à détecter par la méthode d'Edman, et indétectables dans les séquences protéiques prédites à partir des séquences d'ADN [9].

Néanmoins, bien qu'elle soit devenue un outil puissant pour l'analyse des protéines, la spectrométrie de masse présente plusieurs limitations [9] parmi lesquelles on compte les étapes complexes de préparation d'échantillon, la limite de résolution et de plage dynamique qui rendent l'identification et la quantification simultanée d'échantillons protéiques complexes, la limite de sensibilité de détection (proche du femtomole) qui ne permet pas de détecter des protéines présentes en très faibles quantités, ou encore l'utilisation complexe de bases de données incomplètes. De plus, la technique ne permet pas l'analyse en temps réel d'une protéine à l'échelle de la molécule unique.

Plusieurs méthodes d'analyse à l'échelle de la molécule unique ont été développées au cours des dernières années dans le but de parvenir au séquençage des protéines. Parmi celles-ci, on retrouve la technique de détection électrique par nanopore précédemment décrite pour le séquençage de l'ADN. Elle serait une des perspectives les plus intéressantes pour le développement du séquençage des protéines. Nous présenterons à la suite, dans un premier temps, les approches de séquençage de protéines par nanopore, et dans un second temps, quelques exemples d'autres approches de séquençage des protéines à l'échelle de la molécule unique.

2.1 Séquençage des protéines par nanopore

Depuis plusieurs années, la technologie des nanopores a été utilisée avec succès pour l'analyse des protéines à plusieurs échelles.

On peut citer les travaux consacrés à l'étude de la dynamique du transport d'une protéine dépliée à travers un nanopore biologique [24,217–220], ou solide [160,221–223], et au contrôle de sa translocation (entrée et transport) en jouant sur les forces motrices et les conditions expérimentales comme le flux électroosmotique, la tension électrique, et la température [224].

D'autres études ont été réalisées sur le dépliement de protéines sauvages et mutantes [218,225], sur le désordre intrinsèque des protéines [226–228] ; sur l'agrégation de protéines impliquées dans des maladies humaines conformationnelles [226,227,229,230].

Les interactions protéine-ligand, protéine-acide nucléique ou protéine-anticorps [231], et la conformation de protéines natives [232,233] ont également été explorées.

Bien que les protéines et peptides ont largement été analysés par des nanopores, aucune méthode n'est parvenue à réaliser le séquençage d'une protéine par nanopore tel qu'il est possible de le faire avec les acides nucléiques.

La stratégie naturellement suivie pour les protéines est celle qui a mené au succès du séquençage de l'ADN par nanopore et qui consiste à lire la séquence d'une protéine qui est dépliée au cours de son transport unidirectionnel à travers un nanopore. Cependant, les chercheurs se sont heurtés à plusieurs difficultés lorsque la même approche a tenté d'être utilisée pour les protéines. Ces difficultés sont dues aux nombreuses contraintes liées à la structure tridimensionnelle des protéines et à leurs propriétés physico-chimiques. Au contraire de l'ADN qui est uniformément chargé et qui est composé de quatre bases nucléotidiques (A, T, C, G), les protéines présentent une distribution de charge hétérogène, et sont constituées d'au moins 20 acides aminés qui peuvent présenter des modifications post-traductionnelles.

Les défis à relever pour aboutir au séquençage des protéines concernent : (i) le contrôle du dépliement d'une protéine et de son entrée dans un nanopore ; (ii) le contrôle du transport unidirectionnel de la chaîne polypeptidique dans un nanopore ; (iii) le contrôle de sa vitesse de translocation ; (iv) l'obtention d'une résolution de lecture d'un acide aminé ; (v) l'obtention de la distinction des 20 acides aminés naturels ; et (vi) la détection et la discrimination des modifications post-traductionnelles.

Plusieurs stratégies ont été mises en œuvre pour contourner ces difficultés et sont présentées à la suite.

2.1.1 Contrôle du transport unidirectionnel

Pour contrôler le transport unidirectionnel d'une protéine dans un nanopore, une stratégie a consisté à lier des courtes chaînes d'oligonucléotides ($n = 30$ cytosine ou adénine) à l'extrémité C-terminal de la protéine thioredoxin de *Escherichia coli* dans le but d'orienter son entrée et sa translocation à travers le nanopore d' α -hémolysine (**Fig. 16**). L'idée d'utiliser un

oligonucléotide provient du fait qu'il s'agit d'un simple brin d'ADN court négativement chargé qui peut, d'une part, être entraîné vers l'entrée du pore sous l'action d'une tension électrique positive et ainsi faciliter l'entrée de la protéine dans le pore ; d'autre part, apporter une force motrice à la fois suffisante pour le dépliement de la protéine et les étapes initiales de la translocation (estimée à 5-10 pN) ; et enfin, empêcher un mouvement de retour en arrière de la protéine liée à l'oligonucléotide. Avec cette approche, les auteurs ont observé que sous l'action d'une tension électrique positive, l'oligonucléotide s'insère dans la lumière du nanopore et exerce une force sur la protéine native, entraînant le dépliement de son extrémité C-terminal tandis que le reste de la protéine se déplie spontanément et diffuse à travers le pore sous l'action de la force électrique appliquée (**Fig. 16**) [218]. Les expériences de liaison du même oligonucléotide à l'extrémité N-terminal de la protéine, ont révélé que les étapes de dépliement sont différentes selon l'extrémité par laquelle la protéine entre dans le pore [234].

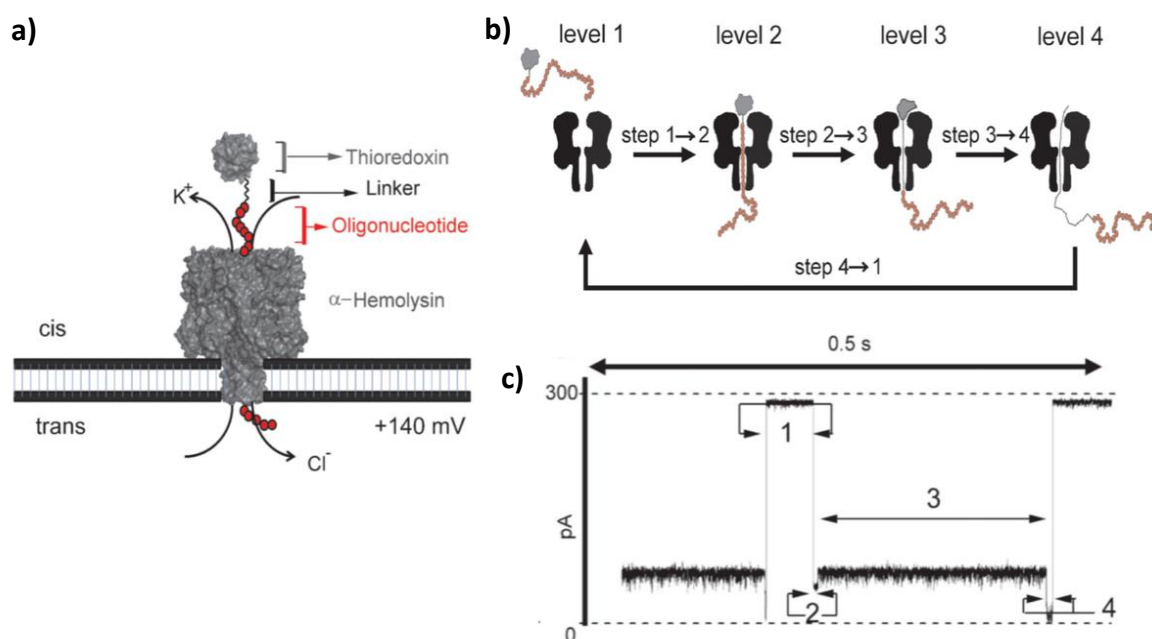


Figure 16 - Stratégie de liaison d'un oligonucléotide à une protéine pour son transport unidirectionnel à travers le nanopore d' α -hémolysine. (a) Schéma illustrant un nanopore d' α -hémolysine inséré dans une membrane lipidique. Une chaîne de 30 oligonucléotides (cytosine) liée à l'extrémité C-terminal d'une protéine thioredoxin, traverse le nanopore et tire la protéine dans la direction de l'entrée du nanopore. (b) Schéma représentatif des différentes étapes de capture et de transport de la protéine par l'utilisation de l'oligonucléotide, et (c) trace de courant associée. Niveau 1 : La protéine liée à l'oligonucléotide est ajoutée en solution du côté cis du pore d' α -hémolysine qui est vide. Niveau 2 : L'oligonucléotide entre dans le nanopore et exerce une force sur la protéine. Niveau 3 : La force exercée sur la protéine entraîne le dépliement partiel de celle-ci, ce qui permet au nucléotide de transloquer à travers le nanopore et de faire entrer la protéine dans le pore. (Adaptée de [218]).

2.1.2 Contrôle du dépliement et de la translocation dans un nanopore

Une autre difficulté majeure provient du fait que la majorité des protéines natives repliées sont trop volumineuses pour pouvoir entrer et être transportées dans un nanopore biologique de diamètre comparable à celui de l' α -hémolysine. La première idée était de travailler avec des protéines préalablement dépliées qui seraient plus adaptées aux dimensions de la cavité des nanopores biologiques utilisés.

Les études de dépliement d'une protéine ont été réalisées en présence de nanopores biologiques et de nanopores solides par l'utilisation d'agents chimiques dénaturants (*e.g* guanidium hydrochloride et urée), ou par la variation de paramètres physiques comme la température ou la tension électrique. La preuve de la translocation des protéines dépliées avec un nanopore protéique (aérolysine) a constitué un point important vers le développement du séquençage des protéines [24,220]. Ces travaux ont également montré la dépendance de la dynamique de la translocation en fonction de la tension électrique et de la taille de la protéine, avec le nanopore biologique d'aérolysine [24], et avec un nanopore solide [235].

Cependant, lorsque la protéine est dépliée, elle passe dans le pore à une vitesse trop rapide (quelques millisecondes) ne permettant pas de lire sa séquence avec une précision de détection d'un acide aminé [236].

En réponse à ce problème, de façon analogue à ce qui avait été proposé dans le cas de l'utilisation de la polymérase du phage T7 pour le séquençage de l'ADN, l'approche a été d'introduire l'utilisation d'un moteur moléculaire enzymatique pour contrôler le dépliement et la vitesse de translocation d'une protéine à travers un nanopore.

Plusieurs études ont tenté d'utiliser le moteur moléculaire ClpX dans le but de parvenir au séquençage des protéines (**Fig. 17**). Le moteur ClpX, en présence d'ATP, possède la capacité de déplier des protéines substrats repliées et marquées par des motifs de reconnaissance en exerçant sur elles une force mécanique de l'ordre de 20 à 30 pN qui est suffisante pour diminuer la stabilité structurale de la protéine, ce qui aboutit à son dépliement mécanique et à sa translocation à travers la lumière du pore. L'utilisation de ClpX présente aussi l'intérêt de pouvoir atteindre une vitesse de translocation (jusqu'à 80 acides aminés par seconde) appropriée à l'analyse de la structure primaire de la protéine étudiée par le nanopore. Il a été montré que le moteur déplie efficacement la protéine et permet son transport, mais ne permet pas la lecture des acides aminés [219].

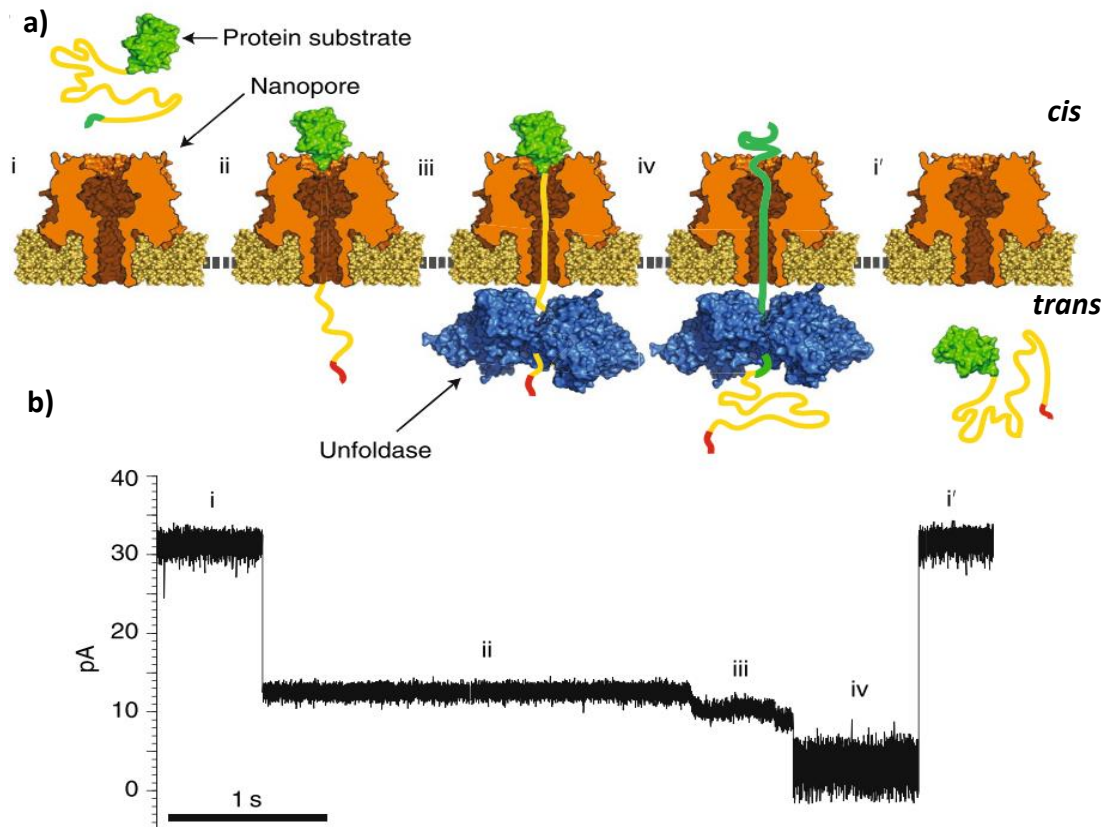


Figure 17 - Stratégie de contrôle de la translocation d'une protéine à travers un nanopore par association d'un moteur moléculaire au nanopore. (a) Schéma représentatif d'un nanopore d' α -hémolysine inséré dans une membrane de lipides en présence d'une protéine substrat présente du côté cis (chapeau) du pore ; et d'une unfoldase ClpX présente du côté trans (tige) du pore ; et des différentes étapes de translocation de la protéine médiée par l'unfoldase. (b) Portion d'une trace du courant électrique en fonction du temps associée aux étapes de translocation de la protéine. (i) Un courant d'environ 30 pA circule à travers le nanopore d' α -hémolysine vide sous l'action d'une tension électrique de +180 mV (ii) Le substrat lié à la protéine d'intérêt est capturé par le nanopore ce qui induit une chute du courant ionique à une valeur d'environ $14 \pm 0,7$ pA. (iii) L'unfoldase qui reconnaît le substrat exerce une force mécanique entraînant le dépliement et la translocation du domaine de liaison à la protéine (Smt3) à travers le nanopore à l'origine d'une légère baisse du niveau de courant qui fluctue. (iv) Cela entraîne le dépliement et la translocation de la protéine qui sont observés par une nouvelle chute du courant ionique à un niveau d'environ $3,8 \pm 1,7$ pA. (i') La translocation complète de la protéine permet de retrouver le niveau initial du courant ionique correspondant au nanopore vide. (Adaptée de [219]).

2.1.3 Détection et discrimination des acides aminés

En plus de la vitesse de transport qui est trop rapide pour atteindre une résolution temporelle de lecture d'un acide aminé, une autre limite provient du fait que le signal de l'acide aminé se trouvant au niveau de la zone de lecture (ou zone de sensibilité) est altéré par celui des acides

aminés voisins. Concernant les nanopores biologiques, une explication à cela provient très probablement du fait que le volume du pore est trop important par rapport au volume d'un acide aminé.

Une alternative pourrait être l'utilisation de nanopores solides d'une faible épaisseur de 1 nm qui peuvent être produits dans des membranes de graphène, mais qui ne permettraient pas d'écarter la détection des acides aminés voisins car le volume de la zone de détection resterait trop important. Néanmoins, certains auteurs ont démontré la possibilité de produire des nanopores solides de diamètre subnanométrique dans des fines membranes inorganiques dont les dimensions, comparables à celle d'un acide aminé, permettraient d'en atteindre la précision de lecture. Par cette approche, les auteurs ont mesuré des fluctuations du courant dont le nombre observé correspond au nombre d'acides aminés de la protéine. Cependant, les niveaux de courant mesurés correspondaient eux à des quadrimères d'acides aminés. La meilleure résolution ainsi obtenue avec cette méthode est par bloc de quatre acides aminés [21,237,238].

En résumé, bien que de nombreux efforts aient été réalisés pour repousser les limites de cette approche, celle-ci n'a jamais permis d'aboutir au séquençage de protéines par nanopore.

En dehors du contexte direct du séquençage de protéines, de nombreuses études se sont intéressées à l'analyse et à la comparaison de protéines ou de peptides à la recherche d'une amélioration croissante de la résolution pouvant être atteinte avec les nanopores.

A l'échelle de la protéine, plusieurs études basées principalement sur l'utilisation de nanopores solides ont consisté à évaluer la taille de protéines natives. Des pores solides formés par des pipettes en verre ont été utilisés pour analyser des protéines ayant des poids moléculaires compris entre 12 et 480 kDa [239,240]. Par cette approche, la meilleure résolution de discrimination obtenue est de 4 kDa dans un cas, et de 50 kDa pour les protéines dans un autre cas [239,240]. Par la suite, l'utilisation de nanopores solides a permis de démontrer que la détection des protéines peut être abaissée à 8,5 kDa [241] avec une résolution de séparation entre deux protéines de 1,5 kDa [233].

Les premiers travaux d'analyse de peptides concernent la caractérisation d'éléments de structures secondaires comme des hélices alpha ou des épingles bêta, mais également l'étude de l'interaction de ces peptides avec le nanopore utilisé. Des structures en hélices alpha [242] ont été étudiées à travers les pores d' α -hémolysine et d'aérollysine. Par la suite, des structures en épingles bêta [243], en hélices de type collagène [244], des hélices alpha [245], et des motifs

en doigts de zinc (zing finger) ce qui inclut des feuillets bêta, des hélices alpha, et des tours, ont été étudiées à travers le pore d' α -hémolysine [246]. Plus récemment, les recherches se sont tournées vers l'identification de peptides par nanopore et sont plus particulièrement motivées par l'implication potentielle de nombreux peptides comme marqueurs biologiques dans l'apparition de cancers, de troubles neurodégénératifs, et d'infections, ainsi que par des corrélations établies entre leurs niveaux d'expression et les maladies.

Les propriétés de transport de peptides positivement chargés d'arginine d'une longueur de 3, 5, et 7 acides aminés, ont été étudiées avec le nanopore NfpA/NfpB de *Nocardia farcinica* [110]. Des études cinétiques réalisées ont permis de montrer qu'il est possible de distinguer les événements de blocage du courant résultants de l'interaction des peptides avec le nanopore de ceux résultants de la translocation des peptides à travers le nanopore [110]. Concernant les peptides d'arginine analysés, bien que l'étude n'avait pas pour objectif direct de démontrer la discrimination des peptides en longueur, on peut néanmoins observer que dans certaines conditions expérimentales, les blocages du courant induit par les peptides de 5 arginines sont distinguables de ceux de 7 arginines en termes d'amplitude [110].

Les travaux de Asandei *et al.* ont démontré que le nanopore d' α -hémolysine permet de détecter des différences par bloc de 3 et 6 acides aminés d'un polypeptide [247]. La construction peptidique est constituée de 12 acides aminés de charge positive (arginine) à une extrémité du peptide et de 12 acides aminés de charge négative à l'autre extrémité (acide glutamique). Cette construction a été établie dans le but de contrôler l'entrée du peptide et son maintien dans le nanopore. Entre ces deux extrémités, les auteurs ont construit une séquence de 6 acides aminés qui correspond à 6 alanines (peptide 1), 6 tryptophanes (peptide 2), ou 3 alanines et 3 tryptophanes (peptide 3). Les auteurs ont montré que chacune des trois séquences étudiées permet d'obtenir des signatures du courant distinguables entre elles.

Dans les travaux de Ji *et al.*, les auteurs ont utilisé le moteur moléculaire phi29 pour détecter des plus gros peptides comme le transactivateur de la transcription TAT (1,66 kDa), MAR (1,83 kDa), et des molécules d'adhésion des cellules épithéliales EpCAM (3,01 kDa) [248]. Ils ont trouvé des distributions de blocage distinctes pour chacun de ces peptides, et ont confirmé la translocation du peptide TAT en le marquant par fluorescence avec la molécule Cy3, et en montrant le dénombrement du nombre de molécules fluorescentes détectées dans le compartiment opposé à celui où a été ajouté le peptide.

Une autre étude est celle de Chavis *et al.* dans laquelle les auteurs ont utilisé des conditions expérimentales comprenant l'utilisation de nanoparticules d'or, de Guanidium Hydrochloride (Gdn-HCl), et de conditions de pH contrôlées, pour optimiser la détection et la séparation d'une série de peptides neutres et cationiques [249]. Dans cette étude, l'objectif des auteurs était de démontrer que la même approche permet également d'estimer avec plus de précision et avec plus de résolution la masse des peptides sélectionnées. Cinq peptides d'intérêt biologique ont été choisis, dont l'encéphaline (LE), l'angiotensine I (A1), l'angiotensine II (A2), le polyglutamine binding peptide (QBP1), et la neurotensine (NT). Il a été montré que l'association de nanoparticules d'or avec le nanopore d' α -hémolysine permet d'augmenter le temps de résidence des peptides dans le pore, mais n'a pas nécessairement permis d'améliorer la distinction des peptides qui sont très bien séparés même en l'absence des clusters d'or (**Fig. 18**). Les peptides sont identifiables à partir de l'amplitude spécifique des blocages du courant qu'ils induisent, ce qui leur permet d'être distingués dans un mélange.

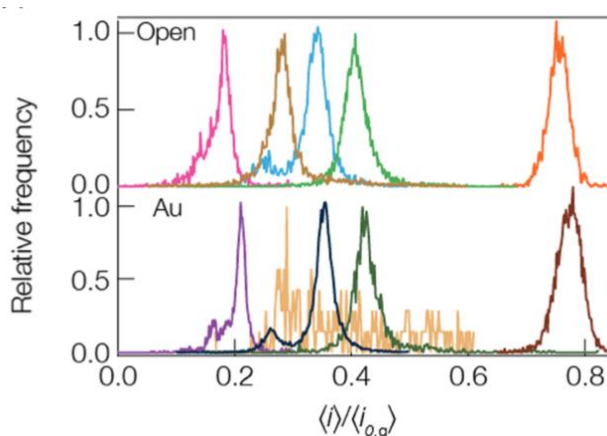


Figure 18 - Discrimination de peptides biologiques par le nanopore d' α -hémolysine en présence ou non de nanoparticules d'or. Représentation des histogrammes du courant relative $\langle i \rangle / \langle i_{o,g} \rangle$ dans le cas de l'analyse de 5 peptides différents dont NT (rose), QBP1 (doré), A1 (bleu), A2 (vert), et LE (orange), qui sont chacun identifiables par une population individuelle sur l'histogramme à une valeur de $\langle i \rangle / \langle i_{o,g} \rangle$ caractéristique, en absence (haut, Open) ou en présence (bas, Au) de nanoparticules d'or. (Adaptée de [249]).

Le canal protéique Fragaceatoxin C (FraC) a récemment été utilisé pour détecter et discriminer des peptides et petites protéines recombinantes d'intérêt biologique (biomarqueurs) de masses moléculaires comprises entre 1,3 kDa et 25 kDa [90]. Dans cette étude, les auteurs se sont intéressés à l'analyse de la chymotrypsine (25 kDa, 245 acides aminés), de l'endothéline, (2,5 kDa, 21 acides aminés), de la β -microglobuline (11,6 kDa, 99 acides aminés), de l'EGF humaine (6,2 kDa, 53 acides aminés), et de l'angiotensine I (1,3 kDa, 10 acides aminés). Il a

été montré que tous les peptides sont entraînés dans le pore par électroosmose et sont détectés par la forme sauvage du nanopore FraC dans des conditions de pH de 4,5, alors qu'à pH 7,5, tous ne le sont pas, ou alors deviennent détectables à partir de différence de potentiel élevées (-200 à -300 mV) [90].

Chacun des peptides a été analysé individuellement et pour chacun d'eux, il a été possible d'identifier une amplitude et une durée moyenne des blocages du courant spécifiques. L'analyse d'un mélange de trois des peptides, la bêta-microglobuline, l'EGF, et l'endothéline 1, a révélé que les trois signaux spécifiques provoqués par chacun des peptides correspondent à des populations distinguables en termes d'amplitudes comme en termes de durées des blocages du courant. Ces résultats indiquent que la meilleure résolution atteinte avec la discrimination de ces trois peptides est de 3,7 kDa. Pour aller plus loin, les auteurs ont comparé l'endothéline 1 et l'endothéline 2, qui présentent une différence d'un seul acide aminé sur vingt et un résidus, mais qui n'est pas localisé à la même position sur le peptide, et ont montré que ces peptides sont distinguables en termes d'amplitudes des blocages du courant (**Fig. 19**)

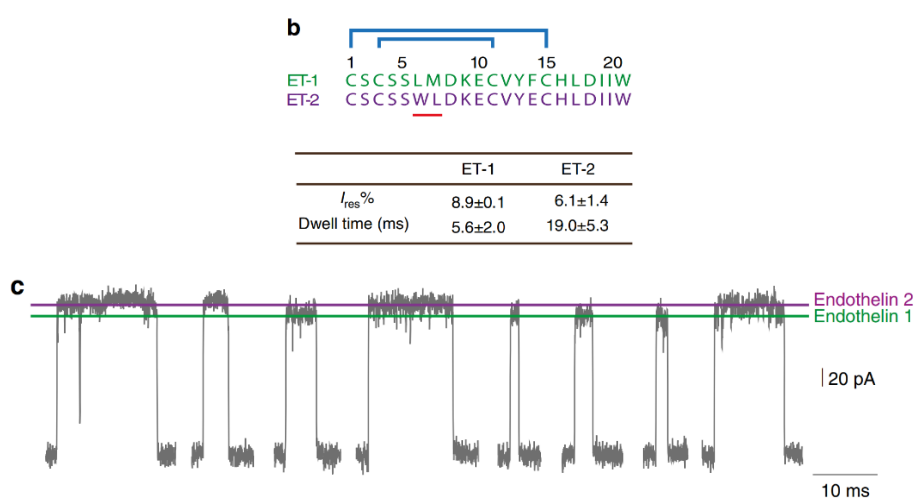


Figure 19 - Discrimination de peptides par le nanopore FraC sauvage à pH 4,5. (b) Séquence en acides aminés des peptides endothéline 1 (ET-1) et endothéline 2 (ET-2). Les lignes bleues indiquent les ponts disulfures existants dans les peptides. Le tableau en dessous indique le courant résiduel (I_{res}) et la durée des blocages du courant induit par chacun des peptides à pH 4,5, sous l'action d'une tension électrique de -50 mV, et dans une solution ionique de 1 M en KCl. (c) Portion d'une trace de courant dans le cas de l'analyse d'un mélange des deux peptides en présence du nanopore FraC à pH 4,5, pour une tension électrique de -50 mV. Les peptides ET-1 et ET-2 sont distinguables dans ces conditions à la fois par leur amplitude et par leur durée de blocage du courant. (Adaptée de [90]).

Cependant, l'étude basée ici sur l'utilisation du FraC ne permet pas d'analyser des peptides d'une taille plus petite que 1,6 kDa car ceux-là traversent le pore trop rapidement pour être détectés, confirmant qu'il faudrait utiliser un nanopore de diamètre plus petit que celui du FraC pour détecter des peptides de plus petits poids moléculaire. La réduction du diamètre du pore par mutagenèse du FraC a permis d'identifier des peptides de plus petites tailles comme l'angiotensine A et II [90].

Les canaux de lysénine ont été utilisés pour détecter les peptides de l'angiotensine 2 comme décrit par les travaux de Shreshta *et al.* [250]. Les auteurs ont vérifié la translocation à travers le pore en effectuant une analyse de trente six heures avec 22700 canaux, puis ont effectué une analyse de chromatographie liquide couplée à la spectrométrie de masse (LC-MS) de la solution récupérée du côté *trans* du pore pour prouver la présence du peptide.

A l'échelle de l'acide aminé, la seule étude ayant démontré la possibilité de détecter un acide aminé individuel a été réalisée avec un nanopore d' α -hémolysine chimiquement modifié [251]. Dans ces travaux, Bayley et ses collègues utilisent le nanopore d' α -hémolysine dans lequel un groupement chimique phénantroline est fixé à l'intérieur du pore par modification chimique [251]. Le but de l'ajout de ce groupement chimiquement est de garder l'acide aminé capturé dans le pore un temps suffisamment long pour être détecté. Les auteurs sont parvenus avec cette approche à détecter cinq acides aminés (L-Tyr, L-Cys, L-Phe, L-Asp, et L-Trp) et leurs énantiomères (D-Tyr, D-Cys, D-Phe, D-Asp, et D-Trp) [251]. Ils ont particulièrement montré que les énantiomères du même acide aminé sont distinguables dans chaque cas. Concernant la distinction des différents acides aminés de même forme, il a été démontré qu'au moins trois acides aminés (L-Phe, L-Cys, et L-Asp) sont clairement discriminables à partir de l'amplitude du blocage du courant qu'ils induisent par leur entrée dans le nanopore d' α -hémolysine modifié [251]. Les mêmes auteurs ont également utilisé cette approche pour montrer la détection de neurotransmetteurs (*e.g* dopamine, norépinephrine, épinéphrine) et d'autres acides aminés comme le glutamate [252].

En dehors de ces deux études, aucun nanopore qu'il soit biologique (sauvage ou modifié), ou solide, n'a permis d'atteindre une résolution de détection d'un acide aminé individuel au commencement de mes travaux de thèse.

2.1.4 Détection des modifications post-traductionnelles

Les travaux de Rosen, Bayley *et al.* portant sur l'étude de la thioredoxine avec le nanopore d' α -hémolysine avaient permis de détecter la forme phosphorylée de la protéine dénaturée et de la distinguer de la forme non phosphorylée en caractérisant la profondeur des blocages du courant et le bruit électrique mesuré lors de l'étape de translocation [248]. Les protéines mutantes seraient identifiées par leur état de phosphorylation (non phosphorylée, monophosphorylée, ou diphosphorylée) et l'emplacement de la modification. En général, il a été observé que la phosphorylation diminue le bruit électrique mesuré et entraîne des blocages d'amplitude moins profonde [248].

Deux études ont été publiées au début de mes travaux de thèse, pour la détection directe de la protéine ubiquitine et de l'ubiquitination de protéines en utilisant soit un nanopore solide [241], soit un nanopore biologique [253].

La première étude est basée sur l'utilisation d'un nanopore solide de 3 nm de diamètre d'une épaisseur inférieur à 10 nm formé dans une membrane de SiNx pour la détection de la protéine ubiquitine de 8,5 kDa [241]. Dans cette étude, les auteurs ont montré que moduler le niveau de pH leur permet de détecter et de discriminer une protéine individuelle d'ubiquitine et des polymères composés de deux (dimère) et de cinq (pentamère) protéines d'ubiquitine [241]. Il a également été démontré qu'il est possible de discriminer deux dimères d'ubiquitine de même poids moléculaire qui ne diffèrent que par la position de la lysine *via* laquelle les deux protéines sont liées (Lys 48 *versus* Lys63) [241].

Dans la deuxième étude, le nanopore biologique ClyA modifié génétiquement (ClyA-AS) a été utilisé pour suivre l'ubiquitination d'une protéine en temps réel [253]. Pour réaliser cela, le choix des auteurs s'est porté sur la protéine Ubiquitin conjugated-enzyme (Ubc4, 19,8 kDa) de *Saccharomyces cerevisiae* qui possède la capacité de s'ubiquitiner elle-même *in vitro* en présence d'ATP. L'analyse de la protéine avec le nanopore ClyA-AS en présence d'ATP a permis d'observer plusieurs niveaux de courant associés à la forme non modifiée, mono-ubiquitinée, ou di-ubiquitinée [253]. Le nanopore ClyA-AS a également été modifié en remplaçant un résidu tryptophane par une glutamine en position 56, ce qui a permis d'augmenter les temps de résidence de la protéine Ubc4 dans le pore d'au moins dix fois, et a potentiellement permis d'améliorer la résolution de séparation notamment pour la comparaison de deux protéines isomères ubiquitinées [253]. Le pore modifié a ensuite été utilisé pour détecter l'addition d'Ub à Ubc4. Lorsque l'ATP est ajoutée à la solution, la protéine Ubc4 est

modifiée par l'ajout d'une ou de plusieurs molécules d'ubiquitine et les différentes formes sont détectées et caractérisées de manière unique [253]. En raison de la nature de la protéine utilisée, les étapes d'ubiquitination en cascade ont été observées en temps réel [253].

2.2 Autres approches de séquençage de protéines à l'échelle de la molécule unique

2.2.1 Séquençage des protéines par courant tunnel

La méthode de détection par courant tunnel se base sur l'utilisation de la microscopie à effet tunnel (Scanning Tunneling Microscope, STM). Le courant tunnel est créé entre deux électrodes métalliques séparées par un espace vide de quelques nanomètres à quelques Angström, dans lequel les électrons d'un métal peuvent se déplacer à travers le vide vers le métal voisin [254]. L'idée d'utiliser du courant tunnel pour détecter des molécules uniques a été introduite dans les années 70 [254]. Le passage d'une molécule unique dans l'espace vide séparant les deux électrodes perturbe le courant tunnel mesuré. Les caractéristiques de cette perturbation sont utilisées pour identifier et caractériser la molécule analysée [254].

La méthode a été reprise en 2014 par Zhao et Lindsay *et al.* dans le but de développer un séquençage de protéines à l'échelle de la molécule unique basé sur l'utilisation du courant tunnel [20]. Le principe de la technique consiste à utiliser deux électrodes métalliques de palladium recouvertes d'un agent chimique, le 4(5)-(2-mercaptoethyl)-1H-imidazole-2-carboxamide (ICA), servant de molécule de reconnaissance (**Fig. 20**). Les électrodes sont espacées d'une distance de 2 nm. L'intérêt de fonctionnaliser les électrodes avec l'ICA est lié à sa capacité à interagir avec des bases nucléotidiques de l'ADN, et à sa capacité à interagir avec des acides aminés par le biais de liaisons faibles comme des liaisons hydrogènes. L'ICA interagit de façon temporaire avec la molécule analysée, ce qui permet de l'orienter et d'obtenir une signature du courant bien définie.

Les auteurs ont choisi d'analyser plusieurs acides aminés présentant des différences de structure minimales comme la glycine (Gly) et la glycine méthyle (met-Gly ou Sarcosine), des énantiomères d'asparagine (L-Asn et D-Asn), et les isomères de structure leucine et isoleucine [20]. Les acides aminés sont positionnés entre les deux électrodes, et lorsqu'une faible tension électrique y est appliquée, l'interaction entre l'acide aminé et la molécule de reconnaissance est observée par l'apparition d'un groupement de pics dont les caractéristiques dépendent de la nature/composition chimique de la molécule piégée dans l'espace entre les deux électrodes [20]. Grâce à l'utilisation d'un algorithme (machine learning), il est possible de distinguer au sein du

signal, des pics spécifiques de chacun des acides aminés analysés (Gly, mGly, L-Asn, D-Asn, Leu, Ile, et Arg) avec une précision d'au moins 80% (**Fig. 20**) [20].

De façon à vérifier s'il est possible d'obtenir une quantification avec cette technique, un mélange des deux acides aminés L-Asn et D-Asn à différents ratios a été réalisé et a été soumis à l'algorithme, pour obtenir une reconnaissance [20]. Cependant, la quantification est difficilement réalisable en raison de l'affinité plus grande de la L-Asn avec l'agent chimique ICA. Avec cette technique, la concentration nécessaire en analytes est de l'ordre de 1 à 100 μ M.

Une approche similaire proposée par la suite a permis de montrer que la diminution de la distance entre les deux électrodes d'or, de 2 nm à 0,55 ou à 0,70 nm, permet d'identifier un acide aminé individuel en absence d'un recouvrement par l'ICA sur les électrodes. Selon la distance inter-électrodes choisie, 12 des 20 acides aminés ont été détectés, alors que les autres acides aminés n'ont produit aucun signal détectable. L'amplitude et la durée du pic qui sont mesurés pour chacun des acides aminés analysés individuellement, ont été utilisées pour identifier de façon spécifique les 12 acides aminés.

De plus, cette stratégie a également permis de faire la distinction entre la forme phosphorylée et non phosphorylée d'une tyrosine par l'utilisation d'une distance inter-électrodes de 0,70 nm [255]. L'analyse individuelle de la forme phosphorylée ou non de la tyrosine a montré que les deux formes présentent des signaux différents. La réalisation d'un mélange de formes phosphorylées et non phosphorylées de la tyrosine produit deux pics distincts correspondant aux signaux mesurés lors de l'analyse individuelle des deux formes [255].

Bien que cette approche ait permis de détecter des acides aminés individuels, un type de modification post-traductionnelle, des énantiomères, et des isomères de structure, son utilisation pour le séquençage de protéines est limitée pour plusieurs raisons [6]. L'inconvénient majeur de cette approche est la nécessité de changer de distance entre les deux électrodes afin de distinguer l'ensemble des 12 acides aminés. Une autre difficulté provient du fait que chaque molécule analysée entre les deux électrodes peut s'orienter de plusieurs manières possibles durant la mesure, ce qui mène à l'obtention de plusieurs signaux différents et complexifie l'identification de la molécule. De plus, les algorithmes de reconnaissance doivent être adaptés en conséquence. Enfin, la concentration nécessaire en analytes est de l'ordre de 1 à 100 μ M.

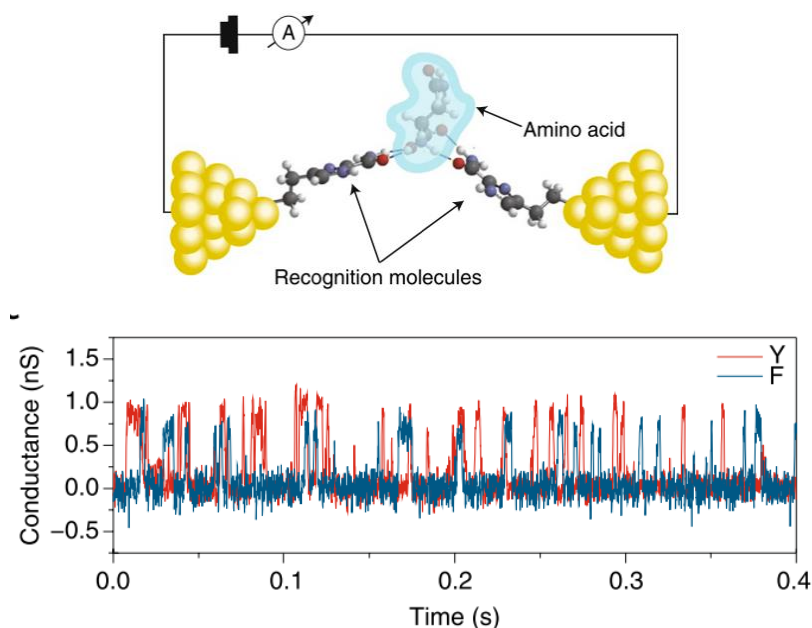


Figure 20 - Stratégie d'identification des acides aminés par effet courant tunnel. Schéma représentatif de la mesure par effet tunnel. La présence d'un acide aminé est observée entre deux électrodes métalliques fonctionnalisées par des molécules de reconnaissance (haut). Portion d'une trace de la conductance en fonction du temps montrant des impulsions électriques d'amplitudes différentes attribuées à la détection de l'acide aminé phénylalanine (F, bleu) ou tyrosine (Y, rouge) (bas). (Adaptée de [20]).

2.2.2 Séquençage des protéines par fluorescence

Le développement d'un séquençage des protéines basé sur l'utilisation de la fluorescence s'est heurté à de nombreuses difficultés dont la principale est celle de pouvoir utiliser des fluorophores différents pour chacun des 20 acides aminés, de façon à les détecter de façon spécifique. En regard de cela, deux approches également basées sur l'utilisation de la fluorescence à l'échelle de la molécule unique ont été développées.

a) FRET fingerprinting

La méthode du FRET fingerprinting propose, plutôt que de tenter d'associer un fluorophore à chacun des 20 acides aminés, d'en utiliser seulement deux pour marquer deux acides aminés (cystéine et lysine) choisis parmi les vingt. L'ordre dans lequel les deux acides aminés marqués s'enchainent, permet d'identifier, par comparaison à une base de données, la protéine analysée.

L'approche requiert de pouvoir lire de façon séquentielle une protéine pour déterminer l'ordre des acides aminés marqués. Pour parvenir à cela, Joo *et al.* ont proposé d'utiliser un moteur moléculaire protéique, ClpX, de la famille AAA+ (ATPases Associated with a variety of

cellular Activities) de *Escherichia Coli* pour « scanner » la protéine [16,18]. Le moteur ClpX possède la capacité de déplier et de dégrader des protéines substrats repliées possédant des tags de reconnaissance, en exerçant sur elle une grande force mécanique de l'ordre de 20 à 30 pN en consommant l'énergie issue de l'hydrolyse de l'ATP. ClpX se présente sous la forme d'un anneau homo-hexamérique. Un partenaire intéressant de ClpX est ClpP qui est une protéase homo-tétradecamérique qui s'autoassemble de façon très stable sous la forme d'un tonneau renfermant une cavité centrale et protégeant ainsi ses 14 sites de clivage de son environnement. Ensemble, ClpX et ClpP forment ClpXP qui peut interagir, déplier, faire transloquer, et dégrader des protéines, faisant de cet assemblage de moteurs moléculaires un candidat idéal pour déterminer la séquence d'une protéine substrat entière [16,18].

Dans les travaux de Joo *et al.*, les auteurs ont choisi de marquer la protéine d'intérêt par des fluorophores accepteurs sur les résidus cystéine (C) et lysine (K) (**Fig. 21**) [16,18]. La protéine est ensuite introduite dans la machinerie ClpXP équipée de fluorophores donneurs et immobilisée sur une surface de quartz recouverte de molécules de PEG par le biais d'une interaction biotine-streptavidine (**Fig. 21**). Après que la protéine cible soit parvenue à se lier à la machinerie ClpXP, sa translocation qui est ATP dépendante est initiée, ce qui oriente le fluorophore accepteur dans le rayon de Förster (2-8 nm), et permet le transfert de fluorescence observé par une chute de l'intensité du signal du donneur et une augmentation concomitante du signal de l'accepteur (**Fig. 21**).

Cette approche a permis aux auteurs de montrer l'identification de plusieurs peptides d'une longueur de 29, 40, et 51 acides aminés, d'une protéine (titine) de 119 acides aminés, et d'un dimère de 210 acides aminés.

Bien que la méthode employée permet de scanner la protéine à une vitesse de translocation stable et de façon unidirectionnelle, celle-ci reste limitée par l'utilisation de bases de données pour identifier la protéine, et ne permet pas de réellement séquencer une protéine. De plus, la méthode présente un taux d'erreur encore trop élevé de 20 à 30% limitant l'identification à 70-80% des protéines.

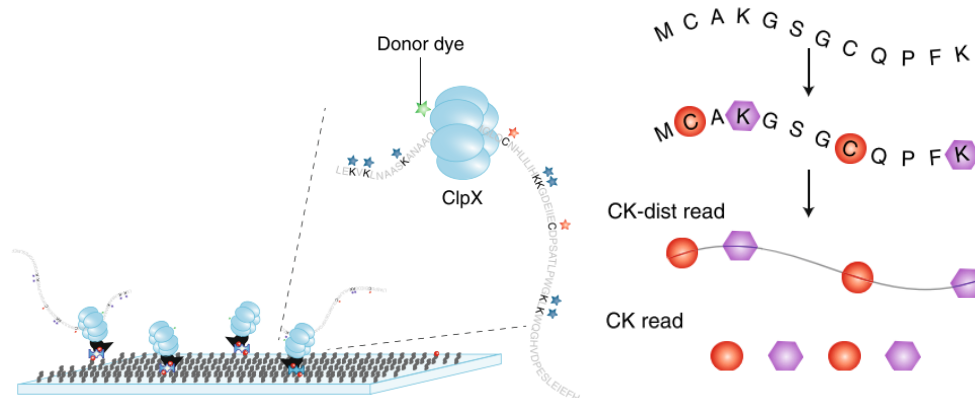


Figure 21 - Stratégie d'identification d'une protéine par la méthode FRET fingerprinting. Schéma représentatif de moteurs moléculaires ClpPX marqués par un fluorophore donneur et qui sont immobilisés sur une surface de quartz fonctionnalisée. Une protéine au préalable marquée par des fluorophores accepteurs au niveau de ses acides aminés cystéine (C) et lysine (K), se lie puis transloque à travers le moteur moléculaire ClpXP. Le rapprochement des fluorophores accepteurs au fluorophore donneur, lors du passage d'acides aminés marqués dans le moteur, permet un transfert de fluorescence permettant d'identifier l'acide aminé. La reconstitution de la séquence en C et K permet de remonter à l'identification de la protéine. (Adaptée de [16]).

b) Fluoroséquençage

Le principe de l'approche de fluoroséquençage est assimilé à une dégradation d'Edman à l'échelle de la molécule unique [17].

La technique du fluoroséquençage propose de fragmenter des protéines cibles en courts peptides (10 à 30 acides aminés) au niveau de sites de coupure contrôlés et prédits à partir du choix de la protéase utilisée (**Fig. 22**). Suite à cela, certains acides aminés sont marqués de façon spécifique par différents fluorophores. Les peptides marqués sont ensuite immobilisés sur une surface et sont analysés par microscopie à fluorescence pour suivre les cycles de dégradation (**Fig. 22**). L'utilisation de la microscopie totale de fluorescence de réflexion interne (TIRF) permet de suivre et de mesurer le signal de fluorescence du fluorophore. Une chute de la fluorescence après un cycle de dégradation d'Edman, indique qu'un acide aminé spécifique a été relargué à l'extrémité N-terminal du peptide analysé. L'acide aminé produit est identifié par analyse spectrale (**Fig. 22**).

Après avoir exploré plusieurs stratégies d'immobilisation, de marquage, et de clivage des peptides, il a été démontré qu'un minimum de quatre acides aminés marqués est nécessaire pour identifier 98% du protéome humain.

Bien que l'approche soit très ingénieuse, de nombreuses difficultés sont rencontrées pour sa mise en œuvre. D'une part, la dégradation d'Edman crée un environnement chimique rude qui n'est pas compatible avec la plupart des fluorophores commercialisés. La méthode peut souffrir de l'inefficacité de la combinaison des fluorophores, du photobleaching, ou encore de l'inefficacité de la dégradation d'Edman. D'autre part, chaque cycle de dégradation d'Edman dure environ 45 minutes ce qui rend le processus de séquençage très long. Enfin, en raison du manque de spectres de fluorescence distinguables et spécifiques, le nombre d'acides aminés distinguables reste limité.

En réponse à ces difficultés, une enzyme a été créée par ingénierie, appelée Edmanase, capable de cliver les acides aminés un par un à partir de l'extrémité N-terminal d'une protéine, et de permettre de réaliser une dégradation d'Edman dans des conditions physiologiques et de façon plus rapide.

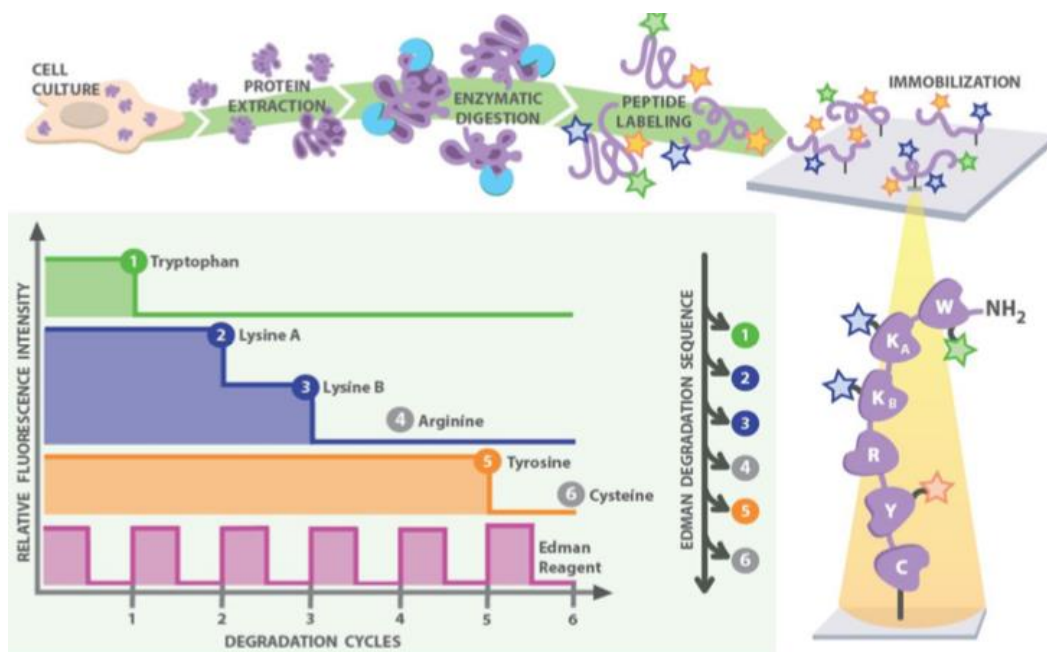


Figure 22 - Stratégie de séquençage des protéines par la méthode de fluoroséquencage basée sur la dégradation d'Edman. Schéma représentatif des étapes d'extraction et de digestion de protéines, de marquage par fluorescence des peptides résultants au niveau de certains de leurs acides aminés, et de l'immobilisation des peptides sur une surface pour qu'ils soient analysés par microscopie à fluorescence au cours des cycles de dégradation d'Edman. Après chaque cycle d'Edman réalisé, une chute de la fluorescence indique le relargage d'un acide aminé spécifique qui est identifié par analyse spectrale. (Adaptée de [17]).

2.3 Résumé

Nous avons présenté dans cette deuxième partie les travaux ayant été réalisés pour développer une stratégie de séquençage par nanopore.

La stratégie qui a naturellement été adoptée est celle qui a mené au succès du séquençage de l'ADN par nanopore, et qui consiste à lire la séquence d'une protéine dépliée au cours de sa translocation à travers un nanopore.

Cependant en comparaison à l'ADN, les travaux d'analyse des protéines se sont heurtés à des difficultés supplémentaires liées à la nature physico-chimique des protéines qui présentent une hétérogénéité de charge, au contraire de l'ADN qui est uniformément chargé. Cela implique qu'en plus de devoir contrôler la vitesse de translocation de la protéine dans le pore pour tenter d'atteindre une précision de lecture d'un acide aminé, il faut également assurer le transport unidirectionnel de la protéine. Par ailleurs, les acides aminés à distinguer sont au nombre de vingt, au contraire de l'ADN qui peut être composé de quatre nucléotidiques, ce qui est plus difficile à atteindre. Enfin, les protéines présentes des modifications post-traductionnelles sur certains de leurs acides aminés, qui doivent également être identifiables.

Malgré les approches ayant proposées d'utiliser des amorces assurant le transport unidirectionnel de protéines dans un nanopore, et d'utiliser des moteurs moléculaires associés aux nanopores pour ralentir la dynamique du transport de protéines ; les nanopores ne permettent toujours pas à ce jour de séquencer une protéine avec cette approche.

D'autres analyses réalisées à l'échelle de petites protéines ou de peptides de différentes natures, dans un régime différent de celui adopté pour le séquençage d'une longue protéine dépliée à travers un nanopore, ont néanmoins montré que les nanopores peuvent permettre de détecter de petites variations de l'ordre de l'acide aminé avec un nanopore modifié chimiquement. Cela indique que l'identification des acides aminés pourrait potentiellement être obtenue avec des nanopores, mais dans un régime d'analyse différent de celui utilisée dans l'approche précédente de séquençage des protéines. Néanmoins, aucune analyse n'avait montré au commencement de mes travaux de thèse, la possibilité d'atteindre une résolution de détection et de séparation d'un acide aminé avec un nanopore biologique sauvage. Le cas de figure idéale serait d'obtenir une telle résolution avec un nanopore sauvage, pour ne pas introduire d'étapes complexes de modifications génétiques.

De plus, d'autres approches de séquençage de protéines à l'échelle de la molécule unique basées sur l'utilisation de la fluorescence ou de l'effet courant tunnel ont été décrites, mais sont techniquement plus complexes que la technologie des nanopores.

C'est dans ce contexte que dans le cadre de mes travaux de thèse, nous nous sommes intéressés au nanopore d'aérolysine présenté juste à la suite, qui a été montré comme permettant, dans sa forme sauvage, de détecter des différences de l'ordre d'une unité monomérique.

Les dernières avancées apparues en même temps que mes travaux de thèse seront présentées dans la discussion du chapitre n°4.

III. Le nanopore biologique d'aérolysine

L'aérolysine est une toxine qui forme des pores (PFT) produite par les bactéries Gram-négatif de l'espèce *Aeromonas*, et qui est en majeure partie responsable de la haute pathogénicité de ces bactéries [72,256]. En ciblant les cellules de plusieurs espèces, dont les cellules humaines, elle a pour conséquence de causer le développement de pathologies comme la gastroentérite ou l'infection des plaies profondes.

La présence d'un facteur hémolytique présent dans le milieu de culture de la majorité des souches de *Aeromonas hydrophila* a été observée pour la première fois dans les années 60 par F. H. Caselitz et R. Gunter [257] qui sera renommé plus tard "aérolysine" par A. W. Bernheimer et L. S. Avigad [258]. Le succès de la purification de l'aérolysine à homogénéité par J. T. Buckley et ses collègues [259] a donné suite à de nombreuses études visant d'une part, à caractériser la protéine et à comprendre son mécanisme de production (synthèse, export, sécrétion, insertion), et d'autre part, à déterminer la séquence du gène de la protéine dans le but premier de le cloner [260,261] et de permettre la production en grande quantité de la protéine dans des systèmes hôtes (*e.g Escherichia coli* [262], *Aeromonas salmonicida* [263]) mais également de pouvoir prédire la séquence en acides aminés de la protéine.

Chez la bactérie *Aeromonas hydrophila*, la toxine est synthétisée dans le cytoplasme de la bactérie sous une forme monomérique dite « pré-pro-aérolysine » [23,70,264,265]. La pré-pro-aérolysine est constituée d'une séquence signal présente à son extrémité N-terminal, qui nécessite d'être clivée pour son passage du cytoplasme vers le périplasma *via* la membrane interne de la bactérie [23,70,264,265]. Elle devient alors « pro-aérolysine ». La pro-aérolysine est ensuite sécrétée par la bactérie sous une forme soluble stable. Cette stratégie est mise en place par la bactérie car celle-ci ne possède pas de mécanisme de défense vis-à-vis de la toxine. Ainsi, pour se protéger elle-même, la bactérie produit sa toxine sous une forme inactive qui, une fois dans le milieu extracellulaire, deviendra active sur des cellules cibles possédant des récepteurs membranaires spécifiques (protéines, lipides, carbohydrates) par le clivage d'un propeptide présent à son extrémité C-terminal. La toxine va se fixer à ses récepteurs et s'oligomériser sous la forme d'un heptamère pour former un pore dans la membrane cellulaire cible [23,70,264,265]. Cela rend la cellule perméable et peut conduire à la mort cellulaire [23,70,264,265].

L'aérolysine a été utilisée comme un nanopore biologique depuis longtemps pour la détection électrique de divers types de molécules [25,266] comme les peptides, les polymères synthétiques, ou les acides nucléiques.

Cette troisième partie a pour objectif de présenter : (i) la structure du monomère de pro-aérolysine ; (ii) le mécanisme de formation du pore d'aérolysine ; (iii) la structure du pore et ses propriétés électriques ; et enfin (iv) la détection électrique de molécules par nanopore.

3.1 Structure de la pro-aérolysine

Comme décrit précédemment, la protéine est produite et relarguée par *Aeromonas hydrophila* sous la forme d'un précurseur inactif d'environ 52 kDa et d'une longueur de 470 acides aminés nommé « pro-aérolysine » dont la séquence a été établie (**Fig. 23**) [267].

Des prédictions réalisées à partir de la séquence de la protéine, confortées par des expériences de cristallographie aux rayons X (à une résolution de 2,8 Angström) et de dichroïsme circulaire, ont montré que la protéine est riche en structures bêta (42%) et présente une plus faible part d'hélices alpha (21%) [267].

La proaérolysine peut exister sous forme de dimères en solution [268]. La cristallographie du dimère de pro-aérolysine a permis de caractériser la structure du monomère qui s'organise en forme de « L » constituée de deux lobes distincts [269]. Le premier petit lobe, plus isolé du reste de la protéine, correspond au premier domaine (domaine 1) de la protéine. Le second lobe plus large se divise en trois autres domaines (domaines 2, 3, et 4) (**Fig. 23**). Les deux lobes sont reliés par un peptide flexible de 10 acides aminés. Également, l'existence de deux ponts disulfures a été identifiée dans le domaine 1 (Cys19-Cys175) et dans le domaine 2 (Cys159-Cys164) [269]. Ces ponts disulfures seraient impliqués dans la stabilisation de la protéine et dans la protection d'une boucle comprise entre les acides aminés 154 et 167 d'une possible dégradation enzymatique par des protéases [269].

Le **domaine 1** constitue les premiers 83 acides aminés à partir de l'extrémité N-terminal. Il a été mis en évidence que l'ensemble du domaine 1 (résidus 1-82) est impliqué dans la liaison au récepteur de la toxine. Les acides aminés du domaine engagés dans la liaison à ce récepteur, plus particulièrement avec un carbohydrate, ont été identifiés par plusieurs travaux (**Fig. 24**) [270,271].

Il a été démontré que les récepteurs de la pro-aérolysine sont des protéines ancrées par un glycosyl-phosphatidylinositol (GPI) appelées protéines à GPI (GPI-anchored proteins) (**Fig. 24**). Ces protéines membranaires utilisent le GPI pour réaliser un ancrage hydrophobe à la membrane plasmique de la cellule où elle servent de récepteur. Le domaine 1 est impliqué dans la liaison à ce type de récepteur par une première interaction avec une N-glycosylation présente sur la protéine engagée dans l'ancrage par GPI (**Fig. 24**).

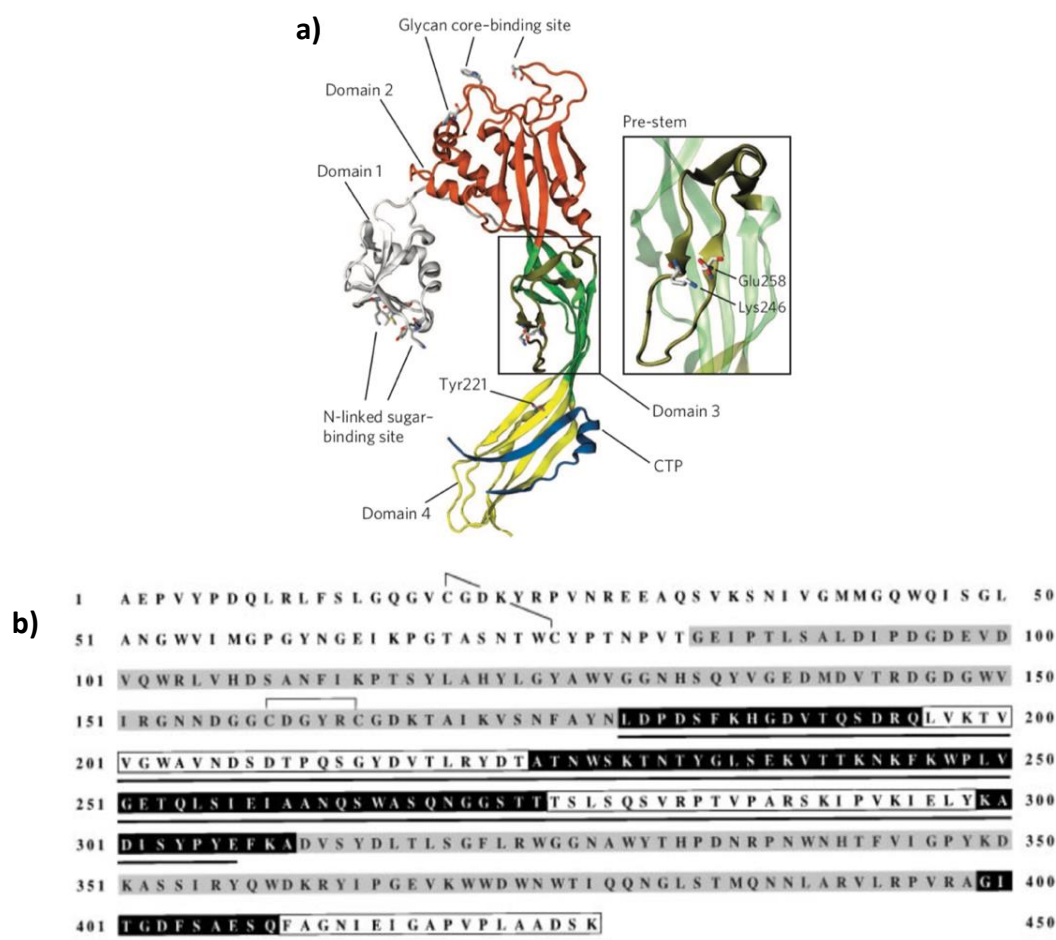


Figure 23 - Représentation de la structure et de la séquence de la pro-aérolysine de *Aeromonas hydrophila*. (a) Représentation en structures secondaires de la pro-aérolysine et de ces quatre domaines avec une visualisation des sites de liaison au récepteur de la membrane (Glycan core-binding site et N-linked sugar-binding site) détaillés dans la Figure 24, de la boucle pré-stem, et du pro-peptide en C-terminal (CTP). (b) Séquence en acides aminés de la pro-aérolysine avec une représentation de la séquence de chaque domaine par une couleur caractéristique (Domaine 1: blanc, Domaine 2: gris; Domaine 3: noir, et Domaine 4: blanc avec un encadré noir), et des ponts disulfures représentés par un trait noir formé entre deux résidus cystéine. (Adaptée de [269,272]).

Il est également important de mentionner la présence dans le domaine 1 d'un peptide signal de 23 acides aminés, présent en N-terminal, et dont le rôle est d'orienter la pro-aérolysine vers la

membrane interne de la bactérie pour son passage vers le périplasma durant le mécanisme de sécrétion de la toxine [265].

Enfin, d'autres études montrent que le domaine 1 joue également un rôle dans le processus de dimérisation du monomère [271].

Le **domaine 2** du monomère possède des régions (résidus 83-178 et 311-398) qui sont également impliquées dans la liaison à son récepteur à travers plusieurs acides aminés (Tyr-162, Trp-324, et His-332). Le domaine 2 se fixe directement sur la partie glycan de la GPI-anchored protéine. Plus particulièrement, une modification de type N-acetylglucosamine se trouve sur le deuxième mannose du glycan qui est cruciale à la liaison de l'aérollysine au récepteur. Le domaine 2 jouerait également un rôle dans l'initiation de l'oligomérisation des monomères. Certains travaux ont permis d'identifier plus particulièrement deux régions présentes dans le domaine 2 qui ont une influence sur le taux d'oligomérisation [262,271,273,274].

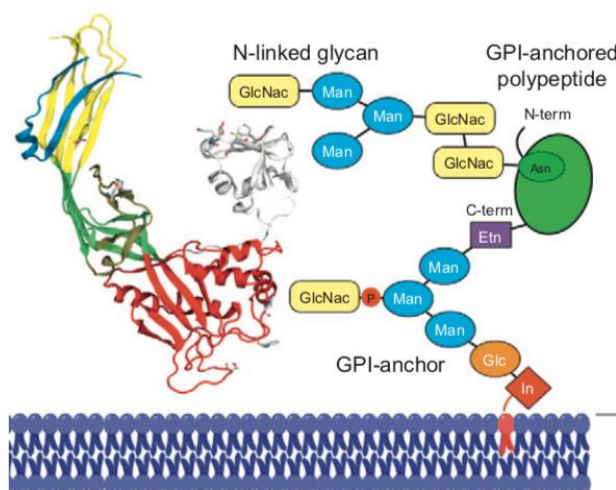


Figure 24 - Représentation schématique d'un monomère de pro-aérollysine lié à son récepteur membranaire à GPI par ses différents domaines. Le monomère d'aérollysine est montré comme dans la Figure 23 avec ses quatre domaines structuraux. Le domaine 1 est lié à une N-glycosylation (N-linked glycan) visible sur le schéma. Le domaine 2 est lié au niveau d'une modification N-acétylglucosamine présente sur le deuxième mannose de la partie glycan de la protéine à GPI ancrée dans le feuillet externe de la membrane plasmique par un phospholipide de type phosphatidyl-inositol. Légende : Asn: asparagine, Etn: éthanol amine, In: inositol, P: phosphate, Glc: glucose, Man: mannose, GlcNac: N-acétylglucosamine. (Adaptée de [275]).

Le **domaine 3** du monomère possède des régions (résidus 179-195, 224-274, 299-310 et 399-409) qui jouent un rôle important dans les interactions entre monomères dans le processus

d'oligomérisation (heptamérisation). Le domaine 3 est constitué de cinq feuillets β contre lesquels se trouve une boucle qui est enroulée. Cette boucle adopte une structure en épingle qui lui permet de diriger l'insertion du monomère et l'ancrage du brin β dans la membrane [269,272]. Certaines régions du domaine 3 semblent également impliquées dans le processus de sécrétion de la protéine. Il a été mis en évidence que la mutation du résidu tryptophane à la position 227 (W-227) dans le domaine 3 peut empêcher la sécrétion de la protéine par la bactérie, et que lorsque la mutation correspond à une substitution du résidu W-227 par une leucine (W227L), la protéine est alors retrouvée bloquée dans la membrane externe de la bactérie [276]. Certaines mutations de ces régions du domaine 3 ont également un effet sur la stabilité de la protéine, qui est en conséquence dégradée dans le périplasme, mais il n'est pas connu si cela peut expliquer l'absence de sécrétion observée.

Enfin, le **domaine 4** du monomère (résidus 196-233, 275-298, et 410-470) contient le peptide C-terminal (CTP) (**Fig. 23**) qui est le pro-peptide présent dans la forme sécrétée soluble de la pro-aérolysine, qui doit subir un clivage protéolytique pour activer les monomères et permettre leur oligomérisation en heptamère. Il n'est donc pas retrouvé dans la séquence finale du pore. Il a également été mis en évidence que le peptide CTP du domaine 4 a une fonction chaperon durant la biosynthèse de la protéine pour permettre le repliement du domaine [277].

3.2 Mécanisme d'insertion du pore

Le mécanisme d'insertion (**Fig. 25**) décrit ici inclut également une brève description du mécanisme de liaison de la protéine à la membrane, de l'activation de la protéine par clivage du pro-peptide, de l'oligomérisation des monomères, et de l'insertion à proprement dit de l'heptamère dans la membrane.

Le mécanisme d'élucidation de la structure et de la formation du pore d'aérolysine a été exploré au cours de plusieurs études. Alors que la structure de la pro-aérolysine est disponible depuis 1994, l'obtention de la structure tridimensionnelle du pore d'aérolysine a constitué un défi. Les premières études structurales à basse résolution (25 Å) de l'aérolysine insérée dans une membrane de lipides ont été obtenues par Parker *et al.* en 1994 [267]. Une des raisons possibles de l'échec des tentatives de cristallisation du pore d'aérolysine est liée à la flexibilité élevée des protéines. Le mécanisme d'insertion original de ce pore qui implique que l'organisation structurale de la protéine change radicalement entre les dimères solubles de pro-aérolysine et la conformation heptamérique du pore, rend difficile sa caractérisation. Les pores d'aérolysine

à l'état sauvage ne peuvent pas être analysés à l'échelle de la molécule unique également car ils s'agrégeraient en amas hétérogènes [272,278].

Face à cette difficulté, la stratégie adoptée a consisté à développer des versions mutantes de l'aérolysine, qui ont permis de mieux comprendre la structure heptamérique du canal et le mécanisme de formation du pore. Ces mutants correspondent à des modèles structuraux de l'aérolysine acquis à différents stades intermédiaires du processus de formation du pore. L'utilisation de plusieurs approches dont la cristallographie aux rayons X, la microscopie électronique cryogénique (16,6–18,3 Å), la simulation moléculaire, et la modélisation informatique, était nécessaire pour obtenir la structure du pore inséré dans une membrane. La combinaison de ces différentes approches expérimentales a permis de fournir des modèles pseudo-atomiques de ces structures intermédiaires et de proposer un nouveau mécanisme moléculaire à l'origine de la formation d'un pore d'aérolysine. Plusieurs structures de transition ont été identifiées dont l'existence est liée à une réorganisation de la boucle pré-stem (pré-tige) entre les domaines 3 et 4 (**Fig. 23**), survenant après le clivage protéolytique de la séquence pro-peptide à proximité de la membrane lipidique, qui contribue finalement à créer un tonneau bêta transmembranaire. Les structures de ces mutants ont été appelées pré-pore, post pré-pore, et quasi-pore [77,272].

Par la suite, des cartes de microscopie cryo-électronique atomique (cryo-EM) à une résolution quasi-atomique (3,9–4,5 Å) des états pré-pores, post-pré-pores et quasi-pores de l'aérolysine ont été fournies. Également, une structure du pore final de l'aérolysine à l'état sauvage a été obtenue en présence d'un détergent à une résolution de 7,9 Å. Par comparaison à celles d'autres toxines comme la lysénine, l'obtention de ces structures permet d'aboutir à une description des étapes majeures impliquées dans le processus de formation des pores par l'aérolysine [77].

Tout d'abord, la séquence pro-peptidique CTP qui maintient la protéine sous sa forme inactive est clivée ce qui permet l'oligomérisation des monomères solubles (**soluble**) (**Fig. 25**) [77]. Le clivage du CTP et l'oligomérisation engendrent une réorganisation de plusieurs domaines de la protéine dont une rotation du domaine 1 qui est orientée et devient capable de se fixer au récepteur de la protéine avec le domaine 2. Il y a également une réorganisation du domaine 4 qui grâce à l'oligomérisation des monomères permet un assemblage en une structure pré-pore heptamérique correspondant à un tonneau bêta, *via* des interactions hydrophobes (**pre-pore**) (**Fig. 25**) [77]. Cet assemblage des monomères est suivi d'un mécanisme d'insertion particulier dit « tourbillonnant » (**swirling mechanism**) permettant le passage du pré-pore au pore (**Fig. 25**)

[77]. Au cours de cette transition, les domaines 3 et 4 subissent une rotation et se positionnent de façon quasi parallèle à la membrane. En même temps, la réorganisation du domaine 4 permet d'extraire la boucle pré-stem (**post pre-pore**) (**Fig. 25**) [77]. La boucle pré-stem se replie et adopte une forme allongée qui a la capacité de s'insérer dans la membrane, ce qui lui permet d'étirer le tonneau bêta sur une longueur d'environ 8,7 nm en prévision de l'enfoncement du tonneau bêta à son tour dans la membrane (**quasi-pore**) (**Fig. 25**) [77]. Cette étape est suivie de torsions au niveau de la partie inférieure des tonneaux bêta, plus précisément à l'interface entre les domaines 2 et 3, et à l'interface entre les domaines 3 et 4, aboutissant à l'enfoncement complet du tonneau bêta dans la membrane sans désorganisation de sa structure (**pore**) (**Fig. 25**) [77].

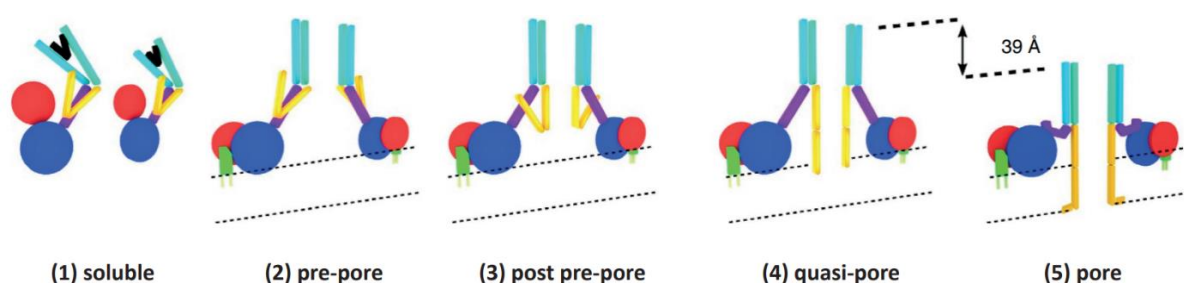


Figure 25 - Représentation schématique des différentes étapes du mécanisme d'insertion du pore d'aérollysine dans une membrane lipidique. Le clivage de la boucle CTP engendre l'oligomérisation des monomères solubles pour former un pré-pore (1-2). Le passage du « pre-pore » au « post pre-pore » implique une extraction de la boucle pré-stem survenant suite à une réorganisation du domaine 4 (2-3). Le repliement et l'allongement de la boucle pré-stem permet de tirer sur le tonneau bêta et de passer à l'étape de « quasi-pore » (3-4). Le tonneau bêta subit des torsions permettant l'enfoncement complet du pore dans la membrane (4-5). Légendes : boucle CTP (noir) ; domaine 1 (rouge) ; domaine 2 (bleu foncé) ; domaine 3 (violet) ; tonneau b (bleu turquoise) ; récepteur (vert) ; boucle pré-stem (jaune) ; (Adaptée de [77]).

3.3 Structure et propriétés électriques du canal

Le pore d'aérollysine présente une structure en forme de « champignon » constituée principalement de deux domaines : un premier domaine *cis* (chapeau) restant extracellulaire lors de son insertion dans la membrane, et un second domaine *trans* (tige) inséré dans la membrane lipidique (**Fig. 26**) [77]. Le domaine *cis* présente des structures secondaires en rubans d'hélices alpha, et le domaine *trans*, des structures secondaires en un tonneau de feuillets bêta. Le pore présente un diamètre qui varie entre 1 et 1,7 nm, sur une longueur totale d'environ 9 nm, et une largeur totale de 14 nm (**Fig. 26**) [77]. Le pore d'aérollysine structuré sous la forme

d'un tonneau bêta présente deux constriction, la première observée à l'entrée et la deuxième observée à la sortie du pore [77].

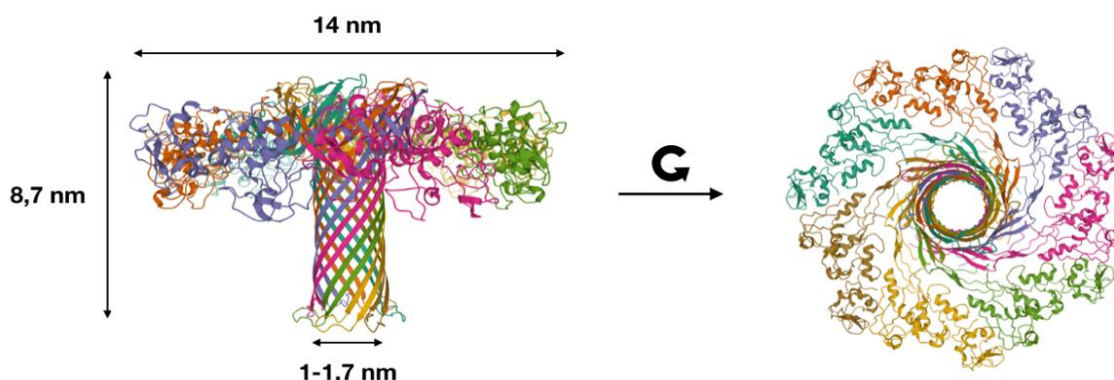


Figure 26 - Représentation du nanopore d'aérolysine produit par *Aeromonas hydrophila*. Représentation du pore en structures secondaires vue de face (gauche) et vue de haut (droite). Chaque monomère du pore est représenté par une couleur. Les dimensions du pore indiquent une largeur totale de 14 nm, une longueur totale de 8,7 nm, et un diamètre interne compris entre 1 et 1,7 nm. (PDB : 5JZT ; [77])

L'une des caractéristiques les plus importantes du pore d'aérolysine pour son utilisation pour la détection électrique de molécules uniques est sa distribution de charge. Le tonneau bêta est composé d'une alternance d'acides aminés positifs et négatifs (lysine et acide glutamique) formant des anneaux de charges [272]. Le motif d'alternance de résidus hydrophobes-hydrophiles (**Fig. 27**) est présent sur une distance d'environ 2,5 nm correspondant à la distance moyenne de la région hydrophobe d'une bicouche lipidique eucaryote. Il se termine au niveau d'un anneau de résidus aromatiques orienté vers l'extérieur formé par Trp-265 et Tyr-233. Ces deux résidus forment une ceinture aromatique idéalement positionnée et fournissant un ancrage idéal de la structure à la limite externe de la bicouche lipidique.

Une autre caractéristique intéressante du pore d'aérolysine est son grand nombre de résidus chargés pointant vers la lumière du pore. On compte un total de 91 charges (de la position 198 à 290) organisées en anneaux (13 par monomères), dont 7 charges positives ne sont pas appariées (charge nette +7e) [272,279,280]. Ces observations sont opposées à celles réalisées chez les PFT issues de la famille de *Staphylococcus aureus* qui manquent de résidus chargés dans la lumière des pores. Ces caractéristiques sont supposées être responsables de l'amélioration considérable de l'interaction électrostatique entre certains analytes comme par exemple les acides nucléiques et le pore, conduisant à un ralentissement de la vitesse de transport de l'ADN dans l'aérolysine, qui est mesurée comme étant trois fois plus faible en

comparaison à la vitesse de transport de l'ADN mesurée dans l' α -hémolysine. La réduction de la vitesse de translocation d'une chaîne d'ADN conduit à une durée de visite prolongée de la chaîne dans le pore ce qui peut permettre d'obtenir des informations plus détaillées et plus précises sur des molécules individuelles (nucléotides).

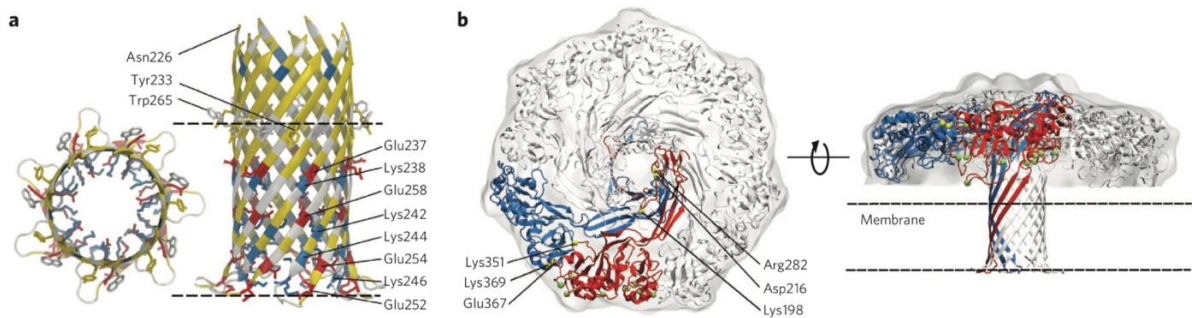


Figure 27 - Modèle quasi-atomique de la structure du pore d'aérollysine sauvage montrant les résidus impliqués dans l'ancrage du pore dans la membrane et les résidus chargés pointant vers la lumière du pore. (a) Représentation en structures secondaires du tonneau bêta du pore d'aérollysine avec la partie représentée entre les lignes en pointillés correspondant à la zone transmembranaire où l'on visualise l'alternance de résidus d'acide glutamique (Glu) et de lysine (Lys), et la zone supérieure où l'on visualise les résidus aromatiques tryptophane (Tyr-265) et tyrosine (Tyr-233) permettant l'ancrage du pore à la limite externe de la membrane de lipides. (b) Vue de haut (gauche) et de face (droite) du nanopore d'aérollysine et l'arrangement de ses sept monomères, avec la visualisation des résidus Glu, Lys, et Arg (Arginine) positionnés à l'interface entre deux monomères et impliqués dans l'oligomérisation (Adaptée de [272]).

L'insertion du nanopore dans des bicouches lipidiques planes reconstituées a permis de réaliser des mesures concernant les propriétés électriques du pore. L'aérollysine présente une conductance unitaire de 420 pS, une légère sélectivité aux anions, une légère rectification du courant, et une sensibilité au voltage [281–283]. Le pore est généralement stable et reste ouvert entre 70 et -70 mV, et peut présenter des comportements d'ouverture et de fermeture (gating) en dehors de cette gamme de tensions. Également, des pores de conductance plus faible ou plus élevée sont parfois observés, ce qui peut être attribué soit à la formation de pores constitués d'un nombre de monomères inférieur ou supérieur à sept, soit à l'insertion simultanée de plusieurs pores dans la membrane.

3.4 Détection de molécules par le nanopore d'aérollysine

Depuis les premières études ayant utilisées l'aérollysine comme nanopore, notamment pour la détection de peptides structurés en hélices alpha [242], et pour la translocation de protéines

dépliées [160]; les travaux proposant l'utilisation du pore se sont multipliés et ont été appliqués à la détection de diverses molécules [25,266].

Le nanopore d'aérolysine a d'abord été utilisé pour l'étude du transport des peptides. Dans ces études, il a été montré que l'aérolysine est sensible à la longueur de l'hélice alpha et à la charge nette du peptide [242], et à la nature de la structure secondaire du peptide [246].

D'autres expériences ont montré qu'il est possible de suivre le dépliement des protéines sauvages et mutantes [217,225], le transport de protéines dépliées [160,284], et le transport de chimères ADN-protéine [220].

L'aérolysine a également été utilisée pour étudier la dynamique de polymères synthétiques non ioniques (PEG) et de molécules cavitaires (cyclodextrines). Les auteurs de cette étude se sont intéressés au transport de ces molécules, par électrosmose ou par électrophorèse, à travers le nanopore d'aérolysine [283], et à la contribution de ces forces motrices au transport (direction et force) de ces molécules dans différentes conditions expérimentales en électrolyte (KCl *versus* LiCl). Ces travaux, auxquels j'ai pu contribuer, ont particulièrement montré que le phénomène d'électrosmose est plus dominant dans l'aérolysine que dans l' α -hémolysine. En l'occurrence, l'utilisation de LiCl dans les expériences réalisées avec l'aérolysine, permet d'étendre la détection à des molécules de PEG de longueurs qui ne parviennent pas à être détectées avec l' α -hémolysine.

Plus récemment, il a été montré qu'il existe une influence de la tension électrique, de la concentration de l'analyte, et de la longueur de l'analyte, sur la dynamique du transport de polymères de dextrane sulfate avec le nanopore d'aérolysine [285].

D'autres études ont porté sur le mécanisme d'entrée et de transport des nucléotides en fonction de la température ou de la force électrique [27,286], la modification de nucléotides [287], ainsi que la discrimination en taille de polymères de PEG [26] et de nucléotides [288].

Une revue concernant l'ensemble des analyses ayant été réalisées avec le nanopore d'aérolysine, et à laquelle j'ai participé durant ma thèse, a été publiée en 2019 dans la revue *ACS Sensors* et est présentée en **Annexe 1**.

Dans le cadre de la démarche scientifique des travaux de cette thèse, on se limitera à la description des travaux récents ayant utilisés le nanopore d'aérolysine pour mettre en avant son pouvoir de résolution, notamment pour la séparation en taille de divers analytes.

A notre connaissance, aucun nanopore biologique à l'état sauvage n'avait permis d'atteindre une résolution de détection d'un acide aminé sans modification chimique du canal. L'aérollysine se présente comme un candidat de choix pour atteindre cette résolution car il existe plusieurs exemples dans la littérature, illustrant la capacité de ce nanopore à atteindre une résolution de discrimination d'une unité monomérique, comme dans le cas de polymères synthétiques [26], ou de courtes chaînes d'oligonucléotides [288].

Les travaux de Kasianowicz & Robertson *et al.* décrits dans la partie n°1 de ce chapitre [175] avaient déjà permis plusieurs années auparavant de démontrer le lien existant entre la longueur d'une chaîne de polymère de PEG et les caractéristiques d'amplitude et de durée du blocage du courant engendré lorsque la molécule est transportée dans le nanopore. Les auteurs avaient plus particulièrement montré la possibilité d'atteindre une résolution de discrimination d'une unité monomérique (44 g/mol) dans le contexte de l'analyse d'un mélange polydisperse de PEG avec le nanopore d' α -hémolysine. Ces travaux étaient à l'origine de l'introduction du concept de la « spectrométrie de masse par nanopore ». Il avait été également montré par la suite que l'augmentation de la concentration en sels (3-4 M) augmente d'une façon spectaculaire les interactions entre les molécules de PEG et l' α -hémolysine [176,289].

Ces travaux ont été repris en 2015 dans le but d'analyser des molécules de PEG avec le nanopore d'aérollysine et de comparer les résultats obtenus avec ceux précédemment observés avec le nanopore d' α -hémolysine [26].

Un paramètre important pour l'obtention d'un spectre de discrimination d'un mélange polydisperse de PEG avec une très bonne résolution est le temps de résidence. Il avait été observé que ce sont les temps de résidence particulièrement longs atteints grâce à l'utilisation de hautes concentrations en sels qui permettaient en partie d'atteindre une telle résolution de séparation des chaînes de polymères.

Dans le cas de l' α -hémolysine, les molécules de PEG sont détectées sous l'action d'une tension électrique positive lorsqu'elles sont ajoutées en solution du côté *cis* du pore. Les temps de résidence les plus longs sont mesurés sous l'action d'une différence de potentiel de +40 mV ce qui a conduit les auteurs à obtenir des spectres de masse permettant de discriminer des molécules de PEG de 24 (767 g/mol) à 48 (2088g/mol) monomères avec une résolution de séparation d'un monomère.

Quand les auteurs ont réalisé les mêmes expériences avec l'aérolysine (**Fig. 28**), il est apparu dans un premier temps que les polymères ne sont pas détectés sous l'action d'une tension électrique positive mais sous l'action d'une tension électrique négative. Au contraire de l' α -hémolysine, le spectre de masse obtenu à -40 mV ne permet d'observer aucune distribution de pics attribuables aux différentes populations ajoutées en solution (**Fig. 28**). En augmentant la tension électrique progressivement, une résolution de détection et de séparation des polymères est obtenue à partir de -80 mV, et la discrimination est meilleure avec l'augmentation de la tension électrique de -80 mV à -120 mV et jusqu'à -160 mV (**Fig. 28**). En comparaison avec l' α -hémolysine, le spectre de masse obtenu avec le nanopore d'aérolysine est étendu et permet de détecter des molécules de PEG de 17 (767 g/mol) à 48 (2088 g/mol) monomères, offrant la possibilité de détecter des chaînes bien plus courtes (**Fig. 28**).

Dans le cas de l'aérolysine, l'augmentation de la tension électrique appliquée permet d'augmenter les temps de résidence des chaînes de PEG dans le pore d'aérolysine, ce qui laisse penser que les interactions entre les polymères et le pore sont favorisées sous l'action de plus hautes tensions électriques. C'est ce qui serait à l'origine de l'obtention de la séparation des polymères avec une résolution d'un monomère, et plus particulièrement pour les plus petites chaînes de polymères non détectées avec le nanopore d' α -hémolysine. Cette hypothèse a été confirmée en comparant la dynamique d'entrée de PEG 1500 g/mol (polydisperse) à travers l'aérolysine en fonction de la nature du sel (KCl *versus* NaCl). Les ions K⁺ se complexeraient aux chaînes de PEG, ce qui indique qu'une augmentation de la force électrique entraînerait les chaînes de polymères dans le pore sans qu'elles puissent être transportées, car le temps de résidence augmente de façon exponentielle avec la force électrique [290]. L'augmentation des durées de bouchage tend à réduire la dispersion des valeurs moyennes du courant de chaque population, permettant ainsi une meilleure discrimination des populations. Toutefois, même si les durées de blocages sont suffisamment longues, la discrimination des populations peut être perdue, si la chaîne n'est pas entièrement logée dans le pore. C'est le cas du PEG de masse molaire 3400 g/mol (PEG 3400) par exemple [177], où une partie de la chaîne est logée dans le pore et l'autre partie reste à l'extérieur du pore. Dans cette situation le bloc est saturé, les amplitudes de courant sont maximales et identiques, ne permettant donc pas la discrimination des différentes espèces du PEG 3400. Cependant, la discrimination des différentes populations du PEG 3400 a pu être observée quand la température avait été augmentée à 45°C. L'augmentation de la température entraîne l'effondrement du PEG par une transition pelote-globule, induisant une diminution de la dimension des différentes espèces du PEG 3400, leur

permettant d'être logées entièrement dans le pore. En conclusion, la discrimination en taille est obtenue si et seulement si les temps de résidence sont suffisamment longs, et à condition que la chaîne soit entièrement logée dans le pore pour que chaque monomère de la chaîne puisse contribuer au blocage du courant [177].

L'aérolysine a également été utilisée pour détecter des courts oligosaccharides composés de 6, 8, et 10 unités monomériques [291]. Dans ce travail, il a été montré que ces oligosaccharides chargés négativement issus de la dégradation enzymatique en temps réel de molécules d'acide hyaluronique, sont discriminables en fonction de leur longueur.

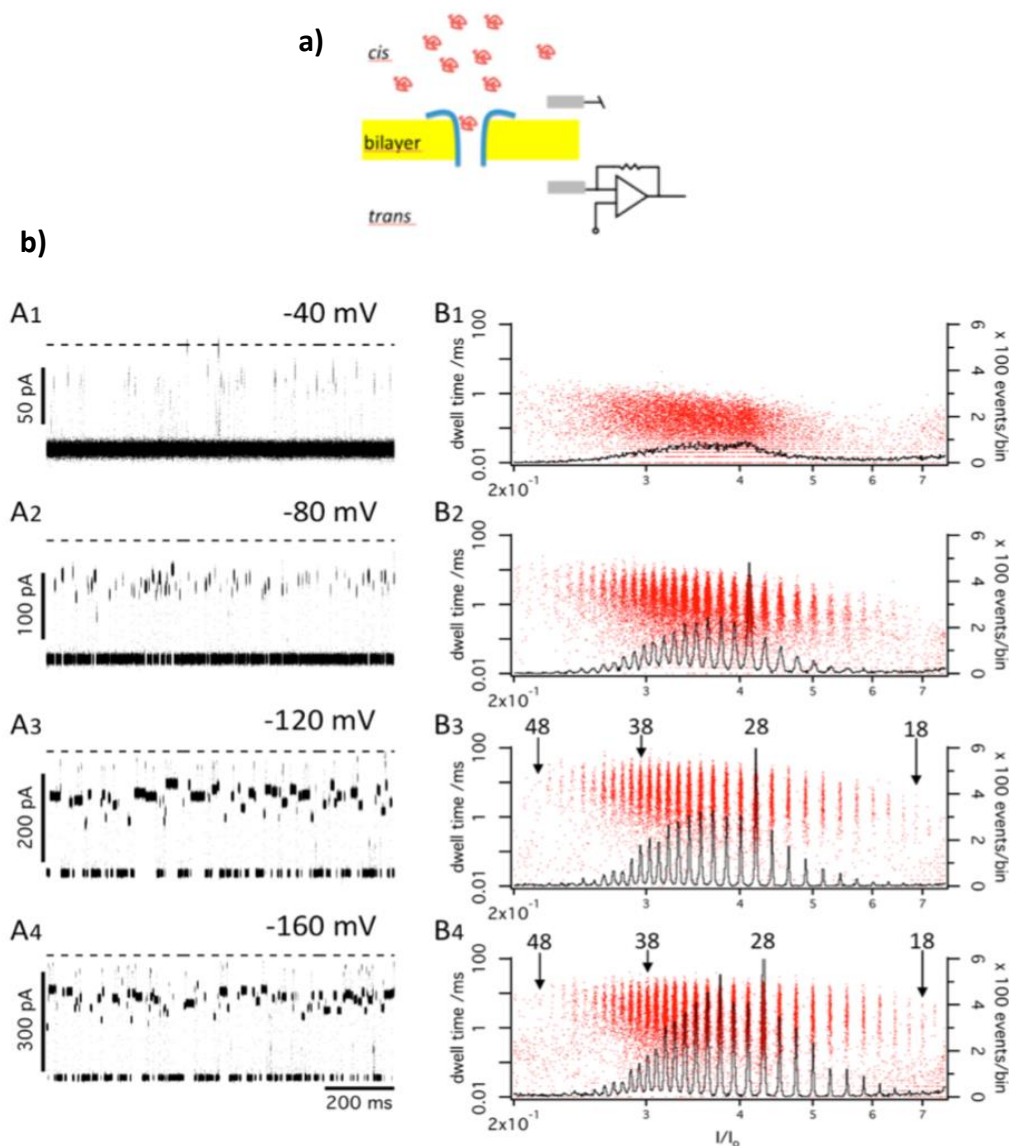


Figure 28 - Discrimination de molécules de PEG avec une résolution de séparation d'un monomère, obtenue avec le nanopore d'aérolysine sauvage. (a) Schéma représentatif d'un nanopore d'aérolysine (bleu) inséré dans une bicouche de lipides (jaune) et mis en présence de

molécules de PEG (rouges) ajoutés dans le compartiment cis (haut). (b) Portions de traces de courant en fonction du temps (gauche), et histogrammes du courant relatif résiduel I_b/I_0 (ligne noir) superposés aux nuages de points (points rouge) de la distribution des durées de blocage en fonction de I_b/I_0 (droite), obtenus dans le cas de l'analyse d'une solution polydisperse de PEG (1500 g/mol) du côté cis du nanopore d'aérollysine, sous l'action d'une tension électrique appliquée en trans de -40 mV (A1), -80mV (A2) ; -120 mV (A3) ; et -160 mV (A4). On visualise que jusqu'à -120 mV, plus l'amplitude de tension électrique est importante, plus les durées de blocages du courant sont longues, et plus la résolution de détection et discrimination est améliorée (bas). Les durées de blocage semblent diminuer à -160 mV, mais la discrimination est conservée. (Adaptée de [26]).

L'ensemble de ces résultats obtenus avec l'aérollysine ont orienté son utilisation pour l'analyse de courts oligonucléotides de tailles et de séquences différentes [27,292,293] mais également pour la détection de modifications chimiques d'oligonucléotides [294,295].

Des auteurs ont démontré en 2016 [27] que de courts oligonucléotides d'adénine d'une longueur de 2, 3, 4, 5, et 10 nucléotides, analysés avec l'aérollysine (**Fig. 29**), individuellement ou dans un mélange, produisent des blocages du courant bien distinguables sur la base de leurs amplitudes de blocage du courant et de leurs durées de blocage du courant. Ainsi, ils parviennent à être discriminés avec une résolution de séparation d'un nucléotide d'adénine (174 g/mol) (**Fig. 29**). Ces résultats ont été obtenus par l'utilisation d'une solution ionique à une concentration de 1 M. Ces conditions expérimentales sont classiquement utilisées dans des expériences avec des nanopores, mais la concentration de la solution ionique est plus basse que dans le cas de la discrimination des molécules de PEG avec l'aérollysine. Les auteurs ont également prouvé que la distinction des populations ne provient pas d'artefacts mais bien de la détection des oligonucléotides qui sont discriminés. Ils ont apporté la preuve de cela en réalisant une dégradation enzymatique en temps réel de l'oligonucléotide d'une longueur de cinq adénines, qui révèle l'apparition progressive de populations supplémentaires correspondant à des chaînes de plus courtes longueurs. La résolution de séparation atteinte dans cette étude a également été attribuée à la géométrie du pore, et à l'état de charge de l'aérollysine, qui aurait de fortes interactions électrostatiques avec les oligonucléotides chargés négativement, détectés sous l'action d'une tension électrique positive.

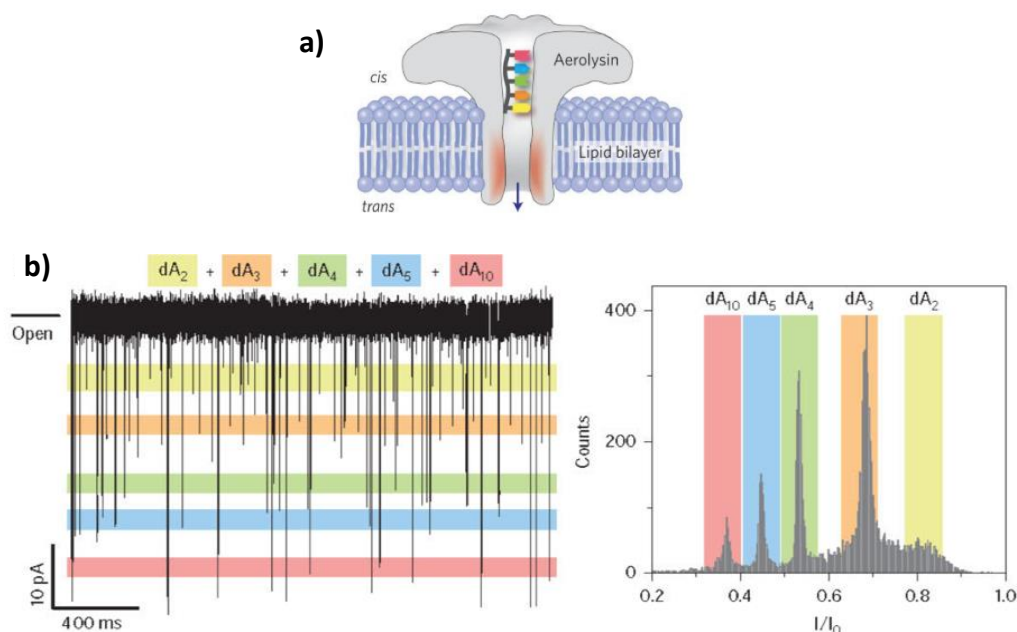


Figure 29 - Discrimination de courts oligonucléotides avec une résolution de séparation d'un nucléotide avec le nanopore d'aérolysine sauvage. (a) Schéma d'un nanopore d'aérolysine inséré dans une bicouche de lipides en présence d'oligonucléotides transportés dans le pore. (b) Portion d'une trace de courant typique obtenue dans le cas de l'analyse d'un mélange d'oligonucléotides d'adénine (dA) d'une longueur de 2, 3, 4, 5 et 10 nucléotides ajoutés du côté cis du nanopore d'aérolysine. Les différents nucléotides sont distingués selon l'amplitude du blocage du courant que chacun engendre. L'histogramme du courant relatif (I/I_0) correspondant à l'analyse de la trace du courant, montrant l'apparition de cinq populations caractéristiques des cinq types de nucléotides présents en solution. (Adaptée de [27]).

Par la suite, les mêmes auteurs ont démontré que le changement d'un nucléotide individuel au sein d'un oligonucléotide est détecté avec le nanopore d'aérolysine. L'analyse d'oligonucléotides constitués de 4 adénines dans lequel le quatrième nucléotide est à chaque fois substitué par une cytosine, une guanine, ou une thymine, ont été analysés avec le nanopore d'aérolysine individuellement ou dans un mélange, et sont parvenus à être discriminés à partir des amplitudes et des durées de blocage spécifiques pour chacune des substitutions (**Fig. 30**). Certaines substitutions restent difficiles à distinguer. C'est le cas de CA3 et TA3 dont les amplitudes de blocage du courant sont très proches pour être distinguées à partir des traces du courant obtenues en temps réel.

Il a été également montré l'identification des différentes bases de l'ADN avec des homoséquences ou heteroséquences d'ADN [279] ; mais également de la méthylation de l'ADN à partir de sérum [266].

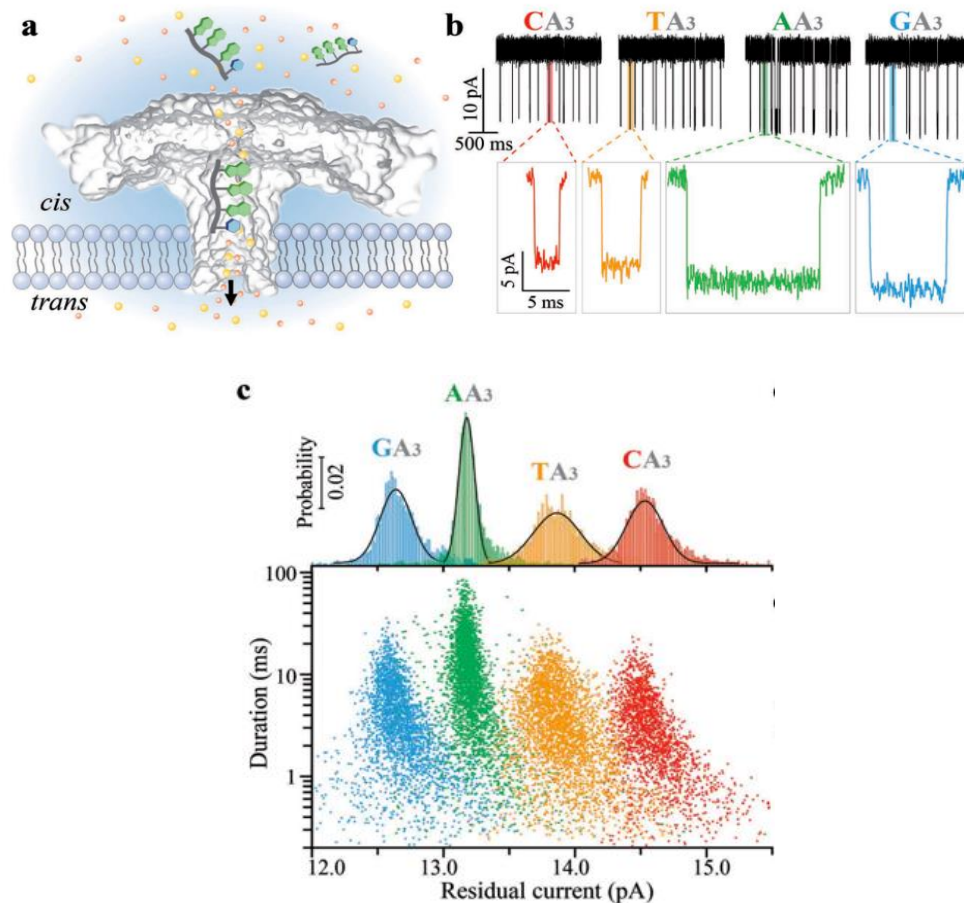


Figure 30 - Discrimination d'oligonucléotides d'adénine présentant une substitution d'un nucléotide individuel. (a) Schéma représentatif d'un nanopore d'aérolysine inséré dans une bicouche de lipides en présence d'oligonucléotides d'une longueur de 3 adénine (A3) auxquels est lié à leur extrémité un quatrième nucléotide dont la nature chimique varie parmi les 4 nucléotides : A (AA3), T (TA3), C (CA3) ou G (GA3). (b) Portions de traces de courant et (c) histogrammes du courant résiduel (haut) et nuages de points de la distribution des durées de blocage en fonction du courant résiduel (bas), dans le cas de l'analyse des oligonucléotides AA3, TA3, CA3, et GA3. Les résultats montrent que chaque type de nucléotide induit des blocages du courant distinguables en termes d'amplitudes des blocages du courant, et de durées des blocages du courant. (Adaptée de [279]).

3.5 Résumé

L'aérolysine est un nanopore qui a montré son utilité pour la détection de différents types de molécules. Il a plus particulièrement montré son pouvoir de résolution supérieur pour la détection et la discrimination en taille de polymères synthétiques, ainsi que d'oligonucléotides, par rapport à d'autres canaux, avec une résolution de séparation d'une seule unité monomérique au sein d'un mélange polydisperse ; et pour la détection de la variation d'un nucléotide individuel au sein d'un oligonucléotide.

Dans les travaux décrits dans cette partie, il a été mis en évidence que l'utilisation de conditions expérimentales de hautes concentrations en sels (1-4 M) et de tensions électriques bien définies, permettent d'être dans un régime d'analyse dans lequel, d'une part, la chaîne analysée est entièrement accommodée à l'intérieur du nanopore. De cette façon, chacun des monomères de la chaîne logé dans le pore contribue au blocage du courant. D'autre part, les temps d'interaction mesurés entre les chaînes analysées et le pore sont maximisés (de l'ordre de plusieurs millisecondes), ce qui permet de réduire la dispersion mesurée autour de la durée moyenne des blocages du courant de chaque espèce. Ensemble, ces critères sont à l'origine des signaux spécifiques mesurés pour chaque espèce de polymère d'une longueur donnée.

C'est dans ce contexte que dans le cadre des travaux de cette thèse, notre choix s'est porté sur l'utilisation du nanopore d'aérolysine que nous avons d'abord utilisé pour évaluer dans quelle mesure la discrimination en taille de molécules avec une résolution d'une unité monomérique peut être appliquée à l'analyse de courts peptides. Ces travaux qui constituent le premier résultat majeur de ma thèse sont présentés par la suite.

Chapitre 2 : Méthode expérimentale

Le principe de la technique utilisée dans cette thèse consiste à mesurer le blocage du courant ionique induit par la présence d'une molécule dans un nanopore inséré dans une bicouche lipidique séparant deux compartiments remplis d'électrolyte. Les caractéristiques des blocages du courant (profondeur, durée, fréquence) dépendent des propriétés de la molécule (nature, longueur, charge, composition chimique, conformation et structure).

Nous présenterons dans ce chapitre la méthode permettant la mise en place des expériences de mesures du transport de molécules uniques à travers un nanopore biologique inséré dans une membrane.

I. Dispositifs expérimentaux de mesure électrique

1.1 Dispositif expérimental classique

Toutes les expériences réalisées reposent sur l'utilisation de la technique d'électrophysiologie, connue aussi sous le nom de la technique des bicouches lipidiques planes. Il existe de nombreuses façons de configurer le dispositif expérimental.

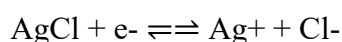
Le premier dispositif expérimental classique qui a principalement été utilisé est décrit à la suite (**Fig. 31**).

1.1.1 Cellule de mesure

Le dispositif inclut l'utilisation d'une cellule de mesure divisée en deux compartiments pouvant contenir chacun un volume d'environ 1 mL de solution ionique (ou électrolyte) (**Fig. 31**). Les deux compartiments sont séparés par une paroi dans laquelle est percé un trou d'un diamètre d'environ 150 μm (**Fig. 31**). La reconstitution d'une bicouche lipidique se fait manuellement par le dépôt d'un film de lipides au niveau du trou micrométrique (**Fig. 31**). La cellule de mesure est fabriquée en polytétrafluoroéthylène (PTFE ou Téflon®) qui est un matériau hydrophobe permettant l'adhésion des lipides. La cellule de mesure est également pré-traitée avant l'expérience de formation d'une bicouche lipidique. Le pré-traitement consiste à étaler et à laisser sécher pendant au moins 30 minutes une couche primaire de la même solution de lipides ($< 1 \mu\text{L}$) au niveau de la paroi et du trou. Cette étape permet de préparer la surface à l'accrochage des lipides pour la formation d'une bicouche lipidique stable.

1.1.2 Electrodes de mesure et chaîne d'acquisition du signal

L'utilisation d'électrodes d'argent (Ag) permet d'appliquer une différence de potentiel et de mesurer le courant ionique circulant entre les deux compartiments. Deux électrodes sont plongées dans chacun des compartiments (**Fig. 31**). Avant d'être utilisées, les électrodes sont préparées en recouvrant leur surface par du chlore (chloridation) pour former les électrodes Ag/AgCl. Le processus de chloridation est important car le potentiel de ces électrodes qui doit rester fixe dépend de la présence des ions chlorure. La réaction chimique correspondante est la suivante :



Il existe différentes méthodes pour convertir les électrodes d'argent en électrodes Ag/AgCl. Les électrodes d'argent sont généralement plongées dans une solution d'hypochlorite de sodium (Javel) riche en chlore. La chloridation des électrodes peut également se faire par voie électrochimique (électrolyse, galvanisation) par exemple en plongeant les électrodes dans une solution de HCl ou de NaCl et en appliquant une différence de potentiel de 0,7 V [296].

Cette étape est indispensable à la mesure d'un courant stable au cours du temps. En effet, le chlorure d'argent est légèrement soluble dans les solutions à haute concentration en chlorure de potassium (KCl). Ainsi, il peut être recommandé que la solution de KCl soit saturée en chlorure d'argent pour éviter la disparition de la couche de chlorure d'argent de la surface de l'électrode.

La qualité de l'électrode a une importance pour le contrôle de paramètres comme la dérive du courant au cours du temps, la stabilité de la mesure, la reproductibilité de la mesure, et le bruit électrique mesuré.

Par convention, l'électrode du compartiment *cis* est reliée à la masse (électrode de référence au potentiel nul) et *via* l'électrode du compartiment *trans*, le potentiel électrique est appliqué et le courant électrique mesuré grâce à une tête de mesure qui contient un convertisseur courant-tension et un préamplificateur (Axopatch 200B patch-clamp amplifieur, Molecular Devices, Sunnyvale, CA, USA) qui reçoit et adapte le signal en améliorant le rapport signal-sur-bruit avant sa transmission vers l'amplificateur principal de la chaîne d'acquisition (**Fig. 31**).

Les électrodes peuvent également être au cours de l'expérience plantées dans des ponts d'agarose pour les protéger. Pour cela un gel d'agarose à 2% (poids/volume) est préparé par solubilisation de la poudre d'agarose dans du KCl à 3M pH 7,5 et est coulé dans des cônes de

10 μ L. Les électrodes sont insérées dans les cônes remplis de gel d'agarose au début de l'expérience et c'est le cône qui est plongé dans le compartiment rempli d'électrolyte.

Le signal amplifié est ensuite envoyé à une chaîne d'acquisition comprenant un système d'amplificateur, un filtre passe bas, un convertisseur analogique – numérique (DigiData 1440A AD-converter, Molecular Devices), une carte d'acquisition et un système de stockage et d'analyse sur ordinateur (**Fig. 31**). La visualisation des traces de courant sur l'ordinateur est assurée par le logiciel Clampex 10.2 (Molecular Devices).

Les paramètres de filtre, de mode de mesure (whole cell ou patch clamp), et de gain, sont contrôlés *via* l'amplificateur. La fréquence d'échantillonnage est fixée *via* le logiciel Clampex. La tension appliquée peut être contrôlée *via* le logiciel Clampex ou manuellement *via* l'amplificateur.

De façon à contrôler la température de l'électrolyte, la cellule de mesure est équipée d'un système de refroidissement à effet Peltier (CL-100, Warner Instruments). Avec cet équipement, la température peut être fixée entre 5 et 45°C.

En fonctionnement normal, la cellule de mesure est enfermée dans une cage de Faraday qui est une cage métallique disposée autour de la cellule de mesure de façon à l'isoler des interférences externes. Durant les expériences réalisées, les courants électriques mesurés sont très faibles (de l'ordre du picoAmpère) et le moindre bruit électrique provenant d'une source extérieure peut perturber et déformer le signal. Il est nécessaire d'installer le dispositif sur une surface qui l'isole des vibrations (tables anti-vibrations) de l'environnement qui peuvent également causer des perturbations du signal.

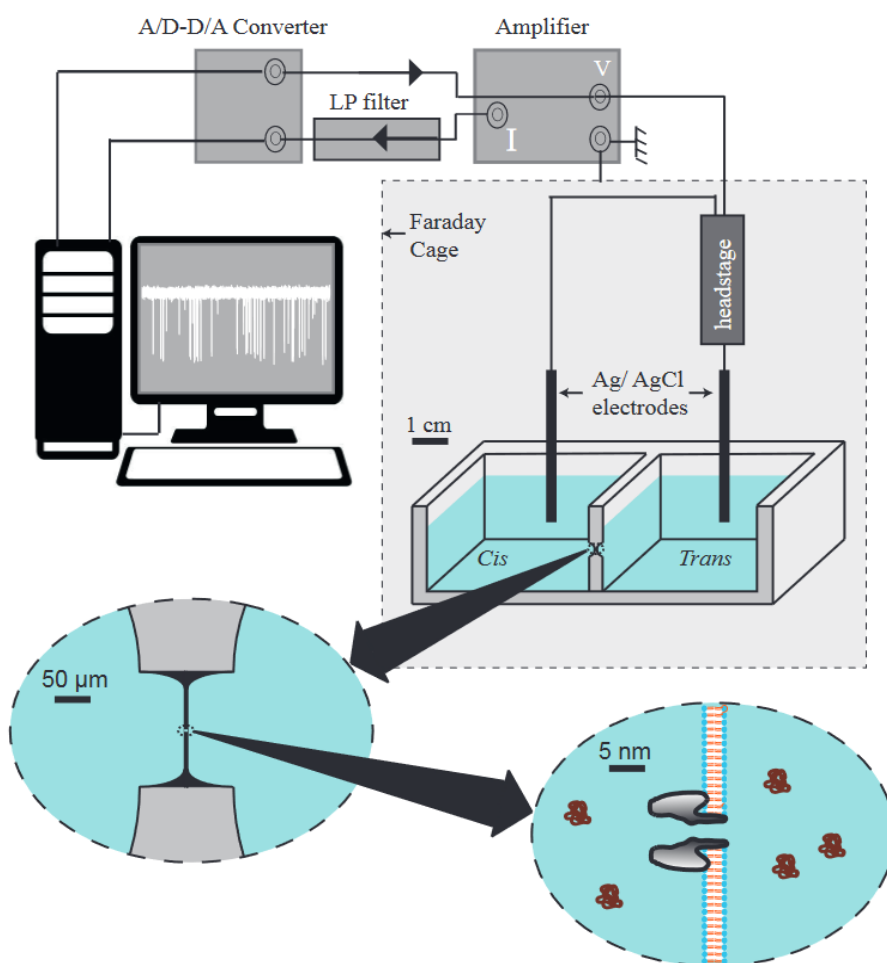


Figure 31 - Schéma d'un montage expérimental classique d'enregistrement électrique par nanopore. (a) Une cellule de mesure est constituée de deux compartiments (cis et trans) qui sont chacun rempli par une solution en électrolyte d'un volume de l'ordre du millilitre et qui sont séparés par une paroi dans laquelle est percée un trou permanent de 150 µm de diamètre. (b) Une membrane isolante correspondant à une bicouche lipidique plane est verticalement formée au niveau du trou micrométrique. (c) Un nanopore individuel correspondant à une toxine qui forme des pores est inséré dans la bicouche de lipides et forme un canal de diamètre nanométrique qui relie les deux compartiments. Des analytes peuvent être ajoutés en solution, dans le compartiment cis, trans, ou dans les deux compartiments. (d) Une paire d'électrodes (Ag/AgCl) est utilisée pour appliquer une tension transmembranaire. L'électrode plongée dans le compartiment cis est reliée à une masse, et l'électrode plongée dans le compartiment trans permet de mesurer le courant ionique circulant à travers le nanopore au cours du temps. (e) Le signal électrique analogique obtenu par la mesure du courant circulant à travers le nanopore est d'abord pré-amplifié par la tête de l'amplificateur, puis de nouveau amplifié par un filtre passe-bas, avant d'être digitalisé et finalement interprété et enregistré par un logiciel informatique d'acquisition contrôlé dans l'ordinateur (f). (g) Les compartiments, les électrodes, et le pré-amplificateur sont placés à l'intérieur d'une cage de Faraday dans le but d'être isolés du bruit électrique environnant.

1.2 Dispositif expérimental haut-débit (Orbit 16, MECA chips)

Le second dispositif que l'on a utilisé est une technologie dont le principe est similaire à celui du dispositif classique mais qui permet la formation simultanée de 16 bicouches lipidiques en parallèles et l'insertion d'un nanopore dans chacune des bicouches formées (**Fig. 32**).

Pour pouvoir réaliser cela, l'entreprise Ionera Technologies a développé des micropuces électroniques (MECA chips, Multi-Electrode Channel Array) constituées de 16 cavités de 50 μm de diamètre et de 8 μm de profondeur reliées chacune à une électrode (**Fig. 32**). Le matériau de base constituant les micropuces correspond à du verre qui est recouvert d'une couche de titane (20 nm) et d'une couche d'or (200 nm) pour former l'électrode. L'or est un des matériaux les plus utilisés pour les électrodes en raison de sa conductivité électrique élevée, sa bonne stabilité chimique, et sa biocompatibilité. Cependant, les métaux nobles comme l'or présentent une faible adhésion aux matériaux comme le verre en raison de leur incapacité à former des interfaces métal-oxyde stables [297]. Pour cela, il est nécessaire d'utiliser une fine couche d'un métal comme le titane qui permet une adhérence de l'or au substrat en verre en créant des interfaces d'oxyde assez facilement. Par-dessus des couches, l'ensemble la surface de la puce est recouverte d'un polymère, le SU-8, sauf au niveau des 16 cavités. Le SU-8 est une résine photosensible communément utilisée dans la fabrication de microsystèmes qui adhère très bien au matériau substrat et est chimiquement inerte une fois qu'elle y est fixée (liaisons covalentes) ; et présente également l'avantage d'être utilisée pour réaliser des motifs par photolithographie.

Les puces sont utilisées dans un dispositif de mesure, l'Orbit 16, développé et commercialisé par l'entreprise Nanion.

Dans le dispositif de mesure, la puce est posée à la verticale sur un support qui permet de positionner ses électrodes aux connecteurs permettant de la relier à l'amplificateur (**Fig. 32**). Chacune des 16 cavités forme un compartiment *trans* individuel. Le compartiment *cis* est commun aux 16 cavités. Le compartiment *cis* est créé en plaçant une chambre en PTFE à la surface de la puce qui encadre les 16 cavités. La chambre est maintenue stable à la surface de la puce par l'utilisation d'un joint en silicone positionné à sa base. La chambre formée est le plus souvent remplie d'un volume de 150 μL bien qu'elle puisse contenir un volume allant jusqu'à 400 μL . En comparaison avec le dispositif classique présenté plus haut, la possibilité de travailler ici avec un volume environ 10 fois plus petit permet de réaliser une économie des analytes étudiés qui peuvent être coûteux comme c'est le cas de certains peptides et de certaines

protéines. L'électrode reliée à la masse est plongée à la surface de la chambre remplie d'électrolyte.

Avec ce dispositif, les bicouches lipidiques ne sont pas formées manuellement. Un processus automatisé de formation des bicouches de lipides par rotation a été mis en place. Le principe consiste à déposer une goutte de lipide à la surface de la micropuce, au plus près des cavités. Un barreau aimanté en PTFE est déposé à la surface de la puce et entre en rotation sous l'action d'un champ magnétique ce qui permet d'étaler les lipides sur l'ensemble des 16 cavités et aboutit à la formation spontanée des bicouches de lipides dans les différentes cavités. La vitesse et le sens de rotation du barreau aimanté peut être contrôlé manuellement. Le dispositif peut être utilisé en deux modes.

Le premier mode permet de visualiser le signal et d'agir sur les 16 canaux en même temps. Dans ce mode, l'acquisition du signal se fait par l'amplificateur interne à l'Orbit 16, le Tecella 16-multichannel (Foothill Ranch, CA, USA), et le suivi du signal par le logiciel Tecella Lab version 0.84. Ce mode est souvent utilisé dans un premier temps pour suivre la formation des membranes simultanément dans les 16 cavités.

Le second mode permet de passer d'un canal à l'autre, de les contrôler un à un, et de visualiser ce qu'il se passe dans chacun d'eux. Dans ce mode, l'acquisition du signal se fait par un amplificateur externe comme celui utilisé dans le dispositif classique (Axopatch 200B patch-clamp amplifier, Molecular Devices, Sunnyvale, CA, USA), et le logiciel GePulse (Michael Pusch, Italian National Research Council) est utilisé pour suivre le signal. Ce mode est généralement choisi lorsque des pores se sont insérés dans des membranes lipidiques et que l'on souhaite suivre et enregistrer le signal du pore vide ou en présence d'analytes d'intérêt.

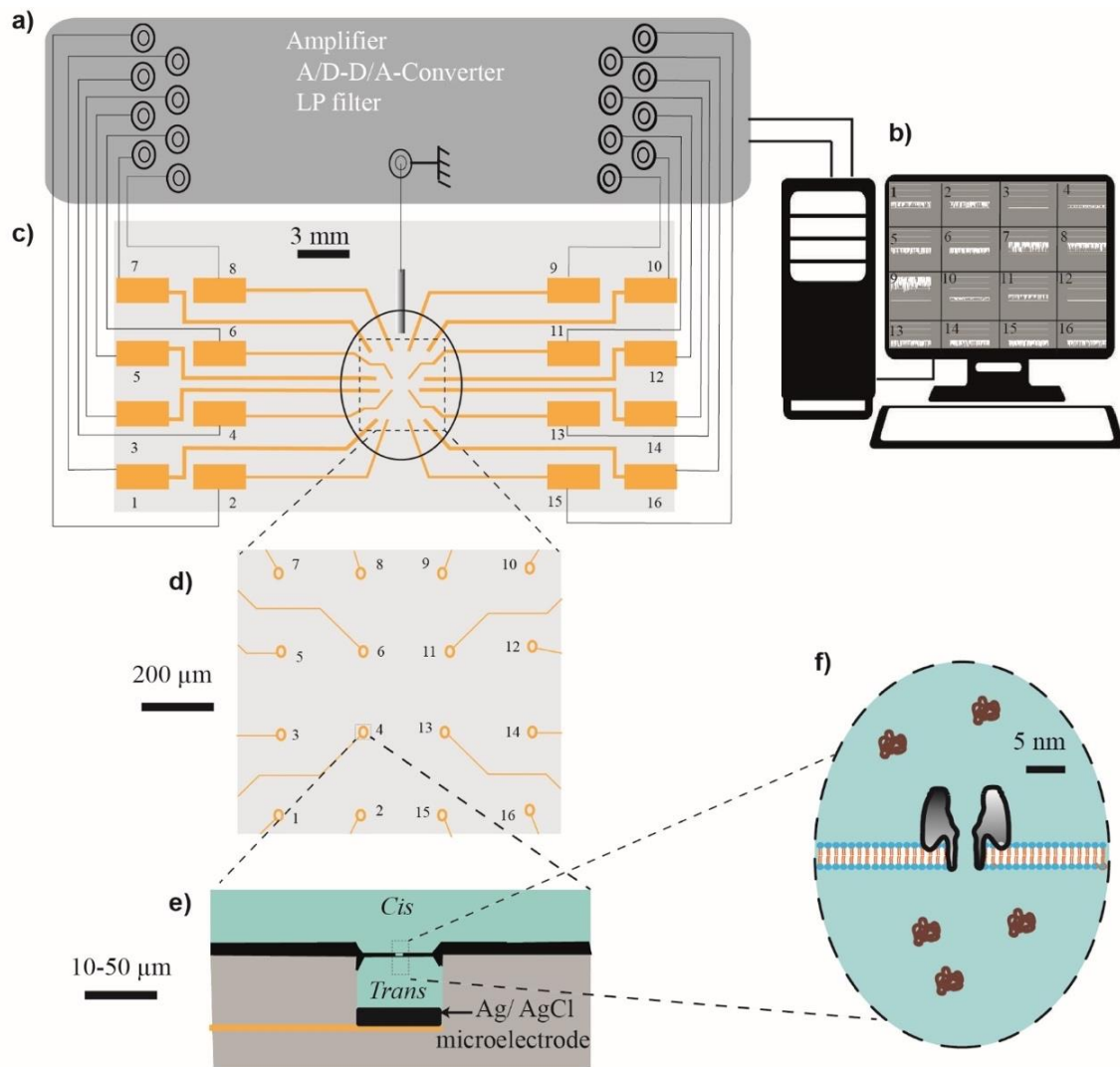


Figure 32 - Schéma d'une micropuce électronique à multiples cavités, utilisée pour l'enregistrement électrique simultané de signaux provenant de multiples nanopores. L'amplificateur multi-cavités (a), avec un convertisseur Analogique-Digital (AD), et un filtre intégré, est contrôlé par un ordinateur (b), et mis en contact avec les cavités de la micropuce (MECA 16) (c) via ses 16 points de contacts numérotés de 1 à 16. Des lignes d'or relient les points de contact aux microélectrodes (MECs) de chaque cavité. Les cavités sont situées dans une zone centrale de 1 mm² (d). L'insert (e) montre le schéma d'une MEC qui est formée par lithographie photorésistante sur un substrat en verre. Le fond de chaque MEC contient une électrode Ag/AgCl nanoporeuse formée par le dépôt microgalvanique d'Argent suivi d'une chloridation. Ensuite, la membrane peut être formée par "peinture" actionnée à distance de lipides ou par la formation de bulles. La reconstitution de pores (f) se fait de la même manière que dans le montage classique.

II. Membranes de lipides

2.1 Composition chimique et organisation des bicouches de lipides

La membrane plasmique des cellules limitant le milieu intérieur du milieu extérieur est principalement composée de lipides et de protéines dont les proportions varient d'une cellule à l'autre. La membrane plasmique est une bicouche de lipides d'un diamètre de 4 à 6 nm, c'est-à-dire un assemblage de deux feuillets de lipides dont les deux constituants principaux sont les phospholipides et les protéines.

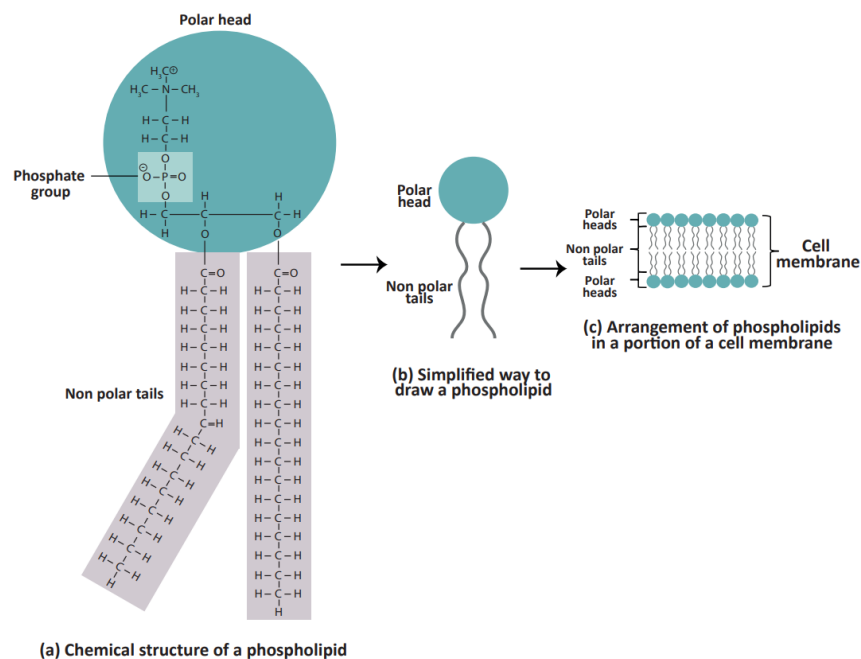


Figure 33 - Représentation schématique et structure chimique d'un phospholipide et de son arrangement dans une bicouche de lipides.

Les phospholipides sont les lipides les plus représentés au sein des bicouches lipidiques. Les phospholipides sont composés d'une tête polaire hydrophile et d'une queue apolaire hydrophobe composé de deux chaînes d'acides gras de longueurs différentes (**Fig. 33**). Une de ces deux chaînes est dite insaturée car elle porte une ou plusieurs doubles liaisons. Dans une bicouche lipidique, les molécules de phospholipides sont réparties de manière asynchrone dans les deux feuillets membranaires. Les têtes polaires sont dirigées vers l'extérieur et les queues apolaires vers l'intérieur (**Fig. 33**). Les bicouches lipidiques peuvent également être constituées d'autres types de lipides, comme les glycolipides ou les molécules de cholestérol. Les glycolipides qui sont généralement exclusivement retrouvés dans le feuillet externe des bicouches de lipides, jouent un rôle dans l'interaction de la membrane plasmique avec

l'environnement grâce à la chaîne glucidique les composant, et qui est accessible dans le milieu extérieur (*e.g* récepteur de protéines). Le cholestérol est un lipide également très abondant (30 à 50%) dans les membranes lipidiques, mais est présent exclusivement chez les cellules eucaryotes et joue un rôle de rigidification des membranes en s'intercalant entre les molécules de phospholipides.

L'acquisition d'une organisation en bicouches de lipides dépend de la structure de la molécule de lipide engagée. Dans le cas où le lipide est constitué d'une queue hydrophobe composée de deux chaînes d'acide gras, l'organisation obtenue correspond à une bicouche lipidique plane ou sphérique (liposomes) (**Fig. 34**). Dans le cas où la queue hydrophobe n'est constituée que d'une chaîne d'acide gras, les lipides s'organisent en une monocouche, qui dans certains cas peut être sphérique pour former des micelles (**Fig. 34**).

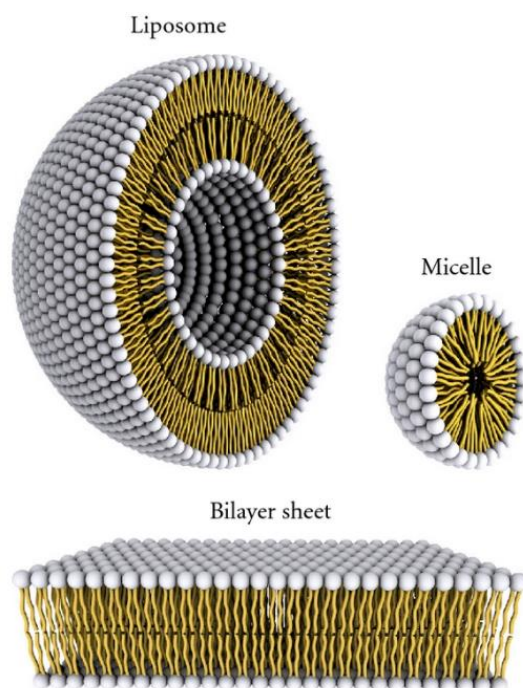


Figure 34 - Modes d'organisation des lipides. Schéma représentatif d'une bicouche de lipides sphérique (a), en micelle (b), ou plane (c). (Adaptée de [298])

Le chimiste Langmuir obtient en 1917 [299] les premières membranes artificielles composées de phospholipides en utilisant les propriétés polaires des molécules et en montrant que les phospholipides peuvent s'étaler à la surface de l'eau en orientant leur tête hydrophile au contact de l'eau. Quelques temps plus tard, en 1925, les scientifiques Gortel et Grendel posent l'hypothèse que la seule limite stable pouvant exister entre deux compartiments aqueux correspond à une bicouche lipidique [300].

Par la suite, les progrès de la microscopie ont notamment permis de montrer que les bicouches lipidiques sont dynamiques en observant que les deux composants majoritaires, les lipides et les protéines, sont mobiles à l'intérieur de la membrane plasmique (**Fig. 35**). Les lipides parviennent à se déplacer dans la membrane de deux façons, soit par une diffusion latérale dans le feuillet dans lequel ils font partie ; soit par une rotation simple sur eux-mêmes (flip-flop) (**Fig. 35**). Ces déplacements sont principalement assurés par des protéines spécifiques. C'est ainsi que les lipides synthétisés dans le cytosol de la cellule parviennent à intégrer le feuillet externe de la membrane plasmique.

Les mouvements de déplacements des lipides dans les membranes dépendent de la fluidité des membranes, qui dépend elle-même de paramètres comme la température, le pH, mais également de la composition lipidique de la membrane.

Par exemple, une augmentation de la température contribue à augmenter la fluidité de la membrane. La présence d'acides gras insaturés dans les queues hydrophobes des phospholipides, diminue les interactions entre les chaînes lipidiques, ce qui aura également tendance à augmenter la fluidité des membranes ; alors que les acides gras saturés auront tendances à rigidifier la membrane, de la même façon que la présence d'une forte proportion de cholestérol.

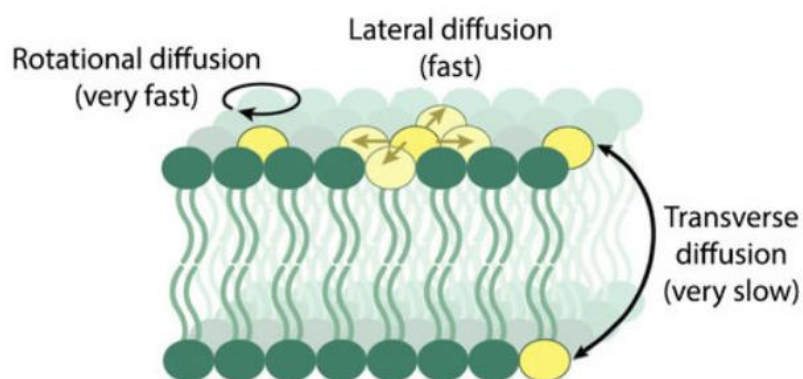


Figure 35 - Schéma représentatif des mouvements de diffusion des lipides dans une membrane. Les lipides diffusent de façon latérale ou par rotation d'un niveau à un autre (flip-flop) dans les membranes lipidiques. (Adaptée de [301])

2.2 Méthodes de reconstitution des bicouches de lipides

En 1962, Montal et Mueller et leurs collègues montrent pour la première fois la possibilité de former des bicouches lipidiques planes « *in vitro* » [43]. La technique utilisée par Mueller consiste à solubiliser des lipides dans un solvant non polaire et stable puis à peindre et brosser ces lipides sur une surface de Téflon disposant d'un trou de l'ordre d'une centaine de microns

en son centre. Ce dépôt de lipides au niveau du trou micrométrique entraîne la formation d'une bicouche lipidique [43,302]. La formation de la bicouche se fait par un assemblage spontané des molécules de lipides qui vont orienter les parties hydrophobes vers l'intérieur et les parties polaires vers le solvant [303] qui lui se réorganise pour être réparti à la fois autour du trou et autour de la bicouche lipidique [304] (**Fig. 36**). Un exemple d'un type de solvant utilisé pour solubiliser les lipides peut être un alcane de type décane ou hexadécane. C'est en fonction du type de solvant utilisé pour solubiliser les lipides que varie l'épaisseur de la bicouche lipidique formée. L'utilisation de l'hexadécane ($C_{16}H_{34}$) permettra par exemple l'obtention de bicouches de lipides plus minces que l'utilisation de décane ($C_{10}H_{22}$). En effet, l'hexadécane qui est une molécule plus grande que le décane se retrouvera en dehors de la bicouche de lipides alors que les plus petites molécules de décane auront tendance à rester entre les feuillettes de phospholipides lors de la formation de la bicouche lipidique. La structure des lipides peut également avoir une influence sur l'épaisseur de la bicouche lipidique [304].

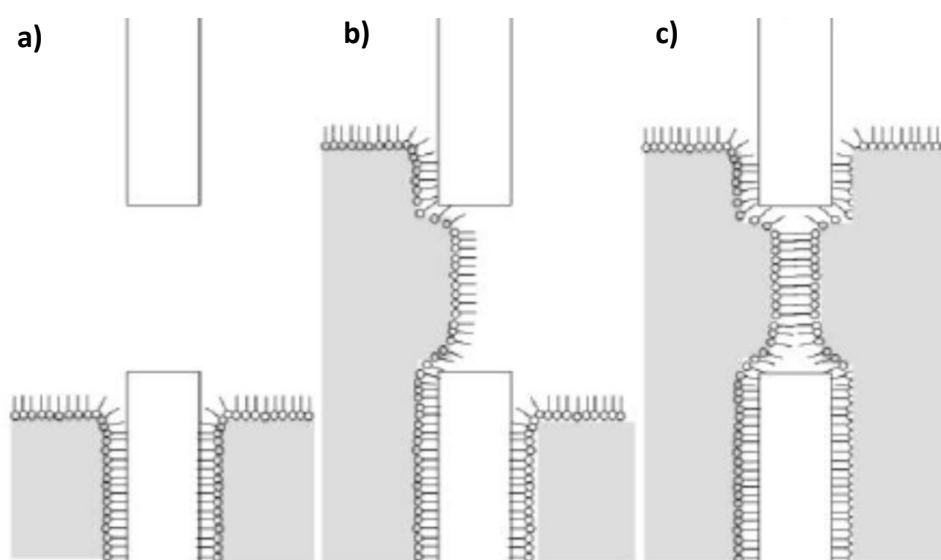


Figure 36 - Formation d'une bicouche lipidique plane par la méthode Montal-Mueller. Schéma représentatif de la formation d'une bicouche de lipides plane par la méthode de Montal-Mueller qui consiste à en élever le niveau de la solution sur le côté gauche de l'ouverture (a), puis sur son côté droit (b), ce qui aboutit à la formation de la membrane (c). (Adaptée de [305])

Dans le cas de l'utilisation du dispositif classique, la reconstitution des bicouches lipidiques planes se fait manuellement.

Le type de lipides utilisé pour former des bicouches lipidiques avec la cellule de mesure utilisée correspond au 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC). Le lipide de synthèse DPhPC a été conçu pour les mesures en électrophysiologie en s'inspirant des lipides

phytanoyles observés dans les membranes cellulaires d'Archée et qui, par leur stabilité et leur durabilité, aident ces organismes à supporter des concentrations extracellulaires élevées de sels ou d'acide, et empêchent la perméabilité à l'eau et le stress osmotique. Cette stabilité physico-chimique élevée des phytanoyles a été attribuée aux chaînes des lipides qui sont entièrement saturées ce qui les rend beaucoup moins sensibles à la dégradation chimique par l'air ou par la lumière. Cela confère aux membranes d'autres propriétés très intéressantes comme une transition des lipides de la phase gel à la phase fluide inexistante à température ambiante. Le lipide DPhPC reste en phase liquide dans une large gamme de température allant de -120 à 120°C. Le lipide DPhPC est également hautement soluble dans les solvants organiques. Enfin, les bicouches lipidiques formées à partir de DPhPC présentent une résistance électrique élevée qui est une condition préalable pour réussir une expérience dans laquelle la conductance du canal ionique est généralement de l'ordre du picoAmpère. Pour ces raisons, les DPhPC sont souvent utilisés comme lipides constitutifs des membranes modèles pour des expériences en électrophysiologie, plus particulièrement pour le suivi de l'insertion de nanopores biologiques dans des bicouches lipidiques et pour leur caractérisation [306].

Les bicouches lipidiques sont formées à partir d'une solution de 1% (poids/volume) de 1,2-diphytanoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids, Alabaster, AL, USA) dissout dans du décane (10 mg/mL, Sigma Aldrich, France) ou dans de l'octane dans le cas du dispositif haut-débit utilisant les micropuces de verres (MECA).

A l'aide d'un pinceau ou d'une micropipette, une goutte de lipide est déposée puis étalée sur le trou de 150 μm de la cuve ayant été préalablement traitée. Sous l'action de la capillarité et des forces de Van der Waals, la bicouche lipidique se forme spontanément.

2.3 Caractérisation électrique des bicouches de lipides

La formation de la bicouche lipidique est suivie électriquement par mesure en temps réel du courant ionique circulant entre les deux compartiments. La mesure d'un courant ionique d'intensité nulle est un indicateur de l'isolation électrique des deux compartiments mais n'est pas forcément un indicateur de la reconstitution d'une bicouche lipidique plane. La qualité de la bicouche lipidique plane est validée par la mesure électrique de sa résistance (qui traduit la résistance de la membrane au passage des ions) et de sa capacité (qui traduit l'épaisseur de la membrane). Dans nos conditions expérimentales, seuls les films de capacité supérieure à 55 pF et de résistance de l'ordre du GigaOhm sont considérés comme des membranes lipidiques et

sont utilisés pour les expériences de transport. Ces membranes ont une épaisseur d'environ 5 nm.

III. Aérolysine recombinante et insertion d'un nanopore dans une membrane

3.1 Production et purification de l'aérolysine

Le pore utilisé dans toutes les expériences réalisées est l'aérolysine. L'aérolysine utilisée est une forme sauvage recombinante produite dans *Escherichia coli* au sein de mon laboratoire d'accueil [24]. La toxine est produite sous forme de pro-aérolysine.

La séquence codante de la pro-aérolysine a été cloné dans le plasmide pET22b-PA (fourni par F. Gisou van der Goot, EPFL, Lausanne). Ce plasmide permet l'expression inductible à l'IPTG de la toxine chez *Escherichia coli* en introduisant un signal de localisation periplasmique. Le système de clonage utilisé ajoute également une étiquette histidine (His-tag) en C-terminal qui facilitera la purification de la protéine. La fraction du périplasme contenant la pro-aérolysine soluble est utilisée pour être purifiée par chromatographie par affinité sur mini-colonnes de Ni^{+2} (His SpinTrap mini-columns, GE Healthcare Life Science).

Pour son utilisation comme nanopore dans nos expériences, la pro-aérolysine doit être activée par clivage du pro-peptide de 41 à 43 acides aminés présent à l'extrémité C-terminal de la protéine. Les monomères activées s'oligomérisent (heptamère) pour former des nanopores d'aérolysine dans les membranes lipidiques. L'activation est réalisée par digestion enzymatique de la pro-aérolysine par la trypsine à température ambiante pendant environ quinze minutes pour cliver le pro-peptide.

3.2 Insertion et propriétés électriques du canal

Une fois active, l'aérolysine est ajoutée à une concentration de l'ordre du nanomolaire en solution dans un seul compartiment pour former des nanopores. Par convention, l'aérolysine est ajoutée dans le compartiment électriquement relié à la masse (compartiment *cis*).

L'insertion d'un nanopore dans la bicouche de lipides formée est suivie en temps réel par le suivi de l'évolution du courant mesuré en fonction du temps. Typiquement, l'insertion d'un nanopore d'aérolysine entraîne un saut de courant d'environ 100 pA pour une différence de potentiel appliquée de +/- 50 mV, à 20°C, et en présence d'une solution ionique de chlorure de

potassium (KCl) à 4 M tamponnée par 5 mM d'HEPES et ajustée à pH 7,5. Cette valeur correspond au courant unitaire du pore d'aérolysine dans ces conditions expérimentales. Ces conditions expérimentales ont été utilisées dans toutes les expériences présentées dans les chapitres n°3 et n°4.

Le nanopore d'aérolysine présente une rectification du courant (**Fig. 37**) dont le pourcentage varie en fonction des conditions expérimentales et se traduit par un courant mieux conduit à travers le pore sous l'action de tensions électriques positives.

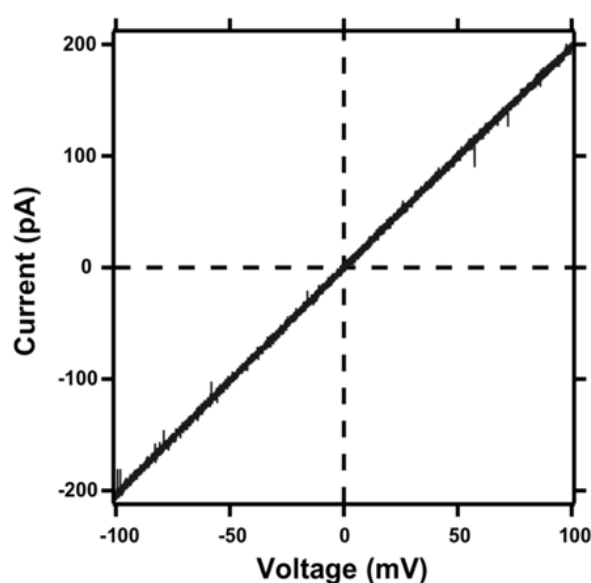


Figure 37 - Rectification du courant dans le nanopore d'aérolysine. Courbe représentative de l'évolution du courant ionique en picoAmpère (pA) en fonction de la tension appliquée du côté cis d'un nanopore d'aérolysine inséré dans une bicouche de lipides en présence d'une solution ionique de KCl à une concentration de 4M, stabilisée à pH 7,5, par une solution tampon d'HEPES à une concentration de 5 mM, et à une température ambiante d'environ 20°C.

Le nanopore d'aérolysine peut s'insérer soit en présentant son chapeau vers le compartiment *cis* ce qui se traduit par la mesure d'un courant légèrement plus grand sous l'action d'une tension négative par rapport à la même valeur de tension positive, ou dans le contraire soit en orientant son chapeau vers le compartiment *trans*. Les pores conservés pour le transport d'analytes correspondent à ceux dont le chapeau est orienté vers le compartiment *trans* et correspondent au cas le plus représentatif.

IV. Détection électrique

Sous l'action d'un champ électrique, une molécule chargée peut être entraînée dans le pore. La présence d'une molécule dans le pore augmente la résistance électrique du pore. Cette augmentation de la résistance électrique est traduite par une chute brusque du courant électrique appelée blocage du courant, impulsion électrique, ou événement.

Pour fixer les ordres de grandeurs et décrire les effets attendus dans nos expériences, nous rappelons ci-dessous quelques éléments du calcul de la résistance électrique d'un pore cylindrique conducteur en présence d'une bille isolante, décrits dans les références [37] et [307]. La géométrie et les variables utilisées sont décrites sur la **Figure 38**.

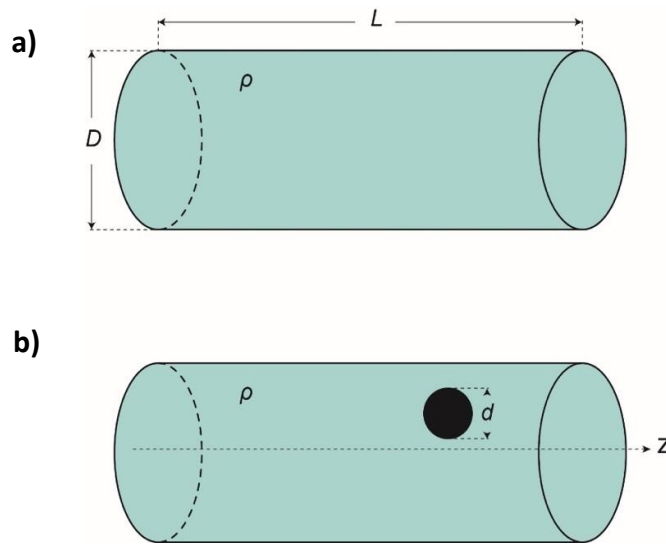


Figure 38 - Schéma représentatif du principe de la résistance électrique d'un cylindre en présence d'une bille isolante. (a) un cylindre de longueur L et de diamètre D rempli d'un fluide conducteur de résistivité électrique ρ . (b) une particule de diamètre d est ajoutée dans le cylindre.

4.1 Résistance électrique d'un pore cylindrique en présence d'une particule sphérique

Maxwell est le premier à avoir calculé la variation de résistance électrique due à la présence de billes sphériques en suspension diluées dans un fluide conducteur [308]. En notant ρ_{eff} la résistivité effective d'une solution de billes infiniment diluée et ρ la résistivité de l'électrolyte pur, Maxwell obtient la relation :

$$\rho_{eff} = \rho \left(1 + \frac{3}{2} f + \dots \right) \quad (1)$$

où f est la fraction volumique de la solution occupée par les billes ($f \ll 1$).

Si le diamètre du pore est très grand devant le rayon de la bille, on peut négliger les effets de confinement et appliquer directement la formule précédente. Comme la résistance d'un pore cylindrique de diamètre D , de longueur L , rempli d'un fluide de résistivité ϱ , est simplement donnée par :

$$R_0 = \frac{4\varrho L}{\pi D^2} \quad (2)$$

En introduisant une particule de diamètre d dans le cylindre, sous les hypothèses de l'équation (1), la résistance R_b du système s'écrit :

$$R_b = \frac{4\varrho_{eff}L}{\pi D^2} = \frac{4\varrho L}{\pi D^2} \left(1 + \frac{3}{2}f + \dots\right) = R_0 \left(1 + \frac{3}{2}f + \dots\right) \quad (3)$$

où f , qui est la fraction volumique de la sphère dans le cylindre, s'écrit :

$$f = \frac{\frac{4}{3}\pi\left(\frac{d}{2}\right)^3}{\pi\left(\frac{D}{2}\right)^2 L} = \frac{2d^3}{3D^2L} \quad (4)$$

D'où :

$$R_b = R_0 \left(1 + \frac{d^3}{D^2L} + \dots\right) \quad (5)$$

Au premier ordre, la variation relative de la résistance électrique s'écrit :

$$\frac{\Delta R}{R} = \frac{R_b - R_0}{R_0} = \frac{d^3}{D^2L} \quad (6)$$

La relation (6) est indépendante de la conductivité de l'électrolyte.

Considérons maintenant l'autre cas limite correspondant à la situation où le diamètre des billes est comparable au diamètre du pore ($d \leq D$), en supposant toutefois la bille centrée sur l'axe du cylindre. Les effets de courbure sont dans ce cas négligeables. Localement, la résistance de la région située entre la paroi du pore et la sphère est celle d'un anneau cylindrique. La contribution dominante à la résistance totale est alors donnée par l'expression générale :

$$R = \varrho \int_{-L/2}^{L/2} \frac{d(z)}{A(z)} dz \quad (7)$$

où $A(z)$ est l'aire de la section annulaire entre la sphère et le cylindre à l'abscisse z . Si l'origine est prise au centre de la sphère, $A(z) = \frac{\pi}{4}(D^2 - d^2 + 4z^2)$ pour $z \leq d/2$. La variation relative de la résistance s'écrit :

$$\frac{\Delta R}{R} = \frac{D}{L} \left[\frac{\sin^{-1}(d/D)}{\sqrt{1 - (d/D)^2}} - \frac{d}{D} \right] \quad (8)$$

Les expressions ci-dessus peuvent être affinées et généralisées à des expressions qui couvrent pratiquement tout le domaine $0 \leq d \leq D$. C'est le cas en particulier de la formule suivante, semi-empirique [37] :

$$\frac{\Delta R}{R} = \frac{d^3}{D^2 L} \left[\frac{1}{1 - 0,8(d/D)^3} \right] \quad (9)$$

La variation relative de la résistance $\frac{\Delta R}{R}$ est proportionnelle à la variation relative du courant $\frac{\Delta R}{R} = \frac{\Delta I}{I} = \frac{I_0 - I_b}{I_0} = 1 - \frac{I_b}{I_0}$, I_0 étant la valeur du courant du pore vide et I_b la valeur du courant du pore en présence d'une particule.

Le courant relatif $\frac{I_b}{I_0}$ s'écrit :

$$\frac{I_b}{I_0} = 1 - \frac{d^3}{D^2 L} \left[\frac{1}{1 - 0,8(d/D)^3} \right] \quad (10)$$

On peut remarquer dans l'équation (10) que les valeurs du courant relatif sont comprises entre 0 et 1.

- La valeur minimale du courant relatif $\frac{I_b}{I_0} = 0$ correspond à la situation où le pore est totalement occupé ($d = D$).
- La valeur maximale du courant relatif $\frac{I_b}{I_0} = 1$ correspond à la situation où le pore est vide ou en l'absence de particule ($d = 0$)

- Les valeurs intermédiaires du courant relatif $0 < \frac{I_b}{I_0} < 1$ correspondent à la situation où le pore est partiellement occupé, elles couvrent pratiquement tout le domaine de taille de particules ($0 < d < D$).

L'expression (10) est en accord qualitatif avec les expériences de discrimination en taille des macromolécules à travers les nanopores protéiques, mais ne décrit pas en toute rigueur les dépendances observées expérimentalement. L'expression (10) est purement géométrique et ne prend pas en compte en particulier les interactions molécule/nanopore, les interactions intramoléculaires, les effets du confinement, et la qualité du solvant sur la conformation des molécules.

4.2 Grandeurs mesurables

Les blocages du courant ou impulsions peuvent être caractérisés par trois grandeurs physiques mesurables ; leur amplitude, leur largeur, et leur nombre :

- L'amplitude des impulsions (blocages du courant) donne une information sur la taille des objets analysés. Plus la taille de l'objet est grande, plus l'amplitude de l'impulsion est importante.
- La largeur des impulsions donne une information sur la dynamique du transport, c'est-à-dire sur le temps de résidence de l'objet dans le pore ou la durée de vie de l'interaction de l'objet avec le pore.
- Le nombre des impulsions reflète la concentration des molécules en solution, et traduit la fréquence de molécules ayant visités l'intérieur du pore.

Les objets détectés sont identifiés à la fois par l'amplitude du blocage du courant et la durée du blocage du courant qu'ils entraînent.

A partir des grandeurs physiques directement mesurables tels que le temps de résidence τ et la fréquence des événements ν , on peut déterminer les constantes cinétiques des interactions bimoléculaires existant entre les analytes et le nanopore, ainsi que l'énergie libre standard ΔG^0 , à l'échelle de la molécule unique. L'observation directe des interactions entre les analytes et le pore protéique revient à décrire l'entrée de l'analyte dans le pore comme la formation d'un « complexe » pore/analyte.

La constante d'association k_{on} est proportionnelle à la fréquence ν des blocages de courant :

$$k_{on}(M^{-1}s^{-1}) = \frac{v}{c} (11)$$

avec c , la concentration de l'analyte.

La constante de dissociation k_{off} est inversement proportionnelle à la durée des blocages de courant τ :

$$k_{off}(s^{-1}) = \frac{1}{\tau} (12)$$

La constante d'association d'équilibre qui donne l'affinité entre le pore protéique et l'analyte s'écrit : $K(M^{-1}) = \frac{k_{on}(M^{-1}s^{-1})}{k_{off}(s^{-1})} (13)$

L'énergie libre standard ΔG^0 s'écrit :

$$\Delta G^0 = -k_B T \ln \left(\frac{1}{K} \right) (14)$$

$k_B T$ étant l'énergie d'agitation thermique.

4.3 Analyse de données

Durant l'expérience en temps réel, les données mesurées sont enregistrées à partir du logiciel Clampex. Les grandeurs physiques mesurées (amplitude, temps et fréquence) nécessitent des analyses statistiques robustes et reproductibles. Pour réaliser ces analyses statistiques, des milliers d'événements de blocage du courant sont enregistrés. On détaille ci-dessous l'ensemble de ces analyses statistiques qui sont toutes réalisées avec le logiciel Igor Pro 6.12A (WaveMetrics, Portland, OR, USA) et des algorithmes développés dans l'équipe.

4.3.1 Histogramme tout point, méthode des seuils, et détection

Comme dans toutes les expériences effectuées sur des molécules uniques, les mesures sont sensibles au bruit thermique ou électrique provenant des appareils. Pour extraire le maximum d'informations du signal électrique enregistré, il faut séparer le bruit du signal électrique qui nous intéresse.

La méthode des « seuils » (**Fig. 39**) est le moyen le plus simple et le mieux adapté pour éliminer le bruit et extraire les informations des traces de courant enregistrées [309]. Cette méthode repose sur le principe qu'une impulsion électrique ne sera prise en compte dans l'analyse

statistique que si son amplitude est en dessous d'un certain seuil correspondant à un niveau de courant défini.

Pour appliquer cette méthode, la première étape consiste à réaliser, à partir d'une trace de courant, un histogramme des intensités de courant (ou la somme du nombre de mesures qui est tracé en semi-log) (**Fig. 39**). C'est l'histogramme « tout point » car il est formé à partir de l'ensemble des points de la trace de courant enregistré dont le nombre dépend de la fréquence d'échantillonnage fixée. La valeur moyenne $\langle I_0 \rangle$ et l'écart-type σ du courant de pore vide sont obtenus par un ajustement (fit) de cette population par une fonction gaussienne (**Fig. 39**).

La seconde étape consiste à appliquer deux seuils. Un premier seuil est fixé à $\langle I_0 \rangle - 4\sigma$, et un second seuil est fixé à $\langle I_0 \rangle - 5\sigma$. Tout événement de chute de courant dépassant le premier seuil (seuil 1 ou $th1$) est compté comme un événement de blocage du courant dans l'analyse (**Fig. 36**). Dans le cas contraire, la chute de courant est éliminée de l'analyse et considérée comme étant due à du bruit électrique. Ainsi, ce premier seuil permet d'éliminer 98% du bruit électrique. Le second seuil (seuil 2 ou $th2$) est nécessaire pour identifier et isoler les événements de blocage du courant uniquement provoqués par l'analyte d'intérêt, c'est-à-dire dans notre cas les homopeptides, et permet d'extraire la valeur de courant de blocage moyen $\langle I_b \rangle$ (**Fig. 39**).

Après avoir identifié les événements de blocage associés au transport d'une molécule dans le pore, une analyse statistique est réalisée à partir de ces données pour extraire plusieurs informations comme le courant I_b lorsque le pore est bouché par une particule permettant ensuite de remonter aux courants relatifs résiduels I_b/I_0 ; la durée de blocage du courant, pour chaque événement de la trace de courant ; et la fréquence de survenue des chutes du courant.

On s'intéresse principalement aux distributions des amplitudes de blocage du courant (histogramme) et aux distributions des durées de blocage du courant (nuage de points).

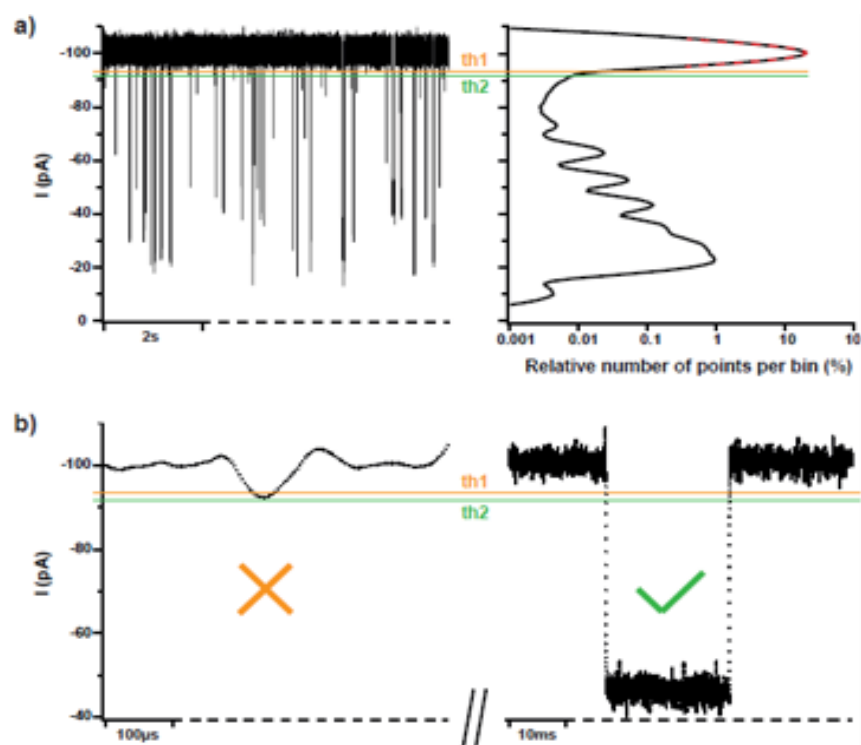


Figure 39 - Méthode d'analyse utilisée pour la détection d'événements de blocage du courant à partir d'un enregistrement électrique. (a) Trace de courant et histogramme de distribution des points de l'ensemble des points de la trace de courant. (b) Trace de courant montrant un événement non retenu pour l'analyse statistique (gauche) et un événement retenu pour l'analyse statistique (droite). La détection des blocages du courant provoqués spécifiquement par les peptides est basée sur une méthode de seuils : un blocage du courant est considéré comme potentiellement provoqué par un peptide si son niveau de courant est plus petit que le seuil 1 (th1) et plus petit que le seuil 2 (th2). Au contraire, si le niveau de courant d'un événement de blocage du courant est supérieur à ces deux seuils, alors il n'est pas considéré comme étant un blocage entraîné par un peptide.

4.3.2 Caractérisation de l'amplitude ou du courant relatif résiduel

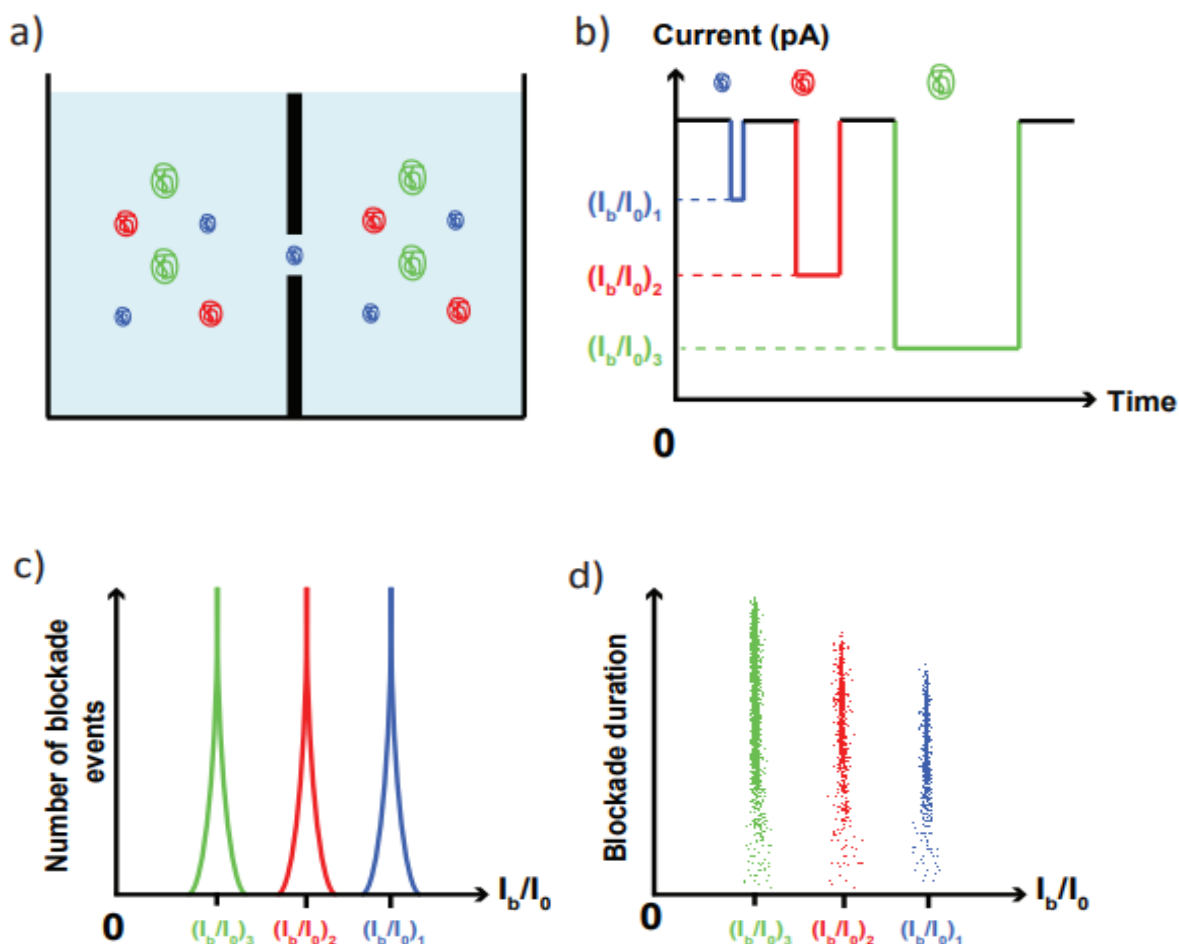
Le premier paramètre auquel on s'intéresse pour caractériser des analytes ajoutés en solution en présence d'un nanopore est l'amplitude. Pour cela, l'amplitude est traduite en courant relatif résiduel qui correspond au ratio du courant mesuré lorsque le pore est vide (I_0), et le courant mesuré lorsqu'un analyte bouche le pore (I_b). Par rapport à l'amplitude de blocage du courant, le courant résiduel reflète le courant restant lorsqu'un blocage du courant survient.

L'histogramme du courant relatif résiduel correspond à la représentation du nombre d'événements en fonction des valeurs de I_b/I_0 (Fig. 40). L'histogramme est construit avec un certain nombre de bins, d'une épaisseur définie, dont le choix dépend du niveau de détail recherché et du nombre d'événements représentés. Dans les résultats présentés par la suite, nous

avons choisi d'utiliser, pour réaliser les histogrammes de I_b/I_0 , une épaisseur de bin de 0,002 (500 bin entre 0 et 1 ; cf. Chapitre n°3), et de 0,001 (1000 bin entre 0 et 1 ; cf. Chapitre n°4).

Cette représentation permet de voir si une population se dessine ce qui indiquerait la présence d'un analyte en solution entraine des événements d'amplitudes reproductibles (**Fig. 40**). Sur la représentation schématique en **Figure 40**, on peut voir sur l'illustration que plus l'analyte est volumineux, plus l'amplitude de bouchage est importante et plus le courant résiduel est faible.

Dans le cas où une population se dessine sur l'histogramme, il est possible de déterminer précisément sa valeur moyenne $\langle I_b/I_0 \rangle$ et sa déviation standard σ en réalisant un ajustement Gaussien de la population. On considérera que l'ensemble de la population est compris entre $\langle I_b/I_0 \rangle \pm 3\sigma$.



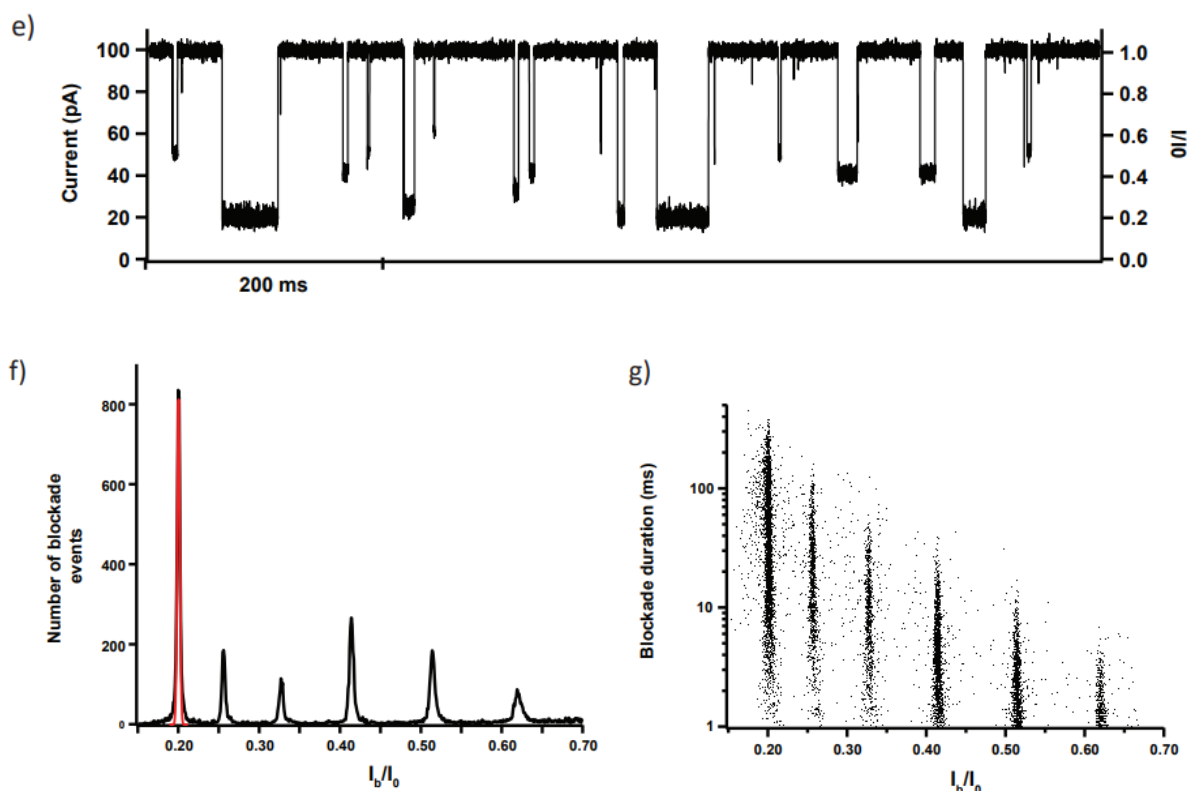


Figure 40 - Principe de l'analyse d'une population polydisperse par un nanopore. (a) Schéma représentatif d'une cuve constituée de deux compartiments séparés par une paroi au niveau de laquelle un nanopore est percé. Un échantillon polydisperse est ajouté en solution. (b) L'entrée d'une molécule d'une taille donnée dans le nanopore engendre un blocage dont l'amplitude est caractéristique de la taille de la molécule. (c) Histogramme du courant relatif résiduel I_b/I_0 dans le cas de l'analyse de la trace de courant obtenue, montrant 3 pics à des valeurs de I_b/I_0 définies et qui correspondent aux trois niveaux d'amplitudes observés sur la trace de courant ; reflétant la présence des trois populations en solution. (d) Distribution des durées de blocage du courant en fonction de I_b/I_0 (nuage de points) associée, montrant un alignement des points au niveau des mêmes valeurs de I_b/I_0 que les trois populations observées sur l'histogramme en (b). La présence de molécules de différents volumes est à l'origine de blocages de courant d'amplitudes (b, c) et de durées (d) variables. (d, e, f) Exemple de l'analyse d'un échantillon polydisperse de peptides avec le nanopore d'aérolysine montrant une portion d'une trace de courant en fonction du temps au niveau de laquelle sont visualisés des blocages du courant d'amplitudes et de durées variables (a). L'analyse statistique de l'ensemble de la trace de courant enregistrée durant cette analyse, permet d'obtenir l'histogramme de I_b/I_0 et la distribution des durées de blocage en fonction de I_b/I_0 , qui montrent l'apparition de 6 pics et de 6 concentrations de nuages de points aux mêmes valeurs de I_b/I_0 , distincts. Ces pics correspondent aux 6 populations présentes en solution. La valeur moyenne $\langle I_b/I_0 \rangle$ de chaque population est déterminée par un fit Gaussien (ligne rouge) sur l'histogramme en (f).

4.3.3 Caractérisation des durées de blocage du courant

Le second paramètre auquel on s'intéresse pour caractériser les analytes étudiés correspond aux durées de blocages des événements enregistrés. Contrairement aux amplitudes, les durées de bouchage présentent le plus souvent une large gamme de distributions de temps ce qui ne permet pas directement d'identifier un analyte à partir du temps de blocage du courant qu'il produit en temps réel. On s'intéresse plutôt au calcul des durées moyennes.

Plusieurs représentations sont utilisées pour caractériser les durées de blocage des événements.

La première représentation la plus classique correspond à la distribution du nombre d'événements comptés en fonction des durées des blocages du courant (**Fig. 41**). Un ajustement exponentielle de la distribution permet d'obtenir une valeur de la durée moyenne de bouchage.

La seconde représentation correspond à un nuage de points de la distribution des durées des blocages du courant détectées en fonction du courant relatif résiduel I_b/I_0 (**Fig. 41**). Cette représentation permet de caractériser de façon qualitative la distribution des temps d'une population.

4.3.4 Caractérisation des fréquences de blocage du courant

Le temps inter-événement (T_i) correspond au temps entre le début d'un premier événement et le début de l'événement suivant. L'inverse du T_i donne une information sur la fréquence d'apparition des événements de blocage du courant. La fréquence de blocage peut ainsi être déterminée de plusieurs manières à partir des T_i .

La première représentation la plus classique consiste à construire la distribution du nombre d'événements en fonction des temps inter-événement (**Fig. 41**) et de faire un ajustement de la distribution permettant de donner une information sur la fréquence moyenne.

Une autre façon de déduire la fréquence moyenne est de façon arithmétique en divisant le nombre total d'événements comptés par le temps total d'enregistrement.

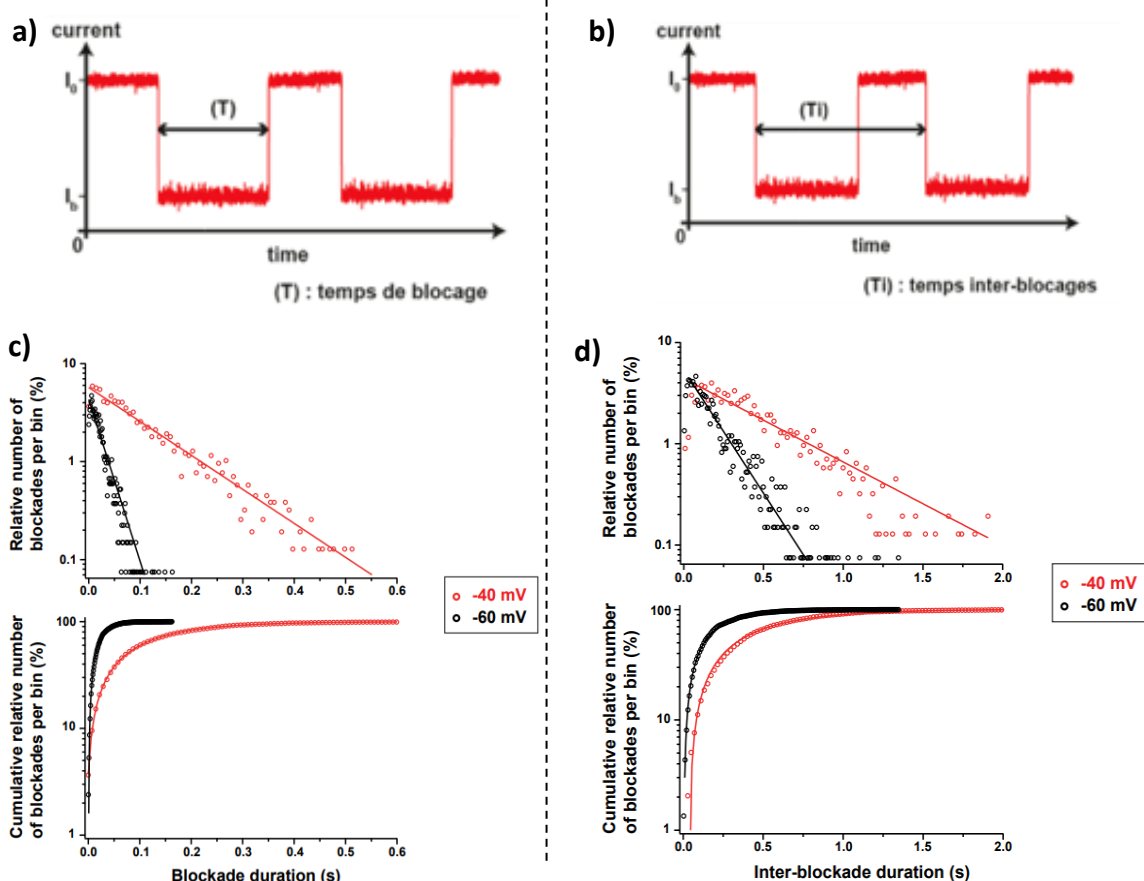


Figure 41 – Exemple d’une analyse statistique des durées et des fréquences des blocages du courant. Représentation schématique d’une portion d’une trace de courant en fonction du temps illustrant la durée T (a) et la durée inter-événements (b) des blocages du courant. Histogrammes (haut de la sous figure) et histogrammes cumulatifs (bas de la sous figure) des durées des blocages du courant (c) ou des durées inter-événements (d) ; obtenus pour deux tensions électriques de -40 mV (rouge) et -60 mV (noir) dans le cas de l’interaction d’homopeptides de 10 acides aminés (1,2 μ M) avec le nanopore d’aérollysine. Les expériences ont réalisées en présence de KCl 4 M, HEPES 5 mM, pH = 7,5, à 20°C.

4.4 Expériences

Pour l’ensemble des expériences et résultats présentés dans les chapitres n°3 et n°4, nous avons utilisé le dispositif expérimental vertical et le dispositif de mesure à haut débit qui ont été précédemment décrits (cf. chapitre 2, 1.1 et 1.2).

4.4.1 Conditions expérimentales

L’ensemble des expériences ont été réalisées selon la méthode expérimentale présentée dans le chapitre n°2. L’électrolyte utilisé dans ces expériences correspond à une solution de chlorure de potassium (KCl) à une concentration de 4 M, tamponnée avec une solution d’HEPES préparée à une concentration de 5 mM (chapitre n°3) ou de 25 mM (chapitre n°4), et ajustée à

un pH de 7,5 par l'utilisation d'une solution de KOH préparée à 6 M. Les membranes de lipides ont été formées comme décrit dans le chapitre n°3 à partir de lipides diphytanoyl-phosphocholine en poudre préparés à 1% par solubilisation dans une solution de décane. Pour l'ensemble des expériences, un dispositif de contrôle de température a été utilisé et nous a permis de fixer la température à $20,0 \pm 0,5^\circ\text{C}$. La préparation d'aérolysine active a été réalisée par digestion par la trypsine à température ambiante durant 15 minutes comme décrit dans le chapitre n°2. Dans ces conditions expérimentales, l'insertion d'un nanopore individuel dans la membrane de lipides génère le passage d'un courant ionique d'environ -100 pA pour une différence de potentiel négative appliquée de -50 mV. Toutes les expériences ont été réalisées sous l'action d'une différence de potentiel de -50 mV, sauf cas contraire indiqué.

4.4.2 Acquisition du signal et analyse de données

L'acquisition du signal a été réalisée à une fréquence d'échantillonnage de 250 kHz (4 μs) ; avec un filtre appliqué de 5 kHz (200 μs) ; et un gain de $\times 10$ (10 mV/pA).

L'analyse des données a été réalisée selon la méthode précédemment décrite (cf. chapitre 2, 4.3).

4.4.3 Analytes

Les objets analysés dans la première partie des travaux, présentée dans le chapitre n°3 sont des homopeptides et des hétéropeptides.

Les homopeptides utilisés sont des peptides dont la composition (Nter-(X)_n-Cter) consistent en l'enchaînement d'acides aminés (X) d'une longueur comprise entre $n=5$ et $n=10$.

Les hétéropeptides utilisés sont des peptides d'une longueur de 10 acides aminés (Nter-(X₁)₅-(X₂)₅-Cter) ici formés par l'enchaînement de 5 acides aminés de même nature chimique (X₁) suivis par l'enchaînement de 5 acides aminés d'une autre nature chimique (X₂).

Nous avons étudié plus précisément :

- i) des homopeptides d'arginine (R_n) d'une longueur de $n = 5, 6, 7, 8, 9$, et 10 acides aminés (R₅₋₁₀) ;
- ii) des homopeptides de lysine (K_n) d'une longueur de $n=10$ acides aminés (K₁₀) ;

iii) des hétéropeptides d'une longueur de 10 acides aminés composés de 5 acides aminés de lysine suivis par l'enchaînement de 5 acides aminés d'arginine (K_5R_5).

En accord avec le pHi des acides aminés étudiés (arginine et lysine), chacun d'eux porte une charge positive à pH 7,5 ($pHi_{(arginine)} = 10,76$; $pHi_{(lysine)} = 9,74$).

L'ensemble de ces peptides sont synthétisés par notre fournisseur (Protéogénix, Schiltigheim, France) avec une pureté supérieure à 98%. Ces peptides sont également tous fournis sous une forme lyophilisée et préparés à une concentration de 1 mM par solubilisation dans une solution d'HEPES préparée à une concentration de 5mM et ajustée à un pH de 7,5.

Les molécules analysées dans la deuxième partie des travaux présentée dans le chapitre n°4 sont des hétéropeptides nommés à la suite X_{R7} ($NH_2-R-R-R-R-R-R-R-X-COOH$) dont la composition correspond à un enchaînement de 7 arginine (R_7) suivi d'un huitième acide aminé (X) dont la nature chimique varie parmi les 20 acides aminés protéinogènes. Nous avons également utilisé des peptides X_{R7} dans lesquels des modifications chimiques ont été introduites sur l'acide aminé X , plus particulièrement les peptides suivants : $(NO_2)-Y_{R7}$, $(P)-Y_{R7}$, et $(sulfoxide)-M_{R7}$.

L'ensemble de ces peptides sont synthétisés par notre fournisseur (Clinisciences, Schiltigheim, France) avec une pureté supérieure à 98%. Ces peptides sont également fournis sous une forme lyophilisée et sont préparés à une concentration de 1 mM par solubilisation dans une solution d'HEPES préparée à une concentration de 5 mM et ajustée à un pH de 7,5.

Chapitre 3 : Discrimination d'homopeptides avec une résolution d'un acide aminé par le nanopore d'aérolysine

I. Introduction

L'objectif de ce chapitre est de présenter la première partie des travaux de recherche de ma thèse sur la discrimination en taille de courts peptides avec une résolution d'un acide aminé obtenue avec le nanopore d'aérolysine sauvage. Nous décrivons d'une façon chronologique, comment l'analyse d'un échantillon peptidique monodisperse nous a mené, d'une part, à comprendre et montrer qu'il est possible d'obtenir la discrimination de courtes chaînes peptidiques avec une résolution d'un acide aminé par l'utilisation de ce nanopore ; et d'autre part, qu'il est possible de détecter des fragments peptidiques présents à l'état d'impuretés dans ces échantillons non identifiés par les techniques conventionnelles d'analyse. Ces résultats constituent la première démonstration d'une telle résolution et sensibilité dans le cas de l'analyse de peptides par nanopore, et ont été publiés en Mars 2018 dans la revue *Nature Communications* (**Annexe 2**). Ils ont contribué à démontrer le haut pouvoir de résolution et la haute sensibilité de l'aérolysine, et ont été déterminant dans notre choix de proposer une stratégie de séquençage des protéines basée sur l'utilisation de ce nanopore.

II. Résultats & Discussion

2.1 Analyse d'une courte chaîne peptidique issue d'un échantillon monodisperse

Des homopeptides d'arginine d'une longueur de 9 acides aminés (R_9), issus d'un échantillon monodisperse de pureté supérieure à 98%, ont été ajoutés dans le compartiment *Cis* à une concentration finale de 1 μ M (**Fig. 42**).

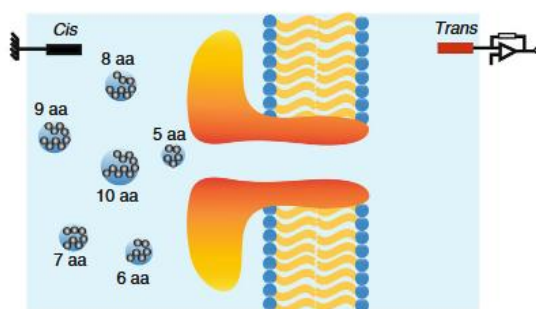


Figure 42 - Schéma illustratif du dispositif expérimental général. Un nanopore d'aérolysine sauvage est inséré dans une bicouche de lipides plane. Des homopeptides d'une longueur de 5 à 10 acides aminés sont analysés individuellement ou dans un mélange. Les homopeptides sont ajoutés dans le compartiment *cis*. Le transport des chaînes se fait sous l'action d'une tension électrique négative appliquée en *trans*. Les expériences sont réalisées en présence d'une solution ionique de KCl à une concentration de 4 M, ajustée et stabilisée à pH 7,5 par l'utilisation d'une solution tampon d'HEPES à une concentration de 5 mM, et à une température contrôlée de 20,0 +/- 0,5 °C.

L'addition de ces peptides entraîne l'apparition d'événements de blocage du courant (chutes ou impulsions du courant), sous l'action d'une tension électrique négative (**Fig. 43**). Aucun événement de blocage du courant lié à la présence de ces peptides n'a été observé sous l'action d'une tension électrique positive. Cette observation nous indique qu'il est dans un premier temps possible de détecter les homopeptides d'arginine. Comme les acides aminés d'arginine portent une charge positive à pH 7,5, les homopeptides d'arginine sont chargés positivement à pH 7,5. Il semble cohérent, selon l'action d'une force électrophorétique, de détecter ces peptides en appliquant une différence de potentiel négative. L'entrée (ou l'interaction molécule-pore) d'une chaîne d'homopeptides R_9 dans le nanopore bloque le courant à hauteur d'environ 70% (**Fig. 43**).

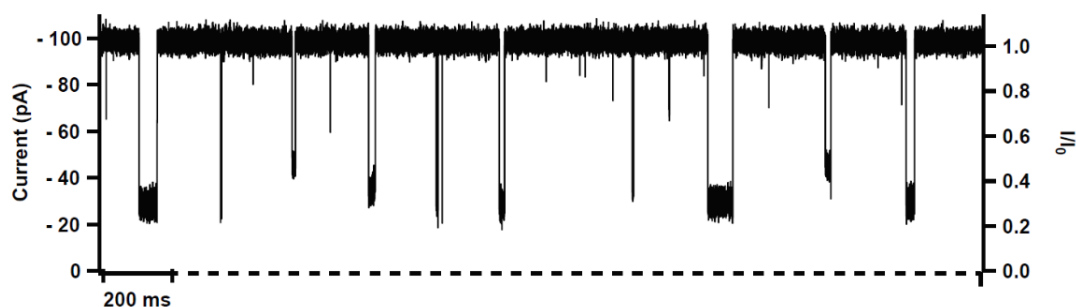


Figure 43 - Détection d'homopeptides d'arginine d'une longueur de 9 acides aminés avec l'aérolysine. Portion d'une trace de courant en fonction du temps obtenue dans le cas de l'analyse avec l'aérolysine d'une solution d'homopeptides d'arginine d'une longueur de 9 acides aminés, d'une pureté supérieure à 98% . Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de +50 mV appliquée en trans.

Cependant, d'autres événements d'amplitudes disparates sont observables sur la trace du courant (**Fig. 43**). La fréquence et l'amplitude de ces événements sont faibles en comparaison aux blocages majoritairement reproduits. La distribution des durées des blocages est très dispersée, allant d'une centaine de microsecondes à plusieurs dizaines de millisecondes pour les temps les plus longs (**Fig. 43**). Il apparaît qu'il existe un lien entre l'amplitude des blocages du courant et leurs durées, c'est-à-dire que plus les événements ont une amplitude de blocage courte, plus la durée du blocage est également courte, et *vice versa* (**Fig. 43**). Cela indique que ces blocages du courant correspondent à l'entrée de molécules, plutôt qu'à de simples collisions non caractéristiques.

En considérant que la solution d'homopeptides R₉ analysée correspond à un échantillon monodisperse, nous nous attendions à la présence de blocages du courant caractéristiques de ce peptide en termes d'amplitude, et reproduits à chaque fois que le peptide est transporté dans le pore. L'observation d'une disparité des amplitudes de blocages du courant était ainsi contre-intuitive.

Pour comprendre ce résultat, nous avons réalisé une première analyse statistique classique présentée en **Figure 44**. Cette analyse consiste à transformer l'ensemble des points du courant enregistrés en « Histogramme tout point » selon la méthode présentée dans le **chapitre n°2**. Rappelons ici que la fréquence d'échantillonnage est de 250 kHz, ce qui signifie qu'un point de courant est enregistré toute les 4 μ s. Cette représentation montre l'apparition de deux populations. Une première population est localisée à un niveau de courant moyen d'environ

100 pA (**Fig. 44**). Cette population correspond au niveau du courant du pore vide nommé « I_0 ».

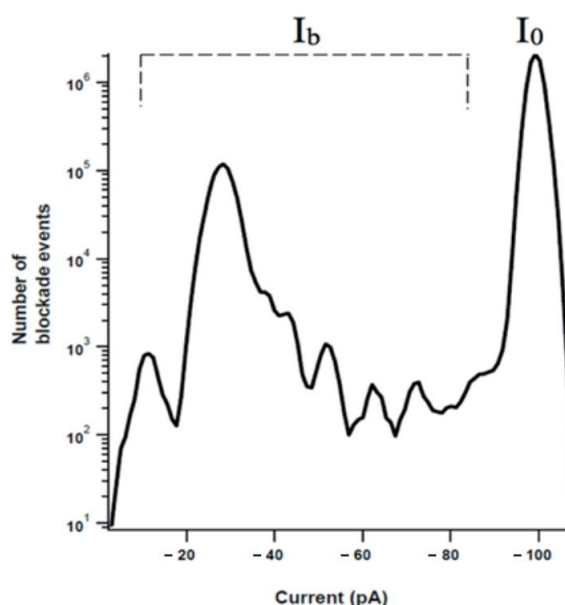


Figure 44 - Analyse statistique de la trace de courant électrique obtenue pour l'analyse d'homopeptides d'une longueur de 9 acides aminés avec l'aérolysine. Histogramme du nombre de points en fonction du courant obtenu lors de l'analyse des homopeptides de 9 acides aminés d'arginine avec l'aérolysine. Deux populations apparaissent : la première correspond au niveau de courant du pore vide (I_0 , à droite sur l'histogramme), et la seconde correspond aux niveaux du courant des blocages apparus lors de l'analyse des homopeptides d'arginine de 9 acides aminés avec l'aérolysine (I_b , à gauche sur l'histogramme).

Une seconde population plus large est localisée plus à gauche de l'histogramme (**Fig. 44**). Cette population correspond aux différents niveaux du courant « I_b » mesurés lorsque le pore est occupé par la présence d'homopeptides d'arginine. Cependant, bien que plusieurs pics soient discernables au sein de cette large population, on peut remarquer que cette représentation ne permet pas de mettre en exergue les différents niveaux des blocages du courant observés en temps réel sur la trace de courant (**Fig. 44**).

La seconde analyse statistique, moins classique mais plus fine, consiste à représenter un histogramme de la distribution de l'ensemble des niveaux de blocage du courant (I_b) normalisés par le niveau de courant du pore vide (I_0) (courant relatif résiduel I_b/I_0) (**Fig. 45**). Cette représentation rend bien compte de la distribution de l'ensemble des niveaux de blocage du courant observés sur la trace enregistrée, et permet de regrouper les niveaux préférentiels des blocages du courant. Dans le cas de l'analyse de R_9 , on voit très clairement sur l'histogramme un seul niveau préférentiel bien défini à une valeur de $I_b/I_0 \approx 0,289 \pm 0,002$ et une distribution

dispersée du reste des niveaux de blocage du courant (**Fig. 45, gauche**). La représentation en nuage de points de la distribution des durées des blocages du courant en fonction de I_b/I_0 a également été réalisée. Un nuage de points se dessinant à la même valeur de I_b/I_0 confirme l'existence d'une population pour laquelle les durées des blocages du courant s'étalent entre environ 1 et 100 ms. Il est naturel d'associer ce niveau préférentiel de I_b/I_0 aux interactions des peptides R₉ avec le pore d'aérolysine. Cependant, on ne peut pas en dire autant du reste des niveaux de blocage du courant qui ne sont pas représentés sur cet histogramme. Ces blocages du courant apparus en temps réel lors de l'analyse qualitative semblent donc être minoritaires.

Traditionnellement, dans les expériences de translocation de macromolécules, ces blocages de courant minoritaires peuvent être associés à des événements de collision entre les molécules et l'entrée du pore [309]. Cependant, il est important de souligner que l'analyse présentée ci-dessus a été réalisée à partir d'une trace de courant ne présentant pas un nombre d'événement de blocages du courant suffisamment important pour nous permettre de caractériser les populations minoritaires observées. La démarche que nous avons mise en place pour expliquer l'origine de ces événements dispersés et minoritaires a consisté à s'intéresser à l'évolution de la distribution de ces événements, en augmentant d'une manière très significative le nombre de blocages du courant (> 2000 événements). Pour y parvenir, nous avons augmenté la durée d'enregistrement de façon à obtenir une statistique plus importante.

L'analyse d'un plus grand nombre d'événements de blocage du courant montre l'apparition de plusieurs niveaux préférentiels bien définis dans l'histogramme de I_b/I_0 (**Fig. 45, droite**). Les niveaux de courant préférentiels sont clairement séparés les uns des autres. On peut y voir un pic principal, de position identique à celui apparaissant sur le premier histogramme, à une valeur de $I_b/I_0 \approx 0,292 \pm 0,002$. La variabilité de la valeur du courant relatif mesuré peut résulter des petites variations liées aux conditions expérimentales, comme par exemple l'évaporation de l'électrolyte au cours de l'expérience. Ce pic a été attribué au peptide de 9 acides aminés. On peut y voir également quatre à cinq pics minoritaires additionnels, de proportions plus faibles, bien séparés les uns des autres, et positionnés à des valeurs de I_b/I_0 bien distinctes, indiqués par des flèches sur la Figure 45.

Il a été ainsi mis en évidence que les blocages du courant minoritaires ne sont en réalité pas négligeables, c'est-à-dire qu'il ne s'agit pas d'artefacts du signal, mais bien de blocages du courant engendrés par l'entrée et/ou l'interaction d'homopeptides d'arginine avec le pore d'aérolysine.

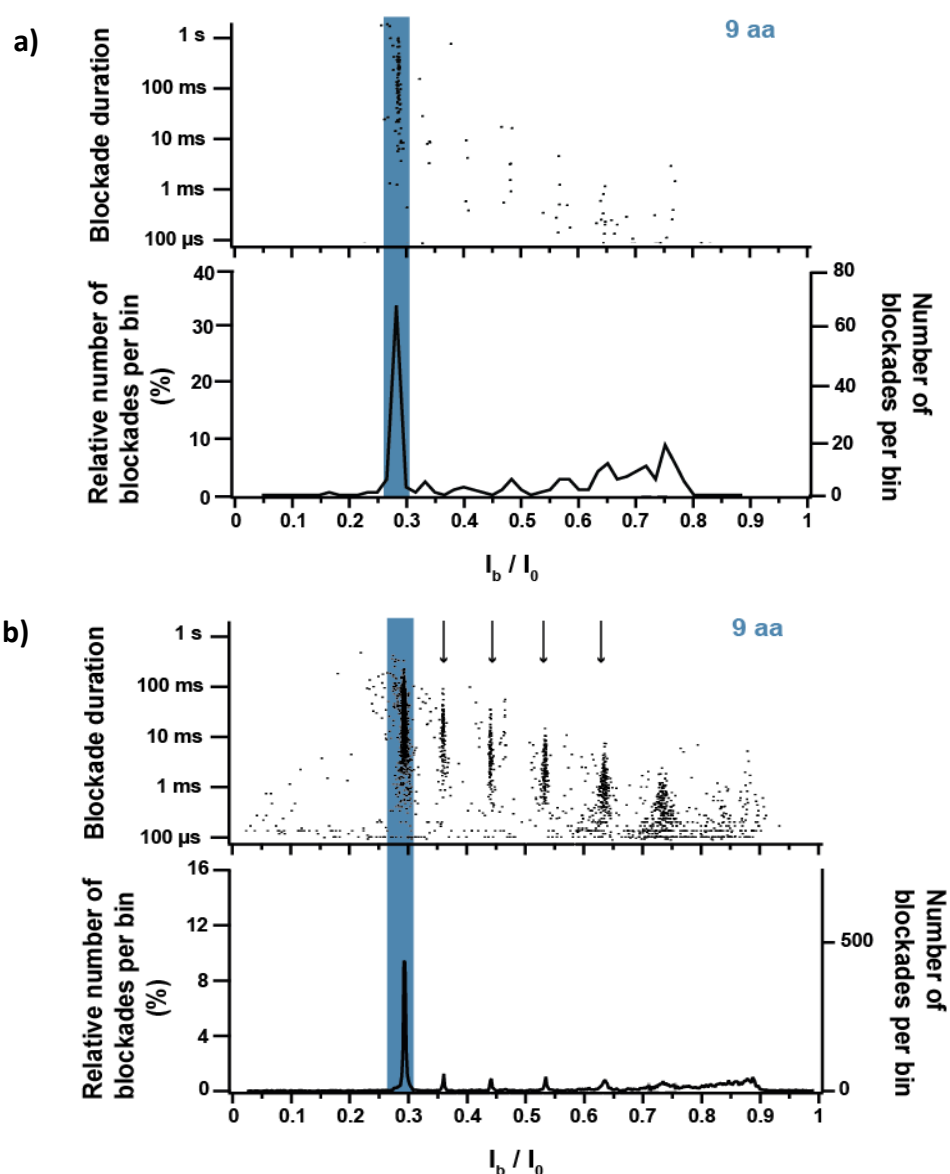


Figure 45 - Analyse d'homopeptides d'une longueur de 9 acides avec le nanopore d'aérolysine. (a) Histogramme de I_b/I_0 et distribution (nuage de points) des durées de blocage du courant en fonction I_b/I_0 dans le cas l'analyse d'un faible (< 2000) nombre d'événements de blocage obtenus pour la detection d'homopeptides R_9 avec l'aérolysine. Les deux représentations montrent l'apparition d'une population majoritaire à une valeur de I_b/I_0 d'environ 0,3. (b) Histogramme de I_b/I_0 et distribution (nuage de points) des durées de blocage du courant en fonction I_b/I_0 dans le cas l'analyse d'un grand (> 2000) nombre d'événements de blocage obtenus pour la detection d'homopeptides R_9 avec l'aérolysine. Les deux représentations montrent l'apparition d'une population majoritaire à une valeur de I_b/I_0 d'environ 0,3 et de quatre populations minoritaires à des valeurs de I_b/I_0 comprises entre 0,3 et 1 (flèches). Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de +50 mV appliquée en trans.

Naturellement, la suite des expériences a consisté à essayer d'identifier ces différentes populations secondaires.

Les échantillons peptidiques analysés par nanopore présentent une pureté supérieure à 98%. Il existe donc seulement 2% d'impuretés dans ces échantillons. Une des hypothèses possibles serait que ces populations correspondent à des peptides de plus courtes longueurs présents à l'état d'impuretés dans l'échantillon peptidique et parvenant à être détectés par l'aérolysine. Cette hypothèse indiquerait que le nanopore d'aérolysine nous permet de détecter des quantités très faibles présentes à l'état de traces dans un échantillon. Pour vérifier cela, nous avons réalisé un calcul théorique de la variation de la résistance électrique du pore en présence d'un objet de taille ou de volume fini dans le pore afin de déterminer les valeurs des niveaux de courant attendues (cf. chapitre n°2). Les valeurs obtenues par calcul coïncident avec les niveaux de courant mesurés expérimentalement correspondant aux quatre populations localisées sur l'histogramme. L'estimation théorique prédit que les quatre populations minoritaires correspondraient à des homopeptides d'arginine de 5, 6, 7, et 8 acides aminés.

2.2 Détection et identification d'impuretés peptidiques

Pour vérifier cette hypothèse expérimentalement, des échantillons d'homopeptides d'arginine de haute pureté (> 98%) d'une longueur de 5, 6, 7, 8 et 10 acides aminés ont été commandés et analysés individuellement. Pour l'analyse de chaque homopeptide d'une longueur donnée avec l'aérolysine, des événements de blocage du courant de niveaux et de durées variables ont été observés, de façon comparable à ce qui avait été constaté pour le peptide de 9 arginines (R₉). De nouveau, comme pour l'analyse du peptide R₉, les histogrammes I_b/I_0 montrent pour chaque peptide d'une longueur donnée, l'apparition d'une raie majoritaire à un niveau de courant relatif I_b/I_0 préférentiel (**Fig. 46**). On peut lire sur la figure les valeurs nominatives de courant relatif I_b/I_0 correspondant aux raies majoritaires des homopeptides d'une longueur de 5 (0,234 +/- 0,001), 6 (0,286 +/- 0,002), 7 (0,353 +/- 0,002), 8 (0,435 +/- 0,002), 9 (0,530 +/- 0,002) et 10 (0,631 +/- 0,004) acides aminés d'arginine (**Fig. 46**).

De façon remarquable, les valeurs nominatives de courant relatif I_b/I_0 correspondant aux raies minoritaires observées dans le cas de l'analyse de l'échantillon d'homopeptides de 9 acides aminés, sont identiques aux valeurs nominatives de courant relatif I_b/I_0 correspondant aux raies majoritaires des homopeptides de 5, 6, 7 et 8 acides aminés. On peut donc attribuer les raies minoritaires observées dans le cas de l'analyse de l'échantillon d'homopeptides de 9 acides aminés pur à plus de 98%, aux fragments d'homopeptides d'une longueur de 5, 6, 7, et 8 acides

aminés, présents en quantité très faible dans cet échantillon. Le peptide de 10 acides aminés a également entraîné des blocages du courant dont l'amplitude est spécifique et distinguable des autres peptides de plus courtes longueurs. Dans les conditions expérimentales utilisées, les homopeptides d'arginine sont détectés uniquement lorsqu'ils sont ajoutés du côté *Cis* du nanopore d'aérollysine (**Fig. 47**).

Ces résultats montrent que la très haute sensibilité du dispositif permet d'identifier des quantités très faibles de peptides. Ces peptides présents à l'état de trace n'ont pas été caractérisés par l'analyse en spectrométrie de masse classique et HPLC réalisée par le fournisseur. Ces résultats nous permettent aussi de quantifier ces peptides malgré leur faible proportion de l'ordre de 1% du nombre total de molécules détectées, soit quelques dizaines de molécules dans l'échantillon.

Un autre résultat est que la distinction de ces peptides signifierait qu'il est possible de discriminer ces peptides avec une résolution de séparation d'un acide aminé, ce qui n'avait jamais été montré en comparaison à des épingles d'ADN [178], à des polymères synthétiques [26], ou encore à des oligonucléotides [288].

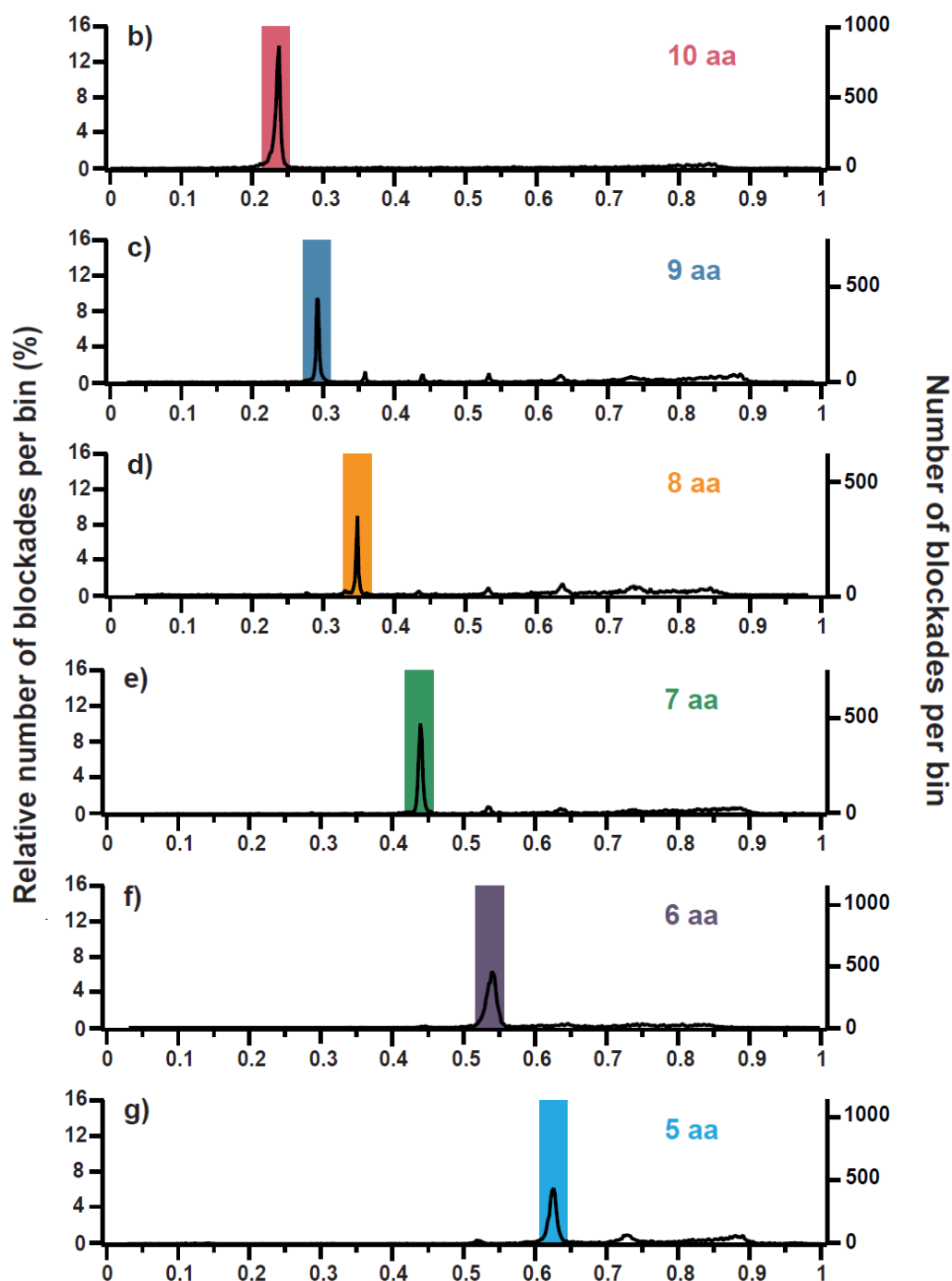


Figure 46 - Analyse individuelle d'homopeptides d'une longueur de 5 à 10 acides aminés avec le nanopore d'aérolysine. Histogrammes de I_b/I_0 obtenus dans le cas de l'analyse individuelle d'homopeptides d'arginine d'une longueur de 10 (b), 9 (c), 8 (d), 7 (e), 6 (f), et 5 (g) acides aminés, sous l'action d'une tension électrique de -50 mV appliquée en trans. L'analyse de chacun des peptides d'une longueur donnée est à l'origine de l'apparition d'un pic majoritaire à une valeur de I_b/I_0 préférentielle et spécifique sur les histogrammes. Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de +50 mV appliquée en trans.

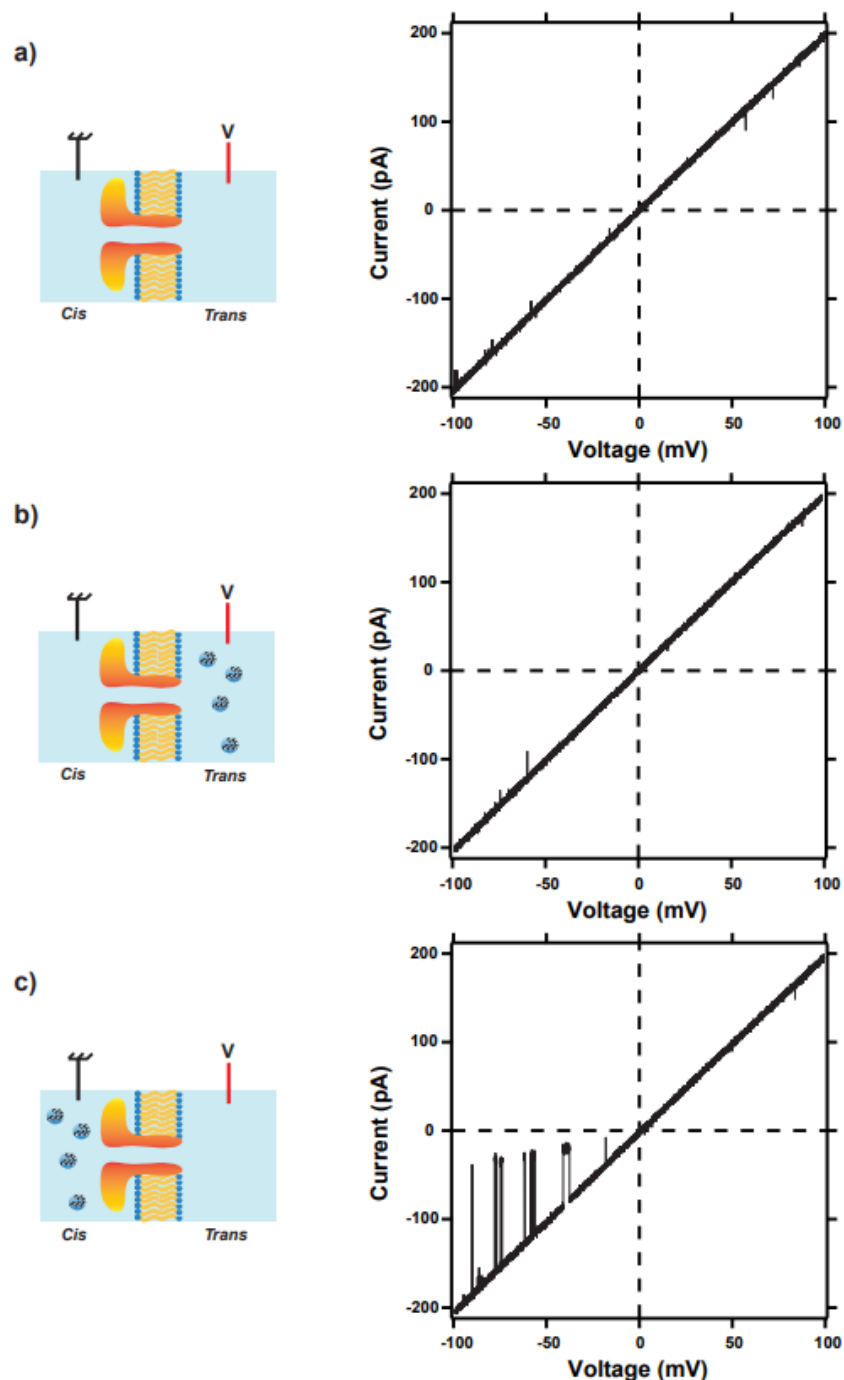


Figure 47 - Détection d'homopeptides d'arginine avec le nanopore d'aérollysine en fonction du côté du nanopore dans lequel sont ajoutés les peptides, et en fonction du signe de la tension électrique appliquée. Courant électrique, en fonction de la tension électrique appliquée en trans, mesuré à travers le nanopore d'aérollysine : (a) en absence de peptides ; ou en présence d'homopeptides d'arginine d'une longueur de 10 acides aminés : ajoutés dans le compartiment trans à une concentration finale de $6 \mu\text{M}$ (b), ou ajoutés dans le compartiment cis à une concentration finale de $1,2 \mu\text{M}$. Les mesures ont été réalisées dans une solution de KCl 4M, 5mM HEPES, pH = 7,5, à 20°C ; et la durée de chaque courbe courant-tension est de 4 secondes.

2.3 Détection et discrimination en taille des peptides avec une résolution d'un acide aminé

Pour confirmer ces résultats, ayant à notre disposition les différentes longueurs de peptides, un mélange de ces différents peptides d'une longueur de 5, 6, 7, 8, 9, et 10 acides aminés a été analysé avec l'aérollysine.

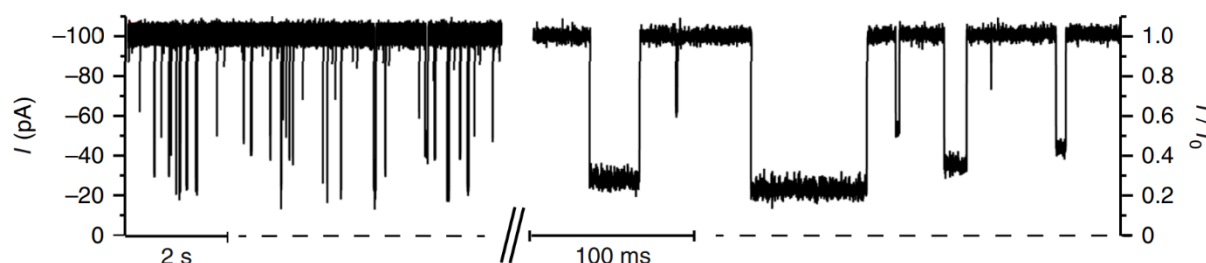


Figure 48 - Détection d'un mélange d'homopeptides d'une longueur de 5 à 10 acides aminés avec l'aérollysine. Portion d'une trace électrique de courant en fonction du temps obtenue lors de l'analyse d'un mélange d'homopeptides d'arginine d'une longueur de 5, 6, 7, 8, 9, et 10 acides aminés avec le nanopore d'aérollysine, sous l'action d'une tension électrique de -50 mV appliquée en trans. Des blocages du courant d'amplitudes multiples sont visualisées sur la trace de courant et ont été attribués aux peptides de différentes longueurs. Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C.

De façon comparable à ce qui avait été observé précédemment, l'addition d'un mélange de peptides a entraîné des blocages du courant d'amplitudes et de durées variables (**Fig. 48**). L'histogramme de I_b/I_0 montre l'apparition de six populations à des niveaux de courant relatif préférentiels distincts (**Fig. 49**). Chacune de ces populations est superposable, à la fois en termes des valeurs du courant relatif I_b/I_0 et des durées de blocages du courant, à une longueur donnée de peptide. Cela confirme que le nanopore d'aérollysine permet dans ces conditions expérimentales, de détecter des courts homopeptides d'arginine d'une longueur comprise entre 5 et 10 acides aminés, mais aussi de distinguer clairement deux homopeptides qui diffèrent en longueur d'un même et unique acide aminé, ce qui correspond ici à une masse molaire d'environ 174 g/mol. L'aérollysine permet donc la discrimination en taille de ces homopeptides avec une résolution d'un acide aminé.

L'analyse du mélange d'homopeptides d'arginine de différentes longueurs a été reproduite au moins quatre fois. Ces analyses montrent que les résultats expérimentaux obtenus sont reproductibles (**Fig. 50**).

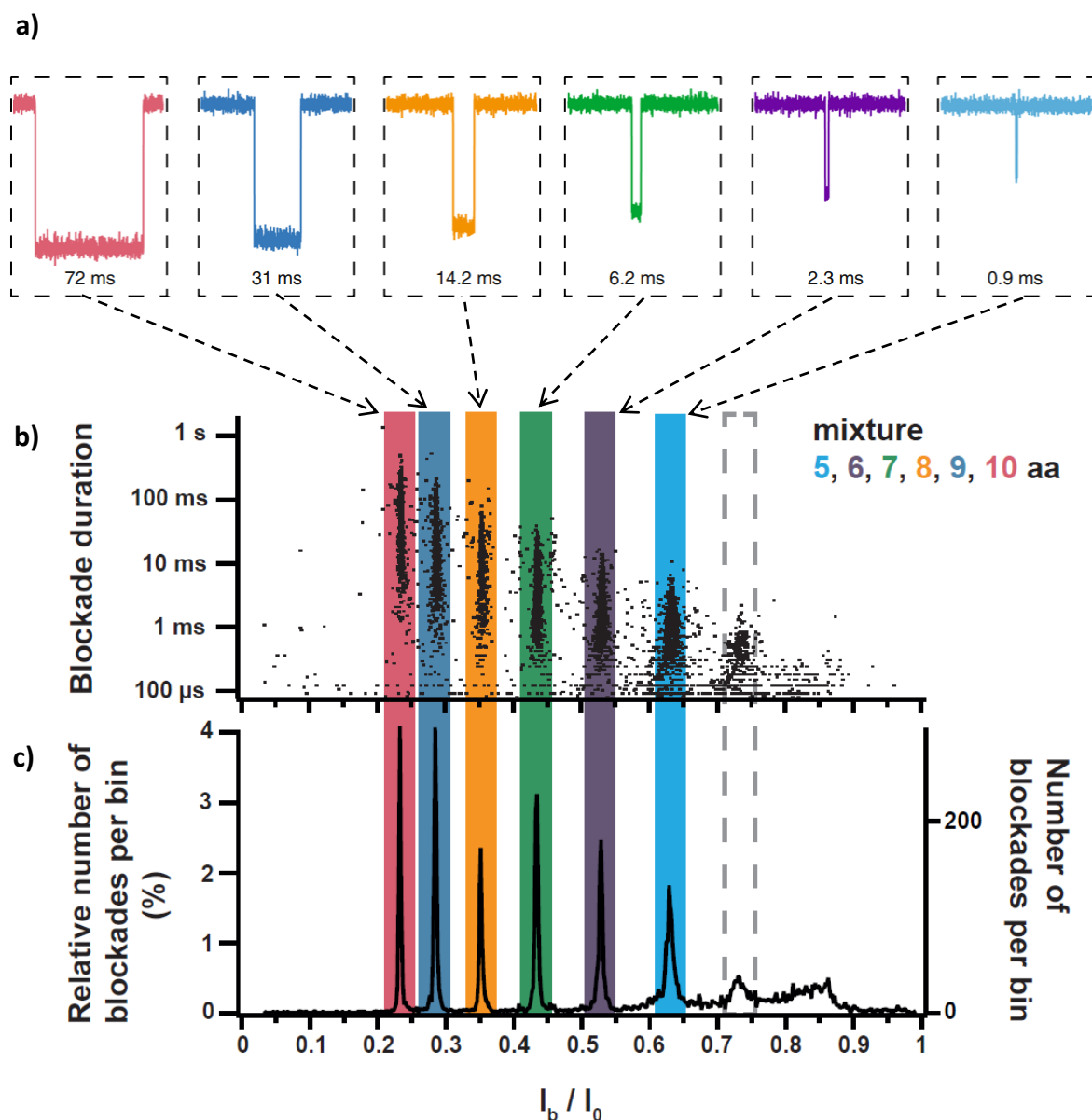


Figure 49 - Analyse d'un mélange d'homopeptides d'arginine d'une longueur de 5 à 10 acides aminés avec l'aérollysine. (a) Illustration d'événements de blocage du courant typiques engendrés par des homopeptides d'arginine de différentes longueurs analysés avec l'aérollysine. (b) Histogramme de I_b/I_0 et (c) distribution des durées des blocages du courant (nuage de points) en fonction de I_b/I_0 , dans le cas de l'analyse d'un mélange équimolaire d'homopeptides d'arginine d'une longueur de 5, 6, 7, 8, 9, et 10 acides aminés avec le nanopore d'aérollysine, sous l'action d'une tension électrique de -50 mV appliquée en trans. L'histogramme montre l'apparition de six pics ayant été attribués aux six populations peptidiques présentes en solution. Le nuage de points montre six populations de points se dessinant aux mêmes valeurs de I_b/I_0 que les populations identifiées sur l'histogramme. Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C.

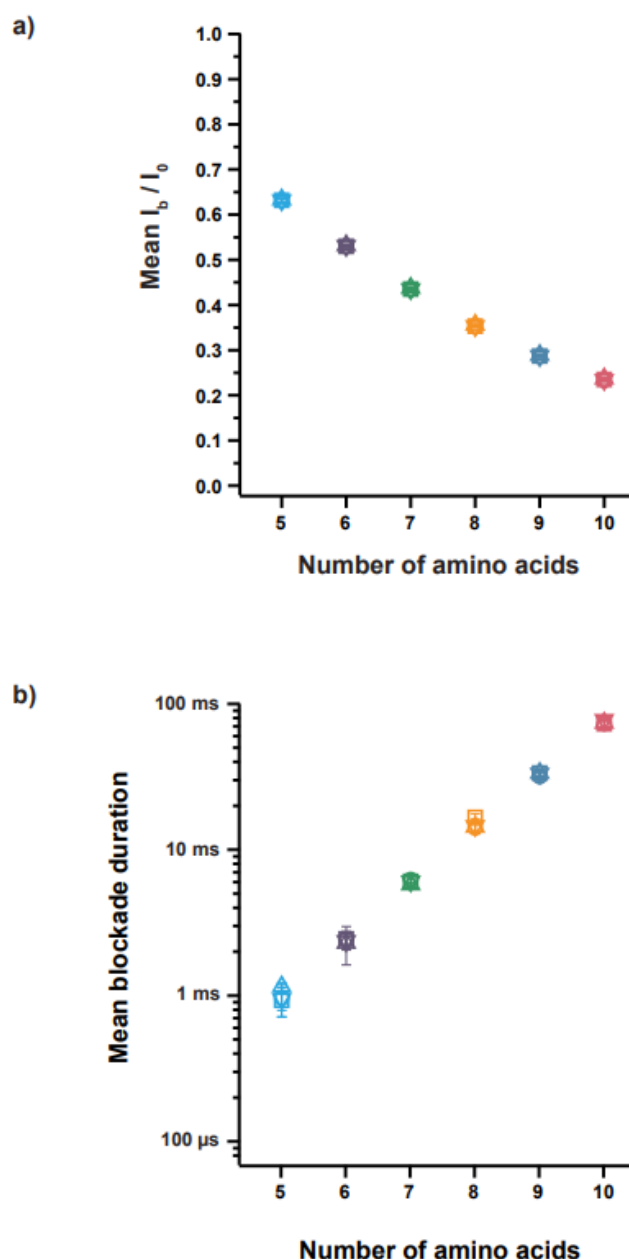


Figure 50 - Reproductibilité des résultats expérimentaux. Superposition de résultats obtenus à partir de quatre expériences indépendantes (symboles $\circ$, $\square$, $\triangle$, ∇) dans le cas de l'analyse d'un mélange équimolaire d'homopeptides d'arginine d'une longueur de 5, 6, 7, 8, 9, et 10 acides aminés avec le nanopore d'aérolysine : (a) Courant relatif résiduel moyen I_b/I_0 , et, (b) durée moyenne de blocage du courant, en fonction du nombre d'acides aminés. Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de +50 mV appliquée en trans.

2.4 Effet de la tension électrique sur la discrimination en taille des peptides

L'effet de la tension électrique a été étudié dans le cas de l'analyse d'un mélange des différentes longueurs de peptides (5 à 10 acides aminés) avec le nanopore d'aérolysine, pour des tensions électriques négatives comprises entre -25 et -80 mV (**Fig. 51 et 52**).

Dans un premier temps, la figure 51 montre, qu'indépendamment de la tension électrique appliquée, la durée moyenne du blocage augmente avec la longueur du peptide de jusqu'à deux ordres de grandeur entre le peptide d'une longueur de 5 acides aminés et celui de 10 acides aminés. Pour chaque peptide d'une longueur donnée, la durée moyenne de blocage du courant dépend de manière non monotone de la tension électrique, augmentant d'abord avec l'amplitude de la tension, atteignant un maximum, puis diminuant (**Fig. 51, b**). La tension électrique correspondant à la durée moyenne de blocage la plus grande, ne dépend que faiblement de la longueur du peptide, et est comprise entre -35 et -40 mV (**Fig. 51, b**). Il a été noté que la durée moyenne de blocage du courant correspondant à chaque peptide d'une longueur donnée est identique, qu'elle soit mesurée lors de d'analyse d'un mélange des peptides, ou lors de leur analyse individuelle.

Dans un second temps, la figure 51 montre que le courant relatif moyen I_b/I_0 diminue légèrement avec l'amplitude de la tension électrique pour les peptides les plus longs (10, 9, et 8 acide aminés), est dépendant de la tension pour les peptides de longueur intermédiaire (7 et 6 acides aminés), et augmente légèrement avec l'amplitude de la tension électrique pour les peptides les plus courts (5 acides aminés) (**Fig. 51, a**).

Les histogrammes de I_b/I_0 (**Fig. 52**) obtenus montrent d'une part, que la discrimination en taille des peptides est préservée dans toute la gamme de tensions électriques explorée, même à une tension électrique aussi basse que -25 mV, et d'autre part, que les différentes populations présentes sur l'histogramme de I_b/I_0 deviennent de plus en plus séparées à mesure que la valeur de la tension électrique appliquée augmente. Il est possible d'observer que la tension électrique pour laquelle les durées de blocage sont les plus grandes (environ -40 mV) ne correspond pas à celle où la discrimination des peptides est la meilleure (**Fig. 52**). De plus, à partir de -50 mV, une septième population qui n'est pas observée à des plus basses tensions électriques, se dessine sur l'histogramme à une valeur moyenne de $I_b/I_0 \approx 0,73$ (**Fig. 52**). A plus basses tensions, cette population n'est pas observée car son signal est probablement masqué par la présence d'un nombre significatif de fluctuations du courant du pore correspondant au bruit du pore ouvert qui sont représentées à des valeurs de I_b/I_0 localisées entre 0,7 et 0,8. L'augmentation de la

valeur de la tension électrique permet d'augmenter le ratio signal/bruit. Cela est à l'origine d'un décalage des valeurs de I_b/I_0 des peptides de 5 à 10 acides aminés, qui permet l'apparition de la septième population.

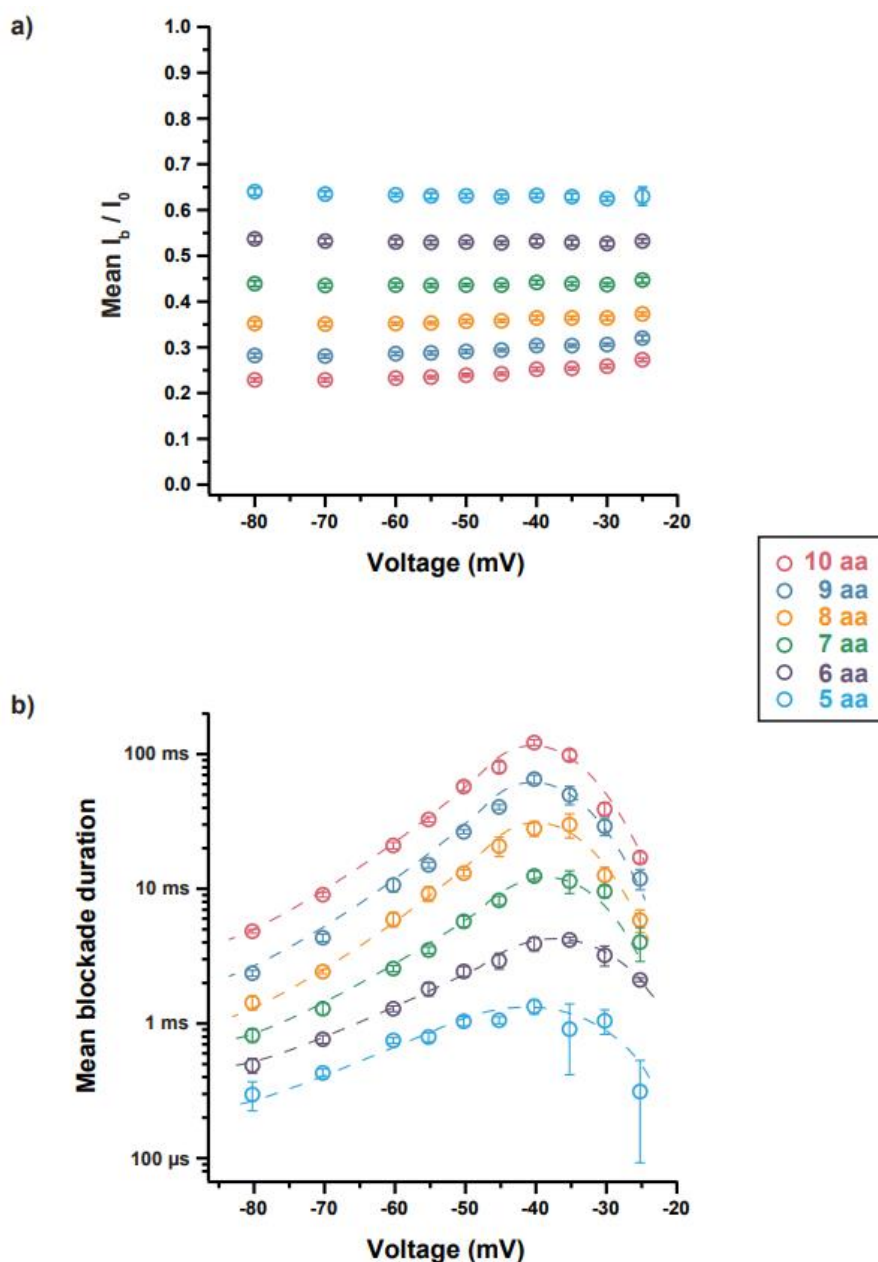


Figure 51 - Effet de la tension électrique sur l'interaction du nanopore d'aérolysine avec un mélange d'homopeptides d'arginine de différentes longueurs. (a) Courant relatif moyen I_b/I_0 en fonction de la tension électrique, et, (b) durée moyenne de blocage du courant en fonction de la tension électrique appliquée en trans ; mesurés pour chaque peptide d'une longueur donnée présente dans le mélange analysé (10 (rouge), 9 (bleu foncé), 8 (jaune), 7 (vert), 6 (violet), et 5 (bleu clair) acides aminés). Les lignes en pointillées servent de guide pour l'œil. Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C.

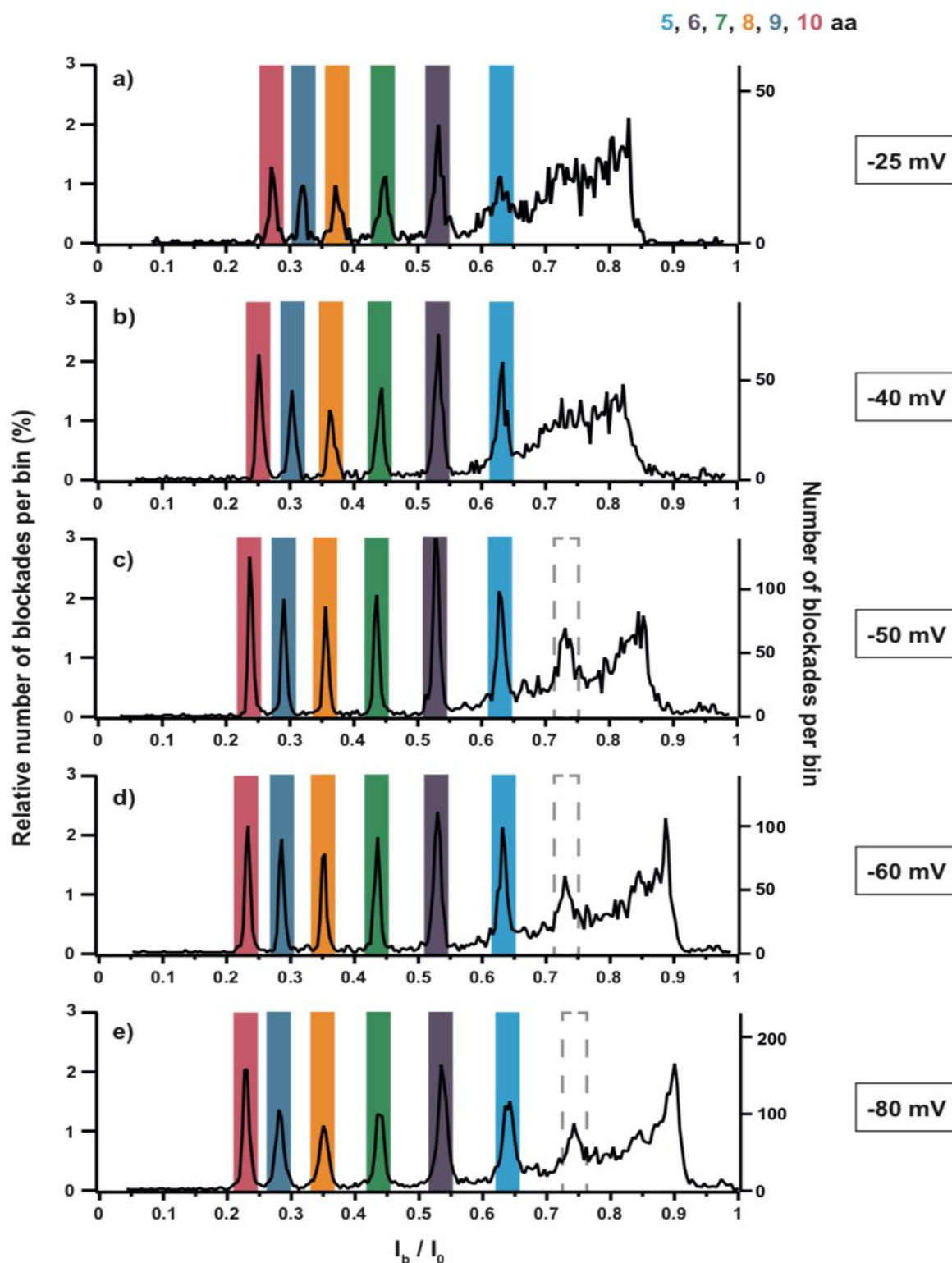


Figure 52 - Effet du voltage sur la discrimination en taille d'homopeptides d'arginine de différentes longueurs. Histogramme de I_b/I_0 dans le cas de l'analyse d'un mélange équimolaire d'homopeptides d'arginine d'une longueur de 5, 6, 7, 8, 9, et 10 acides aminés analysés avec le nanopore d'aérolysine, sous l'action d'une tension appliquée en trans de : (a) -25 mV, (b) -40 mV, (c) -50 mV, (d) -60 mV, et (e) -80 mV.

2.5 Extension de la limite de détection à des peptides d'une longueur de 4 acides aminés

Sur l'histogramme issu de l'analyse du mélange d'homopeptides d'arginine, une population additionnelle est apparue de façon plus nette à un niveau de courant plus faible que celui du peptide d'une longueur de 5 acides aminés ($I_b/I_0 \approx 0,73$).

Cette population additionnelle est également visible sur tous les histogrammes de l'ensemble des homopeptides analysés individuellement. Cette population correspondrait à un homopeptide de 4 acides aminés. Néanmoins, nous n'avons pas à notre disposition le peptide de cette longueur pour pouvoir le vérifier.

Pour démontrer les résultats obtenus d'une autre façon, une expérience de digestion enzymatique d'un homopeptide de 9 arginines (R_9) a été réalisée. La trypsine a été utilisée car cette enzyme est capable de cliver de façon préférentielle les liaisons dans lesquelles une arginine ou une lysine engage sa fonction carboxylique. L'idée est que l'enzyme clive les peptides de 9 acides aminés en peptides plus courts, ce qui se traduirait par une augmentation du nombre de blocages du courant provoqués par ces populations.

L'addition d'homopeptides R_9 a entraîné, comme lors des premières expériences, l'apparition de blocages du courant d'amplitudes et de durées variables. Cela montre la reproductibilité de la détection de ces peptides. Après l'ajout de trypsine, au fur et à mesure du temps de la digestion enzymatique, une augmentation du nombre de blocages du courant provoqués par les peptides les plus courts est observée au détriment d'une diminution du nombre de blocages du courant provoqués par les peptides d'une longueur de 9 arginines. Quantitativement, la même observation est réalisée sur l'histogramme de I_b/I_0 (**Fig. 53**) qui montre au fur et à mesure de la digestion enzymatique une diminution de l'intensité de la raie correspondant au peptide R_9 , au profit de l'enrichissement des raies correspondant aux peptides plus courts. Après trois heures et trente minutes de digestion enzymatique, les blocages du courant engendrés par des peptides d'une longueur de 9, 8, et 7 acides aminés ont quasiment disparu, alors qu'il existe toujours des blocages du courant provoqués par les peptides les plus courts de 6, 5, et 4 acides aminés, ce qui se traduit par une disparition de la population des homopeptides de 9 acides aminés sur l'histogramme. Ces résultats suggèrent que les homopeptides R_9 sont dégradés par la trypsine en peptides de plus courtes longueurs.

Nous montrons donc que les populations minoritaires observées lorsqu'un peptide d'une longueur donnée est analysé correspondent effectivement à des peptides de plus courtes

longueurs, et diffèrent d'un acide aminé individuel entre deux raies successives. Il est possible de conclure avec ces résultats que l'aérolysine permet une discrimination en taille de courts homopeptides d'arginine avec une résolution d'un acide aminé.

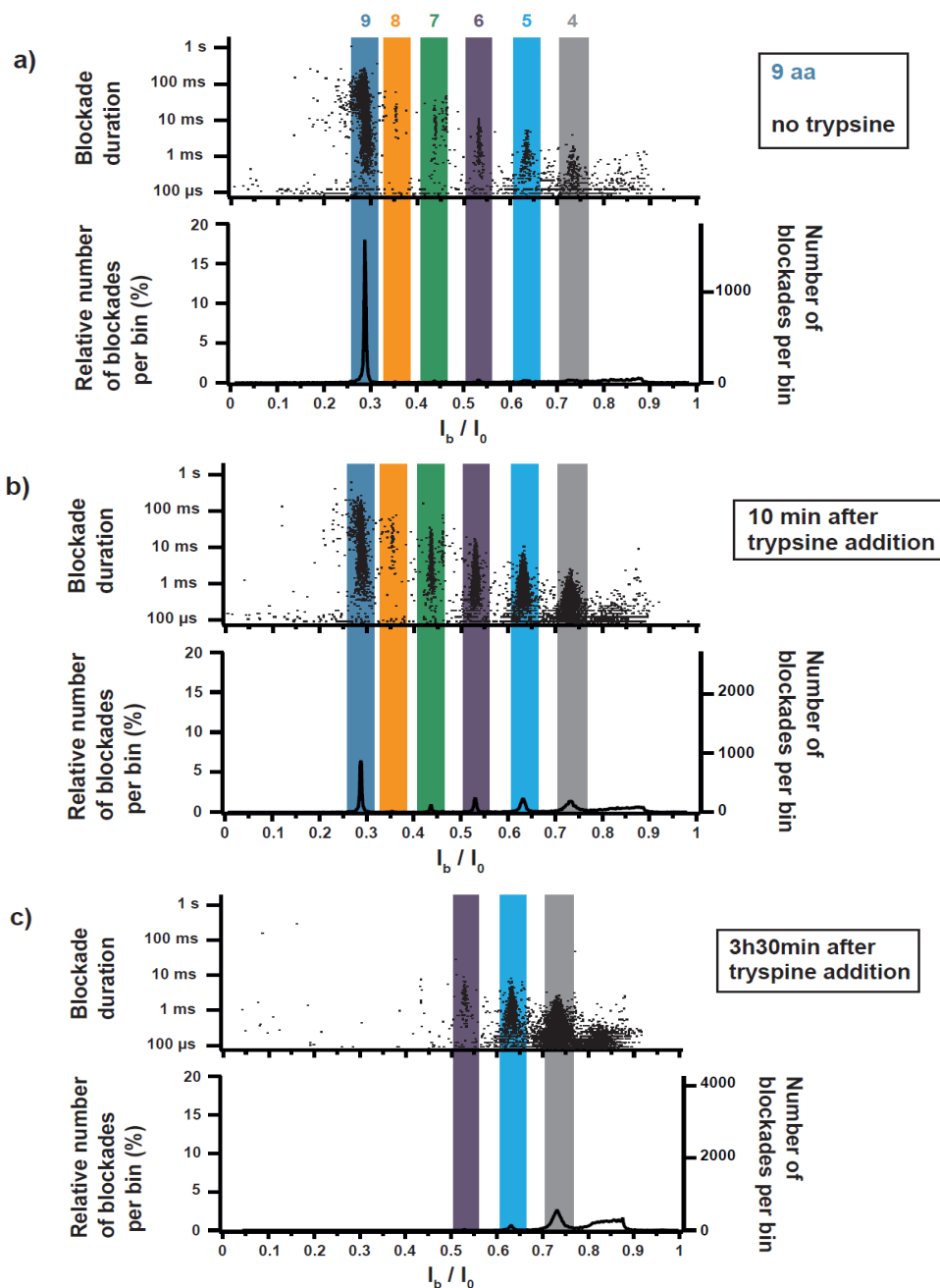


Figure 53 - Expérience de suivi en temps réel de la dégradation enzymatique par la trypsin d'un échantillon d'homopeptides d'arginine avec le nanopore d'aérolysine. Distribution des durées de blocage du courant en fonction de I_b/I_0 (haut de chaque sous-figure) et histogramme de I_b/I_0 (bas de chaque sous-figure) dans le cas de l'analyse d'homopeptides d'arginine d'une longueur de 9 acides aminés en présence ou non de trypsin. (a) En absence de trypsin, l'analyse d'un échantillon d'homopeptides d'une longueur de 9 acides aminés montre la présence d'une population majoritaire correspondant à ce même peptide, et des populations

minoritaires identifiées comme correspondantes aux peptides de 4 à 8 acides aminés. (b) Après 10 minutes de digestion enzymatique, une augmentation du nombre de points est constatée pour les peptides les plus courts : les peptides de 9 acides aminés sont dégradés en peptide de 6 à 4 acides aminés (c) Après 3 heures et 30 minutes de digestion enzymatique, les nuages de points se dessinant pour les peptides de 9, 8 et 7 acides aminés ont disparu contre un enrichissement des populations minoritaires. Cela indique que les peptides d'une longueur de 9, 8 et 7 acides aminés sont totalement dégradés en peptides d'une longueur de 6, 5, et 4 acides aminés. La trypsine est ajoutée dans le compartiment cis à une concentration finale de 25 nM.

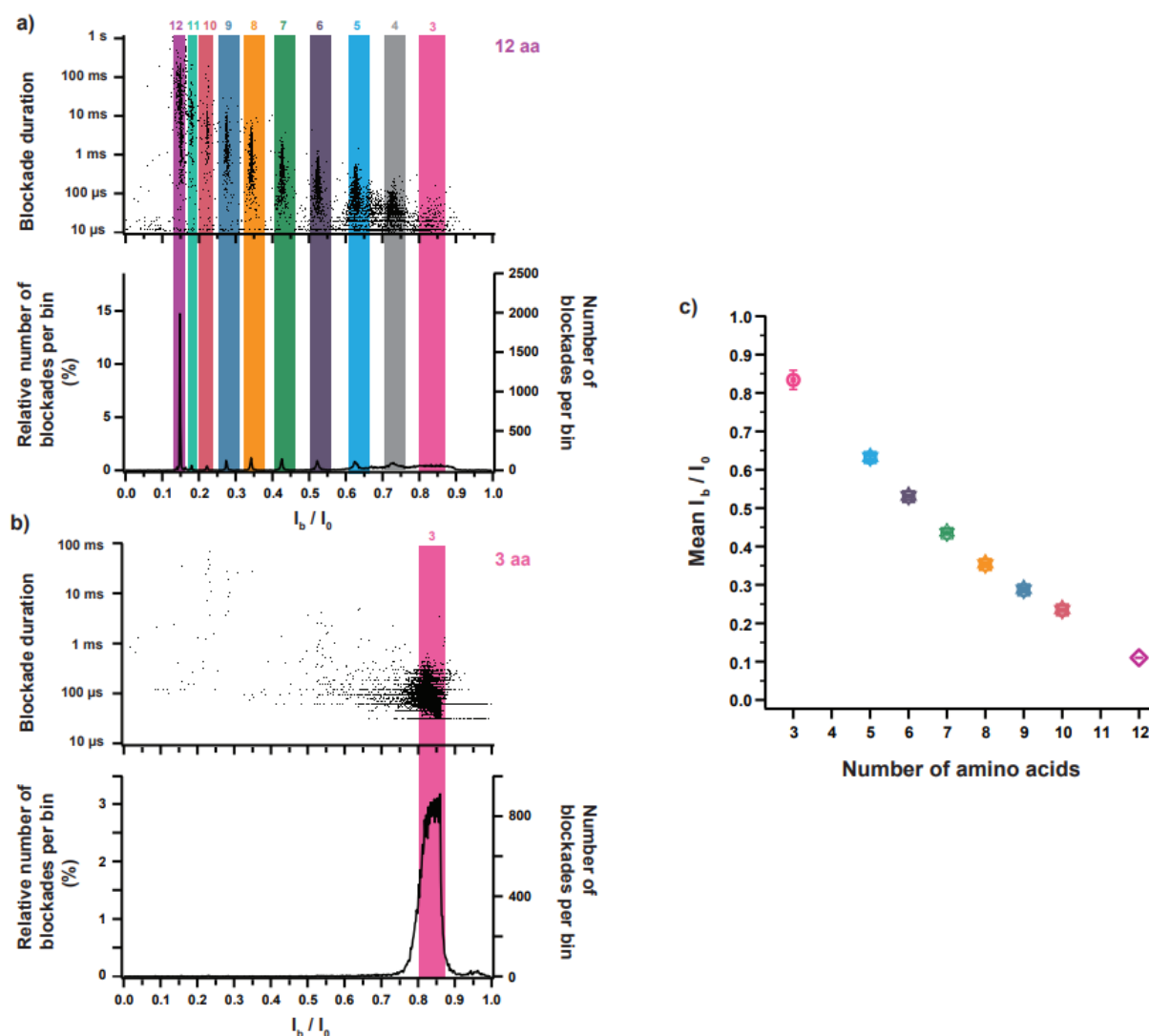


Figure 54 - Détection et identification d'homopeptides d'arginine d'une longueur de 3 et 12 acides aminés. Distribution des durées de blocage du courant en fonction de I_b/I_0 (haut de chaque sous-figure) et histogramme de I_b/I_0 (bas de chaque sous-figure) dans le cas de l'interaction du nanopore d'aérolysine avec : (a) une solution d'homopeptides d'arginine d'une longueur de 12 acides aminés (violet) ; ou (b) d'une solution d'homopeptides d'arginine d'une longueur de 3 acides aminés (rose). (c) Courant relatif moyen I_b/I_0 en fonction du nombre d'acides aminés dans le cas de l'analyse d'un mélange équimolaire d'homopeptides d'une longueur de 5, 6, 7, 8, 9, et 10 acides aminés ; et de l'analyse individuelle des homopeptides d'arginine d'une longueur de 3 et 12 acides aminés. Les mesures ont été réalisées dans une

solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de -50 mV, appliquée en trans.

Dans le contexte du processus de révision de cet article, nous avons également montré que la détection avec l'aérolysine peut être étendue à des homopeptides d'arginine de plus grandes longueurs (12 acides aminés), et de plus petites longueurs (3 acides aminés). Ces peptides ont été analysés individuellement avec le nanopore d'aérolysine (**Fig. 54**). Chaque histogramme présente une population majoritaire unique permettant d'identifier les peptides d'arginine d'une longueur de 12 et de 3 acides aminés. Neuf populations additionnelles sont présentes sur l'histogramme de l'analyse du peptide d'une longueur de 12 arginines, et ont été attribuées aux homopeptides d'une longueur 11, 10, 9, 8, 7, 6, 5, 4, et 3 acides aminés présentes dans l'échantillon comme impuretés peptidiques (**Fig. 54**).

2.6 Application à un autre type d'acide aminé

Après avoir montré que l'aérolysine permet de faire la discrimination de taille d'homopeptides d'arginine, nous étions curieux de savoir si ce pouvoir résolutif peut être obtenu avec des homopeptides constitués d'autres types d'acides aminés.

Pour répondre à cela, nous avons choisi d'utiliser des homopeptides de lysine (K_n) qui ont été mis en présence du nanopore d'aérolysine dans les mêmes conditions expérimentales que pour l'analyse des homopeptides d'arginine. Des homopeptides de lysine (146 g/mol) d'une longueur de 10 acides aminés (K_{10}) ont été ajoutés en solution à une concentration finale identique à celle des homopeptides d'arginine (1 μ M). À l'addition des homopeptides de lysine en solution, aucun événement de blocage du courant n'est apparu, même après avoir attendu un temps assez long pour laisser les molécules diffuser dans le compartiment. La première interprétation était donc que les peptides de lysine ne sont pas détectables car ils n'interagissent pas avec le pore d'aérolysine. Cependant, comme il est bien connu que les molécules de lysine s'adsorbent très fortement sur les surfaces comme celle de la cellule de mesure, on peut penser aussi que l'on a diminué ou même annulé artificiellement la concentration des homopeptides de lysine en solution. Une autre hypothèse est que le contenu peptidique des échantillons d'homopeptides de lysine et d'arginine qui nous ont été fournis n'est pas le même. En conséquence, la concentration finale des homopeptides K_{10} a été augmentée d'environ 200 fois (213 μ M), ce qui nous a permis de détecter les peptides.

L'histogramme de la distribution des niveaux de blocage du courant (**Fig. 55**) montre l'apparition d'une population majoritaire à un niveau de courant préférentiel et un temps

caractéristique que nous avons attribué aux homopeptides de lysine de 10 acides aminés ($I_b/I_0 \approx 0,37$). Les temps de ces blocages du courant sont distribués entre 100 μ s et 10 ms. De façon additionnelle à la raie majoritaire, trois raies minoritaires sont observées mais dont les positions sont beaucoup plus éloignées de la raie majoritaire. Ce résultat est intéressant mais ne valide pas la discrimination en taille avec une résolution d'un acide aminé des homopeptides de lysine.

N'ayant pas à notre disposition les différentes longueurs d'homopeptides de lysine pour analyser un mélange de différentes longueurs de ce peptide (5, 6, 7, 8, 9, et 10 acides aminés) pour valider la discrimination en taille avec une résolution d'un acide aminé, nous avons effectué une expérience de dégradation enzymatique d'homopeptides de lysine de 10 acides aminés (**Fig. 55**).

En absence de trypsine, les résultats obtenus précédemment concernant la détection des homopeptides de lysine ont été reproduits.

En présence de trypsine, au fur et à mesure de la digestion enzymatique, on observe une apparition de populations minoritaires qui n'étaient pas apparus précédemment, et de niveaux de I_b/I_0 plus proches de la population principale. On observe également une augmentation du nombre de blocages des populations minoritaires au détriment d'une diminution du nombre de blocage de la population majoritaire. Après trois heures et trente minutes de digestion enzymatique, la population correspondant au peptide de K_{10} a quasiment disparu, alors que les populations minoritaires sont présentes en proportions encore plus importantes. Ces résultats montrent que les peptides de K_{10} acides aminés ont été dégradés en des peptides de lysine de plus courtes longueurs.

Nous observons donc, de façon indirecte, que la discrimination de taille de courts homopeptides d'arginine de 5 à 10 acides aminés peut être étendue à celle de courts homopeptides de lysine. Cependant, des expériences d'analyse de chacune des longueurs de peptide dans des expériences indépendantes doivent être réalisées pour identifier de façon certaine les populations secondaires observées sur les histogrammes de l'homopeptide de lysine de 10 acides aminés.

Une observation très intéressante réalisée est que pour une même longueur donnée, des homopeptides d'arginine (R_{10} : $I_b/I_0 \approx 0,23$) et de lysine (K_{10} : $I_b/I_0 \approx 0,38$) entraînent des niveaux de blocage du courant différents, dans les mêmes conditions expérimentales. De plus, les durées moyennes des blocages du courant engendrés par les deux types de peptides sont différentes

d'environ deux ordres de grandeur, avec les homopeptides d'arginine qui présentent une durée moyenne de blocage du courant d'environ 100 ms, contre environ 1 ms pour les homopeptides de lysine.

Cette observation permet de dire qu'il est possible de discriminer ces homopeptides dont la nature chimique est différente avec le nanopore d'aérollysine.

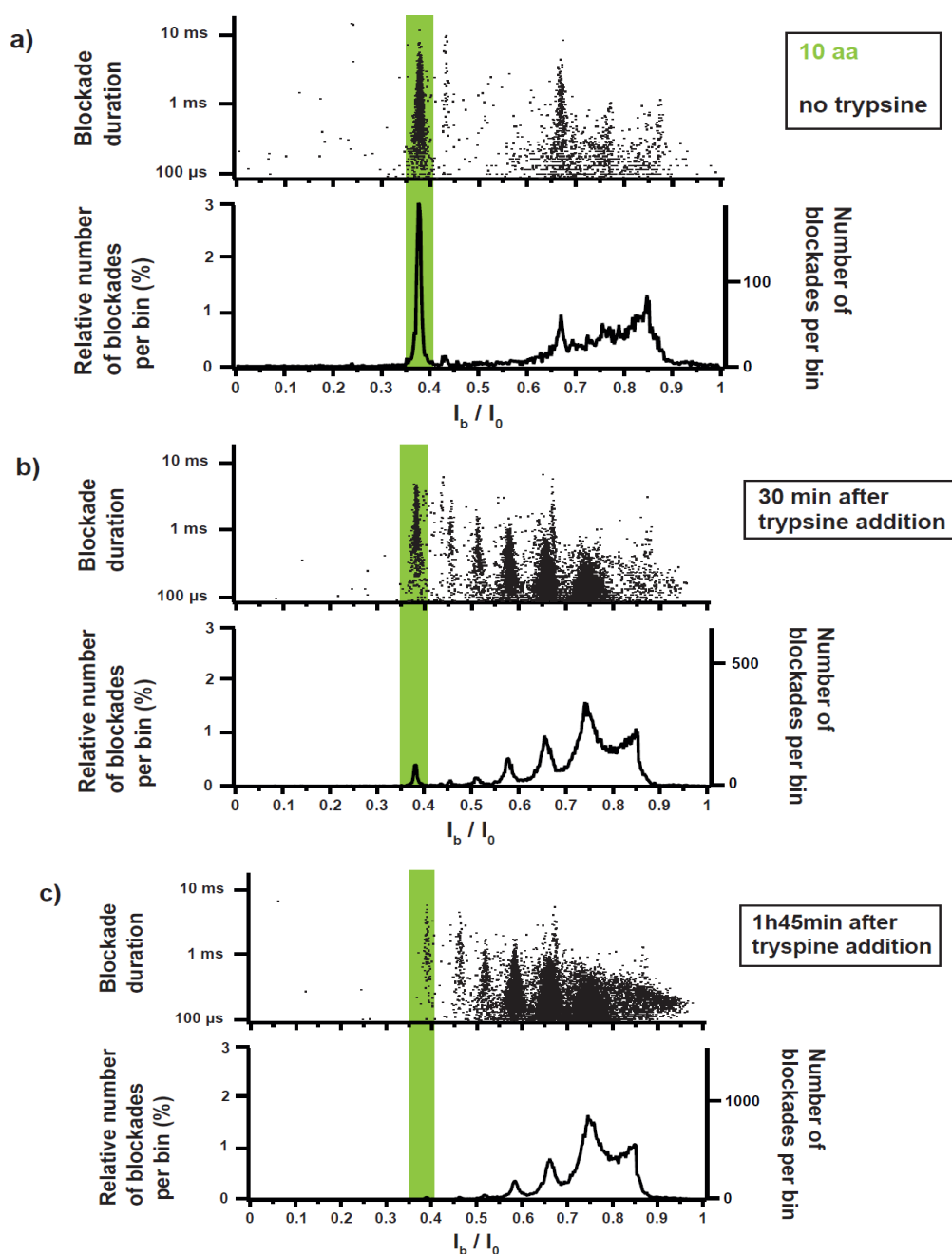


Figure 55 - Analyse d'homopeptides de lysine d'une longueur de 10 acides aminés avec le nanopore biologique d'aérollysine. Histogramme de I_b/I_0 (bas de chaque sous figure) et distribution des durées des blocages du courant (nuage de points) obtenus pour l'analyse d'homopeptides de lysine d'une longueur de 9 acides aminés : (a) en absence de Trypsine. (b)

après 30 minutes de digestion enzymatique par la trypsine. et (c) après une heure et 45 minutes de digestion enzymatique par la trypsine. Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de -50 mV, appliquée en trans. La trypsine est ajoutée dans le compartiment cis à une concentration finale de 266 nM.

2.7 Discrimination de peptides selon la nature chimique des acides aminés

Nous avons voulu aller plus loin en analysant un hétéropeptide d'une séquence composée de 5 acides aminés de lysine suivis de 5 acides aminés d'arginine (K_5R_5). Cette idée nous était d'abord venue après avoir observé que les homopeptides de lysine n'étaient pas détectables avec l'aérolysine lorsqu'ils étaient ajoutés en solution à une concentration finale de 1 μ M. L'idée était donc de lier un peptide 5 acides aminés d'arginine (R_5) à un peptide de 5 acides aminés de lysine (K_5) pour permettre l'entrée de ce dernier dans le pore. Bien que des conditions favorables aient été trouvées pour la détection des homopeptides de lysine comme décrit précédemment, nous avons quand même réalisé l'expérience d'analyse de ces hétéropeptides. L'addition d'une solution d'hétéropeptides K_5R_5 en présence du nanopore d'aérolysine a entraîné l'apparition de blocages du courant d'amplitudes et de durées variables (**Fig. 56, d**). En construisant l'histogramme de I_b/I_0 , une population principale apparaît à une valeur moyenne de I_b/I_0 de $0,297 \pm 0,003$ (**Fig. 56, b**). Cette population principale a été attribuée à l'hétéropeptide ajoutée en solution. Également, on peut observer sur cet histogramme la présence de populations minoritaires dont l'identité n'a pas été déterminée.

La position de la population majoritaire sur l'histogramme a ensuite été comparée à celle des homopeptides de 10 arginines (R_{10} ou RR-10) (**Fig. 56, a**) et des homopeptides de 10 lysine (K_{10} ou KK-10) (**Fig. 56, c**). Les homopeptides de 10 arginines présentent un I_b/I_0 moyen de $0,236 \pm 0,003$; les homopeptides de 10 lysines un I_b/I_0 de 0,38, et l'hétéropeptide de 5 lysines et 5 arginines un I_b/I_0 de $0,377 \pm 0,004$. Ces trois populations sont donc parfaitement discriminables par leurs niveaux de blocage du courant respectifs mais également par le temps moyen durant lequel chacune entraîne un blocage du courant (**Fig. 57**). Ce résultat met en avant qu'il est possible, pour des peptides d'une même longueur de 10 acides aminés, de détecter une différence de 5 acides aminés par le nanopore d'aérolysine, dans nos conditions expérimentales.

La possibilité de voir une différence entre un homopeptide de lysine (ou arginine) avec un hétéropeptide différant de celui-ci de 5 acides aminés, permet de dire que l'aérolysine possède donc une haute sensibilité permettant de voir cette différence de 5 acides aminés.

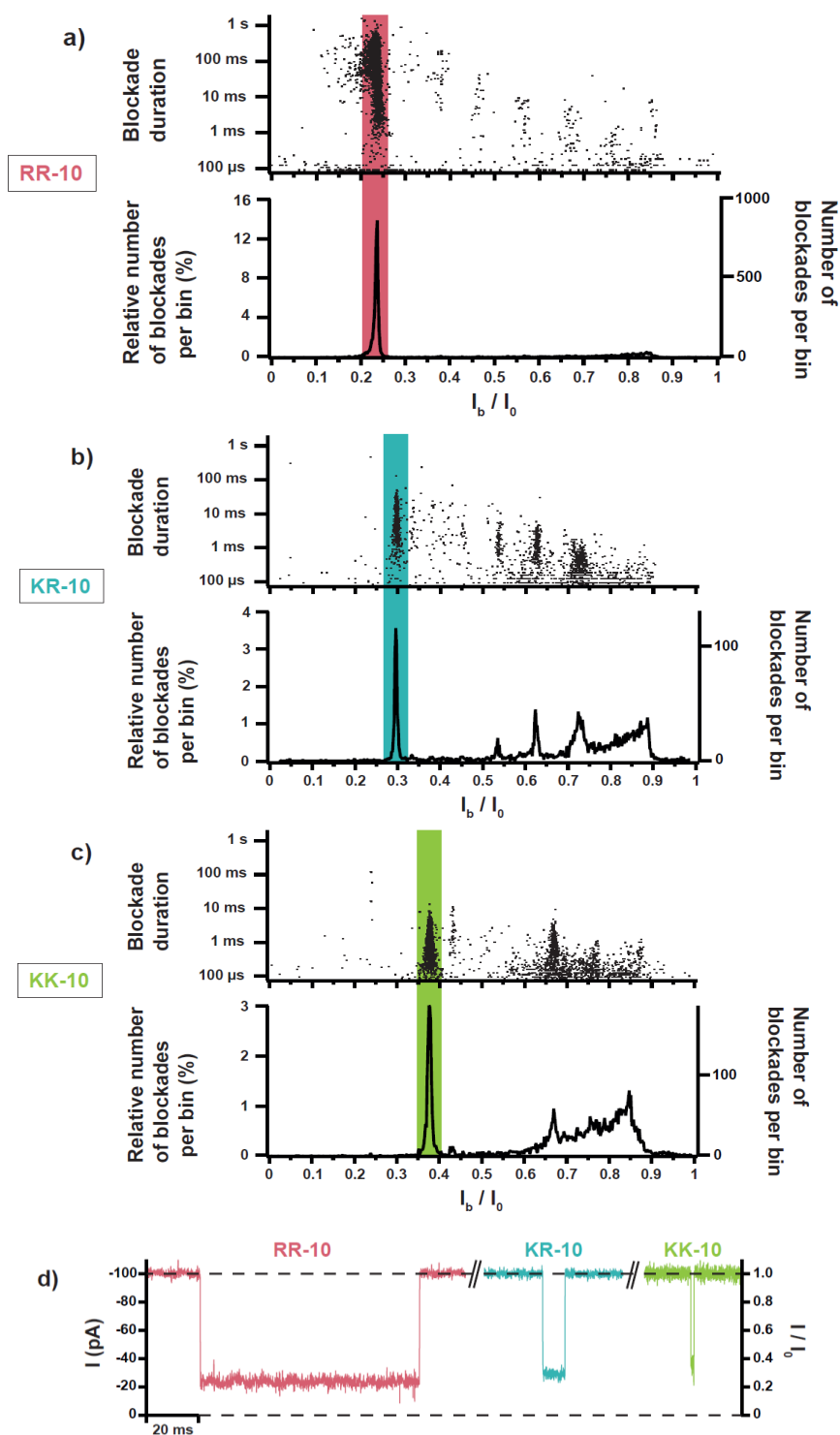


Figure 56 - Analyse d'hétéropeptides avec le nanopore d'aérolysine. Histogramme de I_b/I_0 (bas de chaque sous figure a, b, c) et distribution des durées des blocages du courant (nuage de points) obtenus pour l'analyse : (a) d'un homopeptide de 10 arginines (RR-10 ou R_{10}), (b) d'un hétéropeptide de 5 lysines et 5 arginines (KR-10 ou K_5R_5), et (c) d'un homopeptide de 10 lysines (KK-10 ou K_{10}). (d) Événements de blocage du courant typiques mesurés lors de l'analyse des trois peptides RR-10, KR-10, et KK-10, avec l'aérolysine, sous l'action d'une tension électrique de -50 mV appliquée en trans. Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de -50 mV, appliquée en trans. La concentration finale de chaque peptide en solution est de 1 μ M.

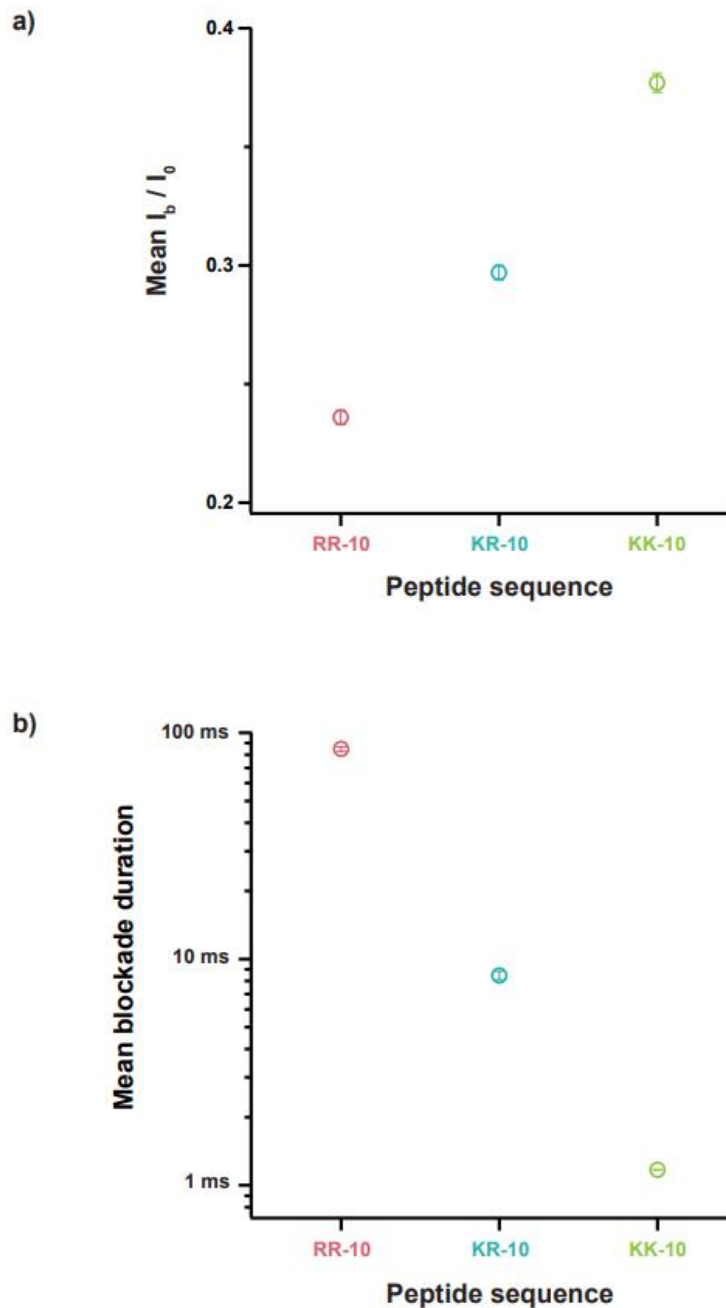


Figure 57 - Discrimination de peptides d'une longueur de 10 acides aminés de séquences différentes avec l'aérollysine. (a) Courant relatif moyen I_b/I_0 en fonction de la séquence peptidique, et, (b) durée moyenne de blocage du courant en fonction de la séquence peptidique, dans le cas de l'interaction du nanopore d'aérollysine avec : une solution de peptides RR-10 (rouge) ; d'une solution de peptides KR-10 (bleu turquoise), et d'une solution de peptides KK-10 (vert). Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C, sous l'action d'une tension électrique de -50 mV, appliquée en trans. La concentration finale de chaque peptide en solution est de 1 μ M.

III. Discussion & Conclusion

Les expériences que l'on a réalisées nous ont permis de montrer que le nanopore sauvage d'aérolysine permet de détecter différents types d'homopeptides courts, ceux d'arginine et de lysine, d'une longueur comprise entre 5 et 10 acides aminés. Cela signifie qu'il existe des interactions entre ces homopeptides et le nanopore d'aérolysine sous l'action d'une différence de potentiel négative. Le signe de cette différence de potentiel est en accord avec la charge positive des peptides à pH 7,5. L'arginine et la lysine ont des pHi supérieurs à 10 justifiant donc que ces homopeptides soit chargés positivement à ce pH. Ces interactions entre les peptides et le nanopore sont aussi probablement dues au caractère hautement chargé de l'aérolysine.

Aussi, nous avons montré que le nanopore d'aérolysine permet clairement de faire de la discrimination en taille d'homopeptides différant en longueur d'un unique acide aminé, dans le cas d'homopeptides d'arginine et de lysine. Il a été démontré dans les études précédentes de spectrométrie de taille de molécules par nanopore, auxquels j'ai participé, que le pouvoir de séparation des molécules de différentes tailles par un nanopore avec une résolution d'une unité monomérique est soumis à certaines conditions [177]. D'une part, la molécule doit entrer entièrement à l'intérieur du nanopore ; d'autre part, chaque unité monomérique doit contribuer à bloquer le courant de façon individuelle ; et enfin, le nanopore doit être très sensible aux changements de taille d'une molécule et il doit exister de fortes interactions entre les molécules et le nanopore.

Pour mesurer l'affinité existante entre les peptides d'arginine et le nanopore, des constantes cinétiques d'association k_{off} (s^{-1}) et de dissociation k_{on} ($\text{M}^{-1}.\text{s}^{-1}$) ont été calculées (**Fig. 58**). Pour les peptides d'arginine les plus longs (10 acides aminés), on a déterminé $k_{\text{off}} \approx 1,60 \pm 0,08.10^1 \text{ s}^{-1}$ et $k_{\text{on}} \approx 7,0 \pm 0,2.10^6 \text{ M}^{-1}.\text{s}^{-1}$; et pour les peptides les plus courts (5 acides aminés), $k_{\text{off}} \approx 1,03 \pm 0,03.10^3 \text{ s}^{-1}$ et $k_{\text{on}} \approx 3,50 \pm 0,06.10^6 \text{ M}^{-1}.\text{s}^{-1}$. Ces constantes ont permis de calculer la constante d'association d'équilibre $K(\text{M}^{-1}) = k_{\text{on}}/k_{\text{off}}$ qui donne une indication sur l'affinité entre les peptides d'arginine et le nanopore : pour les peptides les plus longs, $K \approx 4,38 \pm 0,01.10^6 \text{ M}^{-1}$; et pour les peptides les plus courts, $K \approx 3,50 \pm 0,05.10^3 \text{ M}^{-1}$. L'énergie libre standard correspondante a été calculée et correspond à $\Delta G^0 \approx 13 \pm 0,02 k_B T$ pour les peptides les plus longs (10, 9, 8, et 7 arginines), et $8,2 \pm 0,01 k_B T$ pour les peptides les plus courts (6 et 5 arginines).

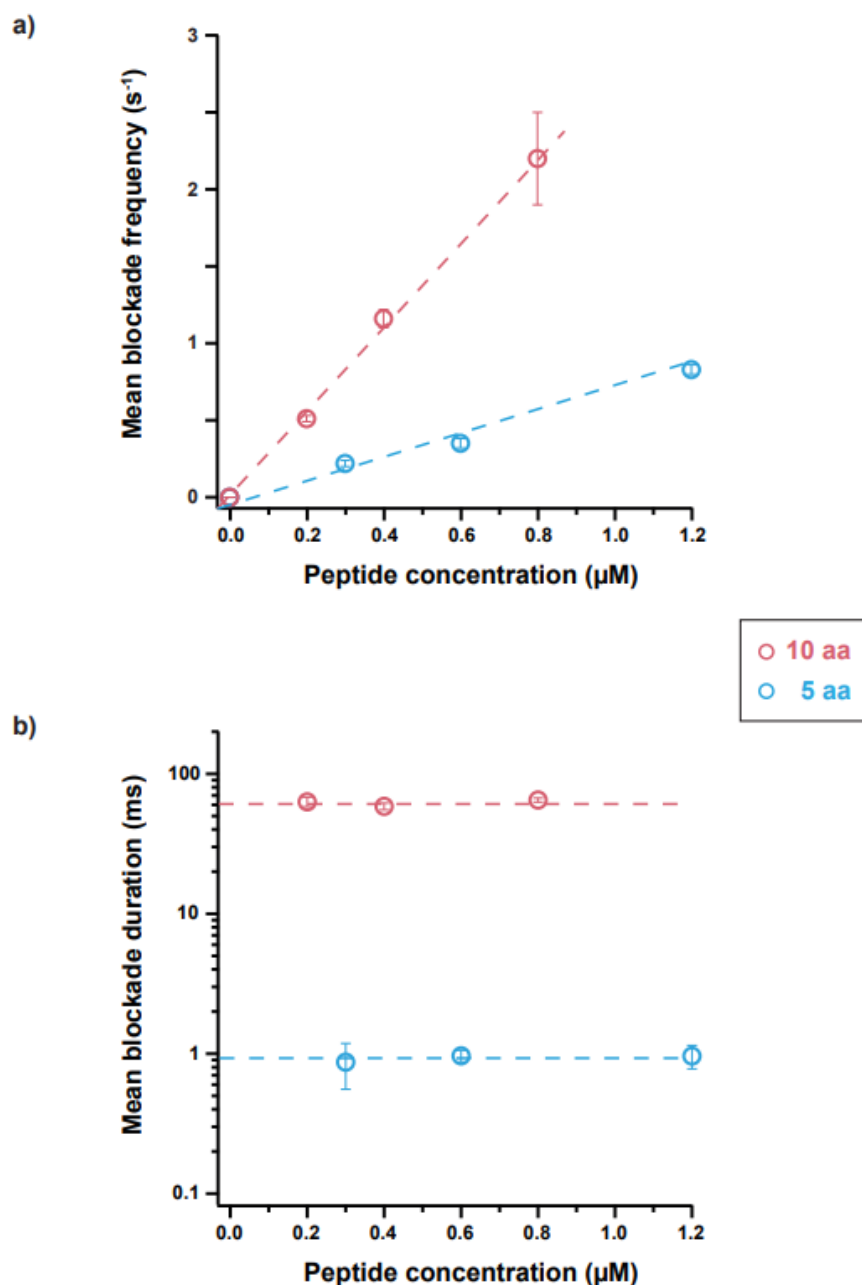


Figure 58 - Effet de la concentration en peptides sur l'interaction du nanopore d'aérollysine avec des homopeptides d'arginine de différentes longueurs. (a) Fréquence moyenne de blocage du courant en fonction de la concentration en peptides dans le cas de l'analyse d'une solution d'homopeptides d'une longueur de 10 acides aminés (rouge), ou d'une longueur de 5 acides aminés (bleu), avec le nanopore d'aérollysine. Les lignes en pointillés représentent les ajustements linéaires. L'augmentation linéaire de la fréquence moyenne de blocage du courant f en fonction de la concentration en peptides c , est interprété comme une réaction bimoléculaire réversible entre les peptides et l'aérollysine, avec une constante de réaction cinétique bien définie. La pente de l'ajustement linéaire permet de déterminer la constante d'association k_{on} ($M^{-1}.s^{-1}$) = f/c pour chaque peptide d'une longueur donnée. (b) Durée moyenne de blocage du courant en fonction de la concentration en peptides dans le cas de l'analyse d'une solution d'homopeptides d'une longueur de 10 acides aminés (rouge), ou d'une longueur de 5 acides

aminés (bleu), avec le nanopore d'aérolysine. La durée moyenne de blocage du courant t est indépendante de la concentration en peptides. L'inverse de t est défini comme étant la constante de dissociation k_{off} (s^{-1}) = $1/t$. La constante d'association d'équilibre est définie comme K (M^{-1}) = $k_{\text{on}}/k_{\text{off}}$, donne une indication sur l'affinité entre les peptides et le nanopore d'aérolysine. L'énergie libre standard correspondante est définie comme $\Delta G^0 = -k_{\text{B}}T \ln(1/K)$, où $k_{\text{B}}T$ est l'énergie d'agitation thermique $\approx 4.10 \cdot 10^{-21}$ J à 20°C . Les mesures ont été réalisées dans une solution de KCl 4M, HEPES 5mM, pH = 7,5, à 20°C , sous l'action d'une tension électrique de -50 mV, appliquée en trans.

Cette étude cinétique permet donc de dire que les interactions entre les homopeptides d'arginine et l'aérolysine correspondent à une réaction bimoléculaire réversible, et que ces interactions peptides-pore sont très fortement dépendantes de la longueur du peptide.

Ainsi, il est possible de penser que la haute capacité du nanopore d'aérolysine à permettre ici la détection de très courts peptides est liée à la forte affinité existante entre les peptides et le nanopore.

De plus, nous avons montré par ces résultats que notre système permet de détecter et caractériser des impuretés présentes à un taux inférieur à 2% dans un échantillon peptidique, alors que deux techniques conventionnelles de caractérisation des peptides (spectrométrie de masse et HPLC) ont permis de détecter la présence d'impuretés mais n'ont pas permis d'identifier, ni de quantifier ces impuretés

Chapitre 4 : Signature électrique des 20 acides aminés protéinogènes par le nanopore d'aérolysine

I. Introduction

Comme mentionné précédemment le séquençage de protéines à l'échelle de la molécule unique par nanopore constitue un nouveau challenge pour le monde de la biologie et de la santé. Plusieurs techniques d'analyse à l'échelle de la molécule unique ont été développées depuis quelques années en alternative aux techniques classiques d'analyse des protéines. Parmi celles-là on compte le fluoroséquençage [17,19], le FRET fingerprinting [16,18], ou encore le séquençage par nanopore [218,219]. Concernant le séquençage de protéines par nanopore, les stratégies décrites dans le chapitre n°1 consistant à lire la séquence d'une protéine lors de son transport unidirectionnel à travers un nanopore se sont heurtées à des difficultés principalement liées à l'hétérogénéité de charge des protéines, à l'existence de 20 acides aminés naturels, et à l'existence de modifications post-traductionnelles.

Pour tenter de répondre à ce défi, nous avons proposé, suite aux premiers travaux de cette thèse décrits dans le chapitre n°3, d'explorer le potentiel de l'aérolysine pour obtenir la distinction des 20 acides aminés.

Nous avons utilisé une amorce peptidique correspondant à un homopeptide de 7 acides aminés d'arginine chargés positivement à pH 7,5, lié à un huitième acide aminé dont la nature chimique varie parmi les 20 acides aminés naturels. Sous l'action d'une force électrique, l'amorce d'arginine entraîne chaque acide aminé dans le nanopore. L'obtention d'une signature électrique pour les 20 acides aminés liés à l'amorce peptidique serait une étape décisive renseignant sur la possibilité de détecter et de discriminer des acides aminés individuels avec l'aérolysine, dans la perspective du séquençage de protéines avec ce nanopore.

Les résultats obtenus montrent que le nanopore d'aérolysine nous a permis d'établir, pour la première fois, une cartographie des 20 acides aminés protéinogènes, et de proposer une stratégie de séquençage. Ce travail a été publié dans *Nature Biotechnology* en Décembre 2019 (**Annexe 3**).

Ces travaux nous ont permis de proposer une nouvelle stratégie de séquençage qui, contrairement aux approches précédentes proposant de lire la séquence d'une protéine au cours de sa translocation à travers un nanopore, serait de cliver la protéine à son extrémité pour libérer

les acides aminés un par un, et de fixer l'acide aminé libéré à une chaîne d'arginine. La chaîne d'arginine permettrait le transport de l'acide aminé vers le nanopore. Le complexe serait détecté par le nanopore et produirait une signature électrique spécifique de l'acide aminé lié permettant de reconstituer la séquence de la protéine successivement.

II. Résultats & discussion

2.1 Détection des 20 acides aminés protéinogènes

Dans un premier temps, nous avons réalisé des expériences d'analyse individuelle de chacun des vingt peptides. Pour cela, dans chaque expérience, le peptide d'intérêt X_{R7} est ajouté en solution dans le compartiment *Cis*, pour atteindre une concentration finale de 1 μM (**Fig. 59**).

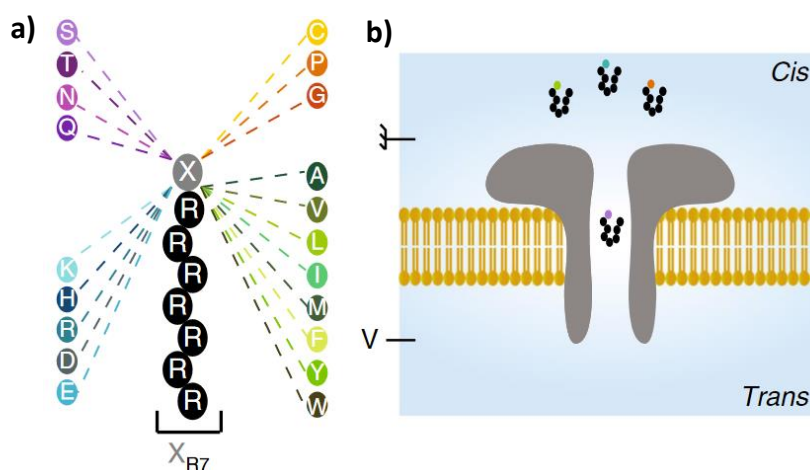


Figure 59 - Schéma illustratif du dispositif expérimental. (a) Schéma de la structure des peptides X_{R7} étudiées (X est : électriquement chargée (bleu) ; polaire non chargée (violet) ; hydrophobe aromatique et non aromatique (vert) ; et le reste des acides aminés que l'on a désigné de cas spéciaux (jaune)), et, (b) du dispositif expérimental montrant un nanopore d'aérolysine inséré dans une membrane de lipides en présence de peptides X_{R7} ajoutés en solution dans le compartiment *cis*. L'entrée des chaînes dans le pore se fait sous l'action d'une tension électrique négative appliquée en *Trans*. Les expériences ont été réalisées en présence d'une solution de KCl à une concentration de 4 M, ajustée et stabilisée à pH 7,5 par l'utilisation d'une solution tampon d'HEPES à une concentration de 25 mM.

L'addition des peptides X_{R7} entraîne l'apparition de blocages transitoires du courant ionique sous l'action d'une différence de potentiel négative (**Fig. 60**). En cohérence avec les résultats obtenus dans le chapitre n°4, il est donc possible de détecter ces peptides chargés positivement en raison de la présence de l'amorce d'arginine, sous l'action d'une tension électrique négative.

L'analyse individuelle des 20 peptides révèle que chacun d'eux est à l'origine d'un blocage du courant d'amplitude spécifique. Pour caractériser les signaux électriques de chacun des 20 peptides, nous avons réalisé des histogrammes du courant relatif I_b/I_0 pour chacun d'eux. Par exemple, le peptide S_{R7} produit des blocages du courant visibles sur la trace de courant (**Fig. 60, d**) présentant une amplitude caractéristique qui est reproduite à chaque nouveau blocage du courant, et qui est à l'origine d'un pic bien défini à une valeur de I_b/I_0 de $0,390 \pm 0,002$ (**Fig. 60, e**) sur l'histogramme de I_b/I_0 . La distribution des durées des blocages du courant en fonction de I_b/I_0 montre également qu'une population se dessine à la même valeur de I_b/I_0 (**Fig. 60, e**).

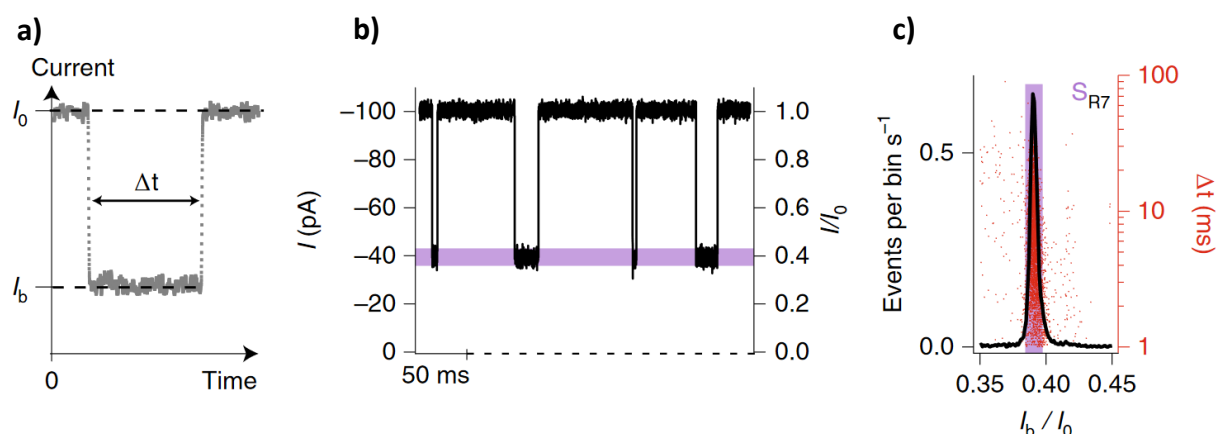


Figure 60 - Détection et analyse d'un peptide X_{R7} avec l'aérolysine. (a) Représentation d'un événement de blocage du courant typique. (b) Portion d'une trace de courant, et (c) histogramme du courant relatif I_b/I_0 superposé à la distribution des durées des blocages en fonction de I_b/I_0 , dans le cas de l'analyse du peptide S_{R7} avec le nanopore d'aérolysine (e). Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

Les vingt peptides ont été classifiés en cinq familles selon l'état de charge de la chaîne latérale de l'acide aminé X du peptide X_{R7} à pH 7,5 (**Fig. 59, a**). Nous avons choisi de définir cinq familles pour lesquelles la chaîne latérale de l'acide aminé X est : électriquement chargée (bleu) ; polaire non chargée (violet) ; hydrophobe aromatique et non aromatique (vert) ; et le reste des acides aminés que l'on a désigné de cas spéciaux (jaune).

De façon à déterminer si les signaux engendrés par l'entrée de ces peptides dans le nanopore sont distinguables, nous avons réalisé une représentation dans laquelle les histogrammes de I_b/I_0 obtenus pour l'analyse individuelle de chacun des vingt peptides sont superposés. Les peptides de la famille sont regroupés sur un même histogramme pour plus de clarté (**Fig. 61, f-j**).

La superposition des histogrammes montre pour chaque famille, l'apparition de populations à des valeurs de I_b/I_0 caractéristiques et bien définies et qui apparaissent dans certains cas très

bien séparées, alors que dans d'autres cas les populations observées sont trop proches pour être facilement distinguables.

Concernant la famille des peptides dont l'acide aminé X présente une chaîne latérale électriquement chargée (R_{R7} , K_{R7} , H_{R7} , E_{R7} , et D_{R7}), les populations obtenues pour l'analyse des cinq peptides R_{R7} , K_{R7} , H_{R7} , E_{R7} , et D_{R7} , sont localisées à des valeurs de I_b/I_0 qui sont clairement distinguables les unes des autres (**Fig. 61, f**). Il est facile d'identifier chaque peptide de cette famille par l'amplitude du blocage du courant qu'il induit par sa présence dans le nanopore (**Fig. 61, f**).

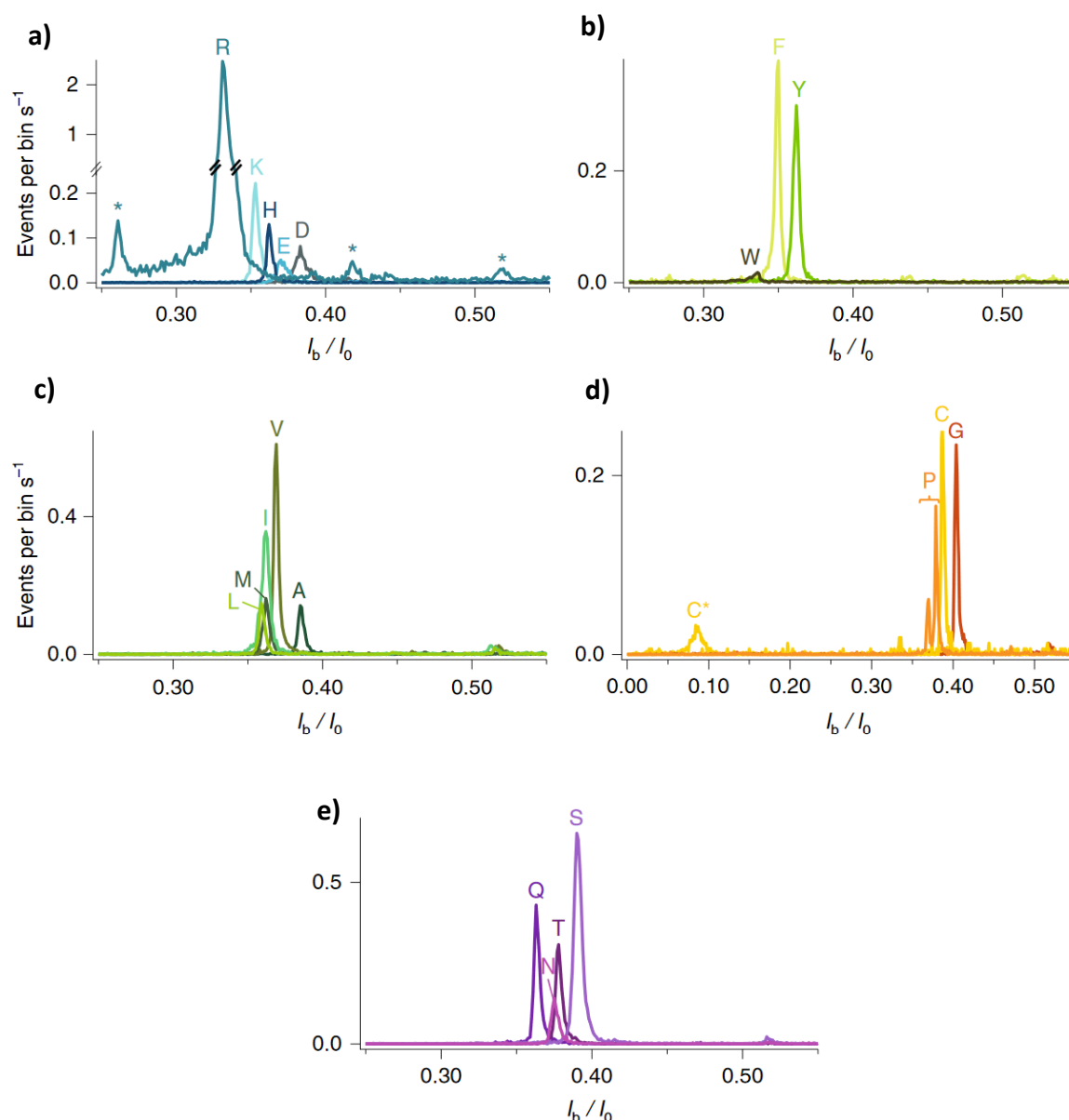


Figure 61 - Analyse des peptides X_{R7} avec l'aérolysine. Superposition des histogrammes de I_b/I_0 obtenus par l'analyse individuelle des vingt peptides X_{R7} , avec le nanopore d'aérolysine, sous l'action d'une tension électrique positive de -50 mV. Les peptides sont groupés selon les

propriétés de la chaîne latérale de l'acide aminé X : électriquement chargée (a, bleu) ; hydrophobe aromatique et non aromatique (b et c, vert) ; cas spéciaux (d, jaune) ; et polaire non chargée (e, violet). Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

C'est également le cas pour la famille des peptides dont l'acide aminé X présente une chaîne latérale hydrophobe aromatique (F_{R7}, Y_{R7}, et W_{R7}) pour laquelle trois populations sont bien définies et distinctes (**Fig. 61, g**). On peut observer que la population correspondant au peptide W_{R7} présente une intensité très faible directement liée à la fréquence d'entrée (taux de capture) de ce peptide dans le pore qui est beaucoup plus faible que pour les autres peptides.

Concernant la famille des peptides dont l'acide aminé X présente une chaîne latérale polaire non chargée (**Fig. 61, h**), il est également possible de distinguer quatre populations correspondant aux quatre peptides de cette famille. Contrairement aux familles précédentes, deux populations correspondant aux blocages du courant engendrés par les peptides T_{R7} et N_{R7} présentent des valeurs de I_b/I_0 qui sont trop proches pour être assez facilement distinguables (**Fig. 61**).

C'est également le cas pour la famille des peptides dont l'acide aminé X présente une chaîne latérale hydrophobe non aromatique (**Fig. 61, i**) dans laquelle les populations correspondant aux peptides L_{R7}, M_{R7}, et I_{R7} sont également trop proches pour être facilement discriminables (**Fig. 61 et 62**).

La mesure de signaux très proches peut provenir des volumes de ces acides aminés qui sont très proches dans chaque cas comme discuté plus loin. Une autre hypothèse pourrait être que ces acides aminés, en interagissant avec l'amorce peptidique d'arginine, pourraient adopter une conformation de volume proche.

De façon générale, ces résultats montrent qu'il existe un lien entre la nature de l'acide aminé X des peptides X_{R7} et l'amplitude des blocages du courant engendrés par la présence de chacun de ces peptides dans le nanopore d'aérolysine. Les différents peptides présentent des valeurs de courant relatif I_b/I_0 caractéristiques, qui leur permettent d'être dans de nombreux cas distinguables (**Fig. 62**).

Enfin, concernant la dernière famille des peptides dont l'acide aminé X fait partie des cas spéciaux (C_{R7}, P_{R7}, et G_{R7}), quatre populations sont observées sur l'histogramme de I_b/I_0 alors que l'on s'attendait intuitivement à voir l'apparition de trois populations caractéristiques de

chacun des trois peptides de cette famille (**Fig. 61, j**). C'est l'analyse du peptide P_{R7} qui est à l'origine de blocages du courant de deux amplitudes caractéristiques, ce qui a pour origine l'apparition de deux populations voisines sur l'histogramme de I_b/I_0 . Plusieurs explications peuvent justifier ce résultat. Il est connu d'une part, que l'orientation par laquelle entre le peptide dans le pore (N-terminal ou C-terminal) peut influencer le signal obtenu comme cela a déjà été montré dans le cas de l'ADN avec des différences entre l'extrémité 5' et 3' [310], mais également dans le cas des protéines [24,220]. D'autre part, ce résultat peut être lié à l'existence de deux sites préférentiels d'interactions du peptide avec le nanopore, ce qui serait à l'origine de deux signaux différents. Enfin, une autre explication peut provenir du fait que la proline est un acide aminé particulier dit en réalité alpha-iminé en raison de son extrémité N-terminal qui est directement liée à sa chaîne latérale.

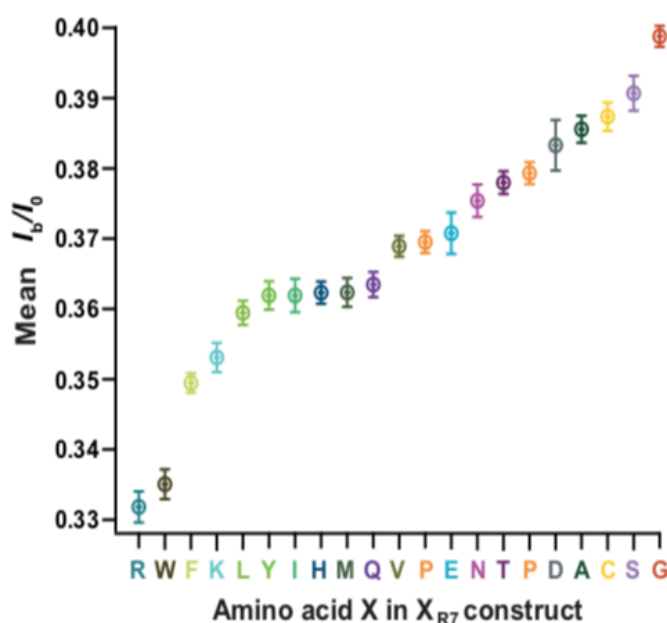


Figure 62 - Vue d'ensemble des valeurs du courant relatif résiduel I_b/I_0 pour les peptides X_{R7} analysés avec le nanopore d'aérolysine.. Distribution des valeurs moyennes de I_b/I_0 et leurs déviations standards obtenus dans le cas de l'analyse individuelle des peptides X_{R7} , rangées dans l'ordre. Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

Dans cette dernière famille, un autre résultat particulier concerne le peptide C_{R7} pour lequel la population observée sur l'histogramme présente une valeur moyenne de I_b/I_0 proche de 0,1 très différente des dix-neuf autres peptides qui présentent tous une valeur moyenne de I_b/I_0 comprise entre environ 0,3 et 0,5. Une des explications proposée à ce résultat est directement liée à la nature chimique de l'acide aminé cystéine qui, de par la présence d'une fonction thiol (-SH) sur

sa chaîne latérale, est le seul acide aminé capable de former des ponts disulfures. Ainsi, nous avons interprété ce résultat comme étant lié à la présence de dimère de C_{R7} .

Pour vérifier cette hypothèse, nous avons réalisé une expérience d'analyse individuelle du C_{R7} , en absence, puis en présence d'un réducteur de ponts disulfures, le dithiothréitol (DTT), ajouté dans le même compartiment (*cis*) à une concentration finale de 25 mM afin d'observer l'évolution du signal en temps réel (**Fig. 63**). L'analyse du peptide C_{R7} seul laisse de nouveau apparaître des blocages du courant, dont l'amplitude qui est caractéristique et reproduite, est à l'origine d'une population localisée à une valeur de I_b/I_0 de $0,084 \pm 0,005$ sur l'histogramme de I_b/I_0 . Quinze minutes après l'ajout de DTT, une diminution de l'intensité de la première population localisée à la valeur de I_b/I_0 de $0,084$ est observée au profit de l'apparition d'une nouvelle population localisée à une valeur de I_b/I_0 de $0,383 \pm 0,005$. Trente minutes après l'ajout de DTT, la première population a presque totalement disparu au profit d'une augmentation de la seconde population. Cette expérience confirme le fait que les peptides C_{R7} présents en solution s'assemblent en dimères *via* la formation de ponts disulfures.

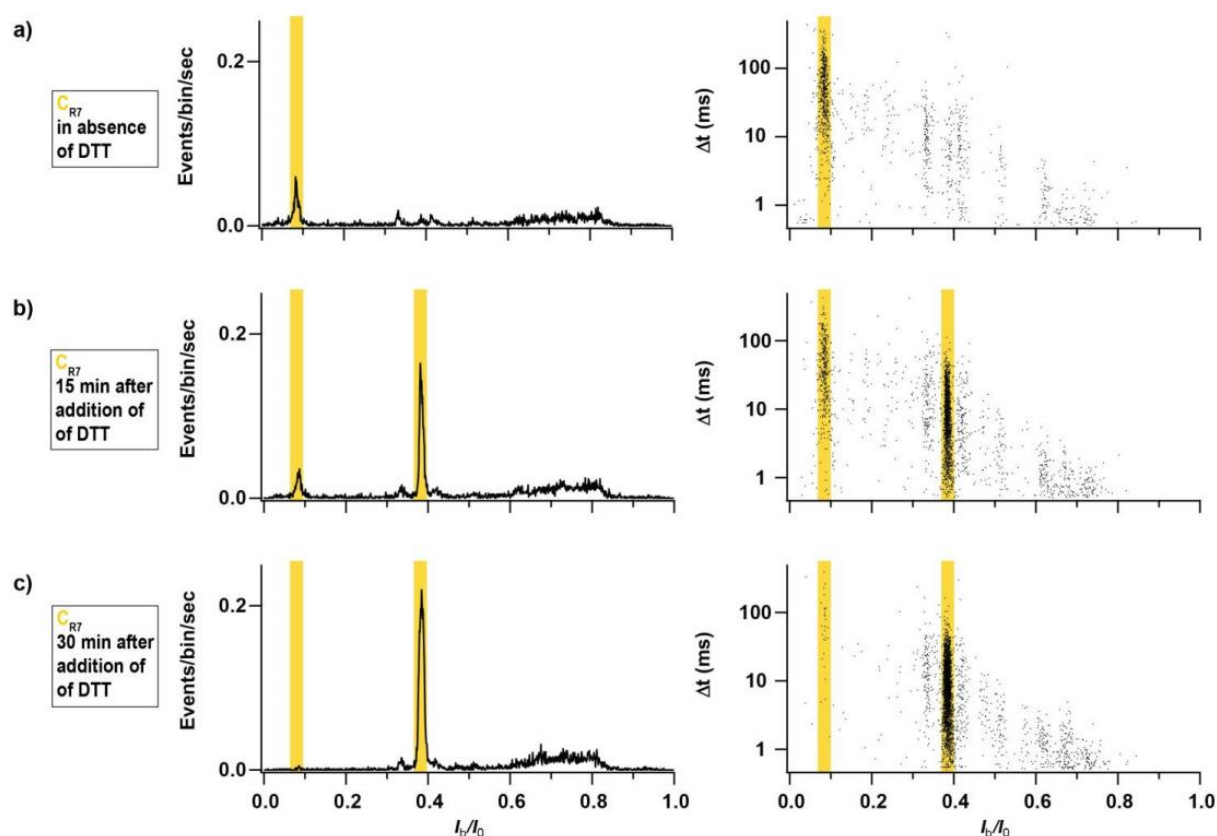


Figure 63 - Analyse des peptides C_{R7} en présence ou en absence d'un agent réducteur de ponts disulfures. Histogrammes de I_b/I_0 et distribution des durées des blocages du courant en fonction de I_b/I_0 obtenus pour l'analyse du peptide C_{R7} en absence de l'agent réducteur DTT

(a), 15 minutes après ajout de DTT (b), et 30 minutes après ajout de DTT (c). Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

2.2 Discrimination de modifications chimiques subtiles

En analysant de façon plus précise nos résultats et en comparant la structure des acides aminés entre eux, deux résultats surprenants ont été montrés.

D'une part, nous avons précédemment montré qu'il est possible de distinguer les peptides F_{R7} et Y_{R7} dans la famille des acides aminés dont la chaîne latérale est hydrophobe aromatique (**Fig. 61, g**). La différence existant entre les acides aminés phénylalanine (F) et tyrosine (Y) correspond à la présence d'une fonction hydroxyle (-OH) observée dans la composition chimique de la tyrosine mais pas dans celle de la phénylalanine. Cette différence correspond à une variation en masse molaire de 17 g/mol. Ce résultat montre que le nanopore d'aérolysine permet donc de détecter des différences aussi subtiles que l'ajout ou la suppression d'une fonction hydroxyle.

D'autre part, nous avons également montré qu'il est possible de discriminer les peptides L_{R7} et I_{R7} (**Fig. 61, i**), bien que les populations apparaissant sur l'histogramme de I_b/I_0 pour chacun d'elles sont très proches et moins facilement distinguables (**Fig. 62**). Dans le cas des peptides L_{R7} et I_{R7} , la différence existant entre la leucine (L) et l'isoleucine (I) est uniquement structurale. Les deux acides aminés sont des isomères de structures possédant exactement la même composition chimique et la même masse molaire. Ce résultat montre que le nanopore d'aérolysine permet donc de discriminer deux peptides de masses molaires identiques. Ainsi, la discrimination des peptides X_{R7} observés avec le nanopore d'aérolysine n'est pas uniquement liée aux différences de masses qui existent entre les peptides.

Pour la suite de l'étude, de façon à renforcer nos résultats, nous avons vérifié que les résultats de discrimination des peptides X_{R7} , préalablement analysés individuellement, n'ont pas pour origine des artefacts du signal mais qu'ils sont bien liés à la sensibilité du pore à détecter les différences entre les acides aminés X de chacun des vingt peptides. Pour cela, nous avons réalisé des mélanges de peptides en solution en présence du nanopore d'aérolysine.

2.3 Discrimination des 20 acides aminés protéinogènes dans un mélange

Des mélanges de peptides ont été analysés avec le nanopore d'aérolysine. Pour cela, les peptides à comparer ont soit été ajoutés simultanément dans le compartiment *cis* ; ou soit été ajoutés

successivement, un par un, dans le compartiment *cis* de façon à suivre l'apparition progressive des populations correspondant aux peptides. Dans ces expériences, les peptides ont également tous été ajoutés de façon à obtenir une concentration finale de 1 μ M dans le compartiment *cis*. Les mélanges réalisés dans chaque expérience sont équimolaires.

L'addition d'un mélange équimolaire des cinq peptides K_{R7} , H_{R7} , E_{R7} , D_{R7} , et R_{R7} , correspondant à la même famille, entraîne l'apparition de blocages du courant qui peuvent être individuellement attribués à chacun des peptides en se basant sur la signature électrique mesurée lors de leur analyse individuelle (**Fig. 64, a**). L'histogramme correspondant à ce mélange montre l'apparition de cinq populations (**Fig. 64, b**) à des valeurs de I_b/I_0 de $0,353 \pm 0,002$ (K_{R7}) ; $0,362 \pm 0,002$ (H_{R7}) ; $0,371 \pm 0,003$ (E_{R7}) ; $0,383 \pm 0,004$ (D_{R7}) ; $0,332 \pm 0,002$ (R_{R7}) ; se superposant avec les populations obtenues lors de l'analyse individuelle de chacun des peptides. Les cinq populations peuvent donc être attribuées aux cinq peptides à partir de leur valeur de I_b/I_0 .

Une vérification supplémentaire a consisté à ajouter successivement chacun des peptides de cette famille de façon à observer l'apparition des populations une par une (**Fig. 65**). De la même façon, cette expérience montre que les résultats sont reproductibles. Les populations sont apparues successivement à des valeurs de I_b/I_0 superposables à la fois à celles obtenues dans le mélange précédemment décrit et à celles obtenues lors de l'analyse individuelle des peptides.

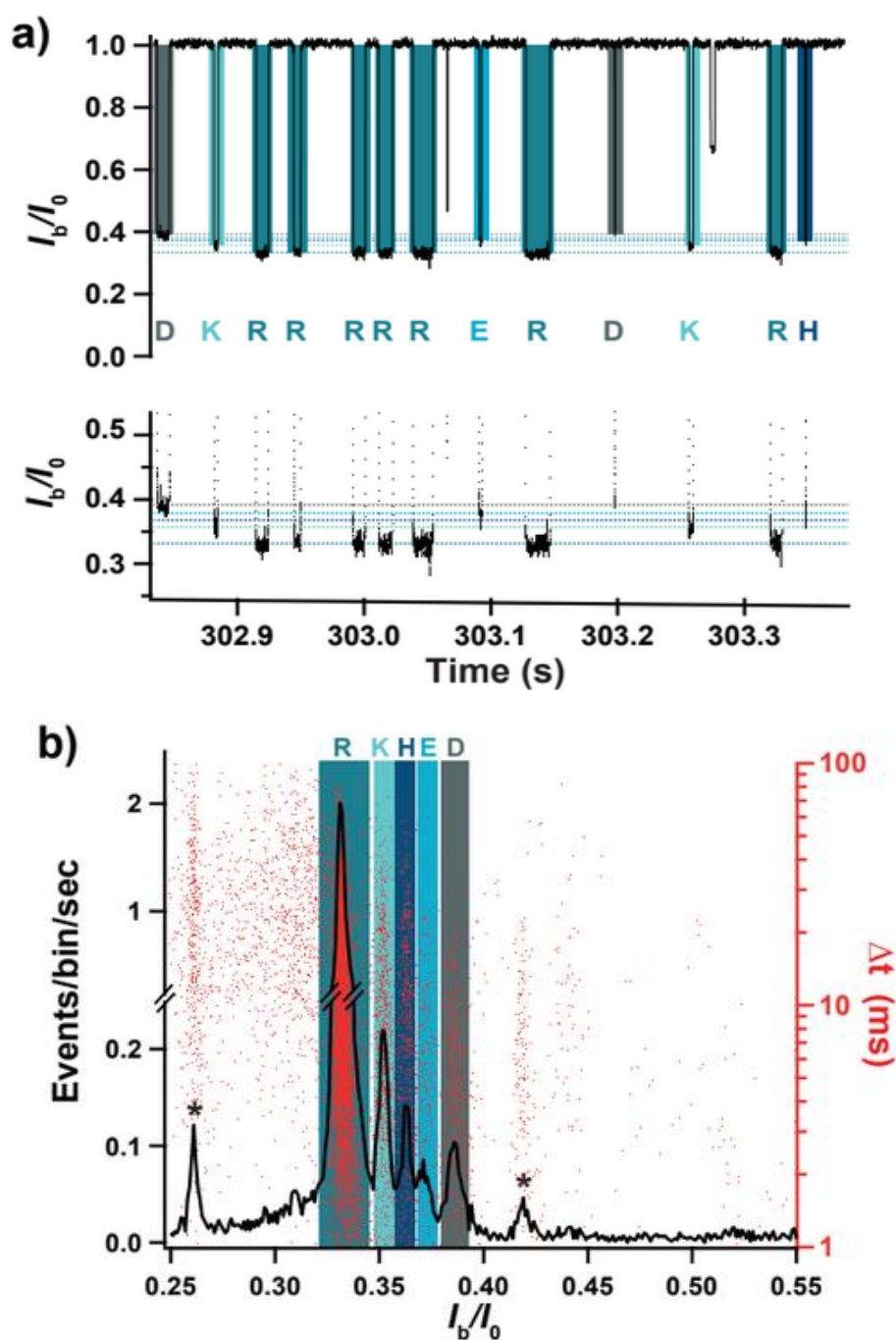


Figure 64 - Discrimination de peptides individuels au sein d'un mélange avec le nanopore d'aérollysine. (a) Portion d'une trace de courant en fonction du temps obtenue pour l'analyse d'un mélange équimolaire de peptides R_{R7} , K_{R7} , H_{R7} , E_{R7} , et D_{R7} avec le nanopore d'aérollysine et élargissement de la portion de la trace du courant montre les différents niveaux de courant. (b) Histogramme de I_b/I_0 (ligne noire) et distribution des durées des blocages du courant (points rouges) en fonction de I_b/I_0 , obtenus pour l'analyse du mélange des peptides mentionnés en (a). Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

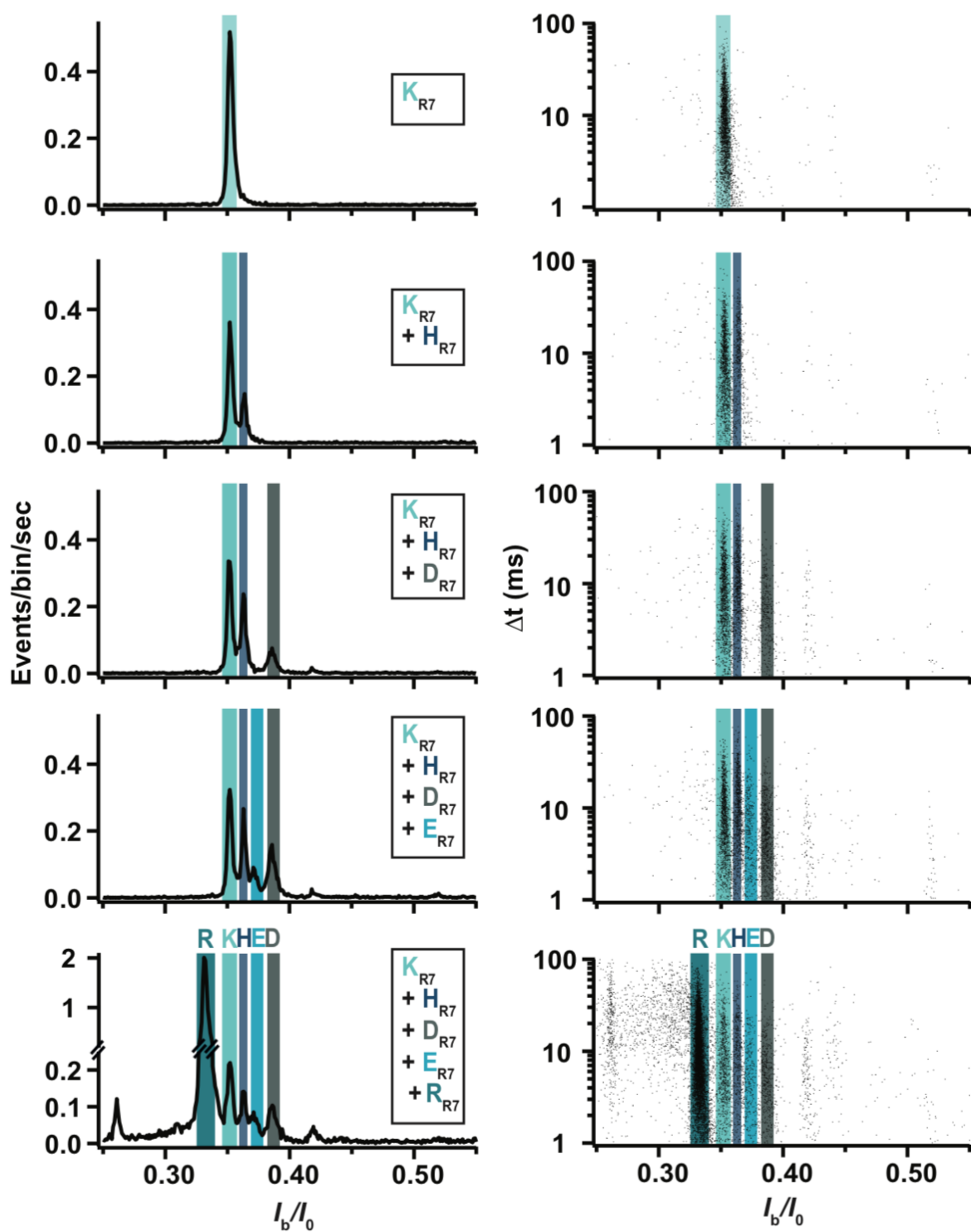


Figure 65 - Identification de peptides individuels au sein d'un mélange avec le nanopore d'aérolysine. Histogramme de I_b/I_0 (gauche) et distribution des durées des blocages du courant en fonction de I_b/I_0 (nuage de points, droite) dans le cas de l'analyse avec l'aérolysine d'un mélange équimolaire des peptides K_{R7} , H_{R7} , E_{R7} , D_{R7} et R_{R7} , qui ont été ajoutés successivement en solution. Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

Nous avons également réalisé des expériences d'un mélange des peptides F_{R7} et Y_{R7} , et d'un mélange des peptide L_{R7} et I_{R7} , de façon également à vérifier et confirmer la discrimination de ces peptides.

L'addition d'un mélange équimolaire des peptides F_{R7} et Y_{R7} entraîne l'apparition de blocages du courant se traduisant par l'émergence de deux populations localisées à des valeurs de I_b/I_0 clairement distinctes sur l'histogramme de I_b/I_0 ainsi que sur le nuage de points de la distribution des durées des blocages du courant (**Fig. 66, haut**).

C'est également le cas pour l'addition d'un mélange de peptides L_{R7} et I_{R7} qui est responsable de l'apparition de deux populations qui sont néanmoins beaucoup plus proches que dans le cas de l'analyse des peptides F_{R7} et Y_{R7} (**Fig. 66, bas**) mais qui restent distinguables. Nous avons également vérifié que ces populations sont distinguables lors de l'ajout successif des peptides.

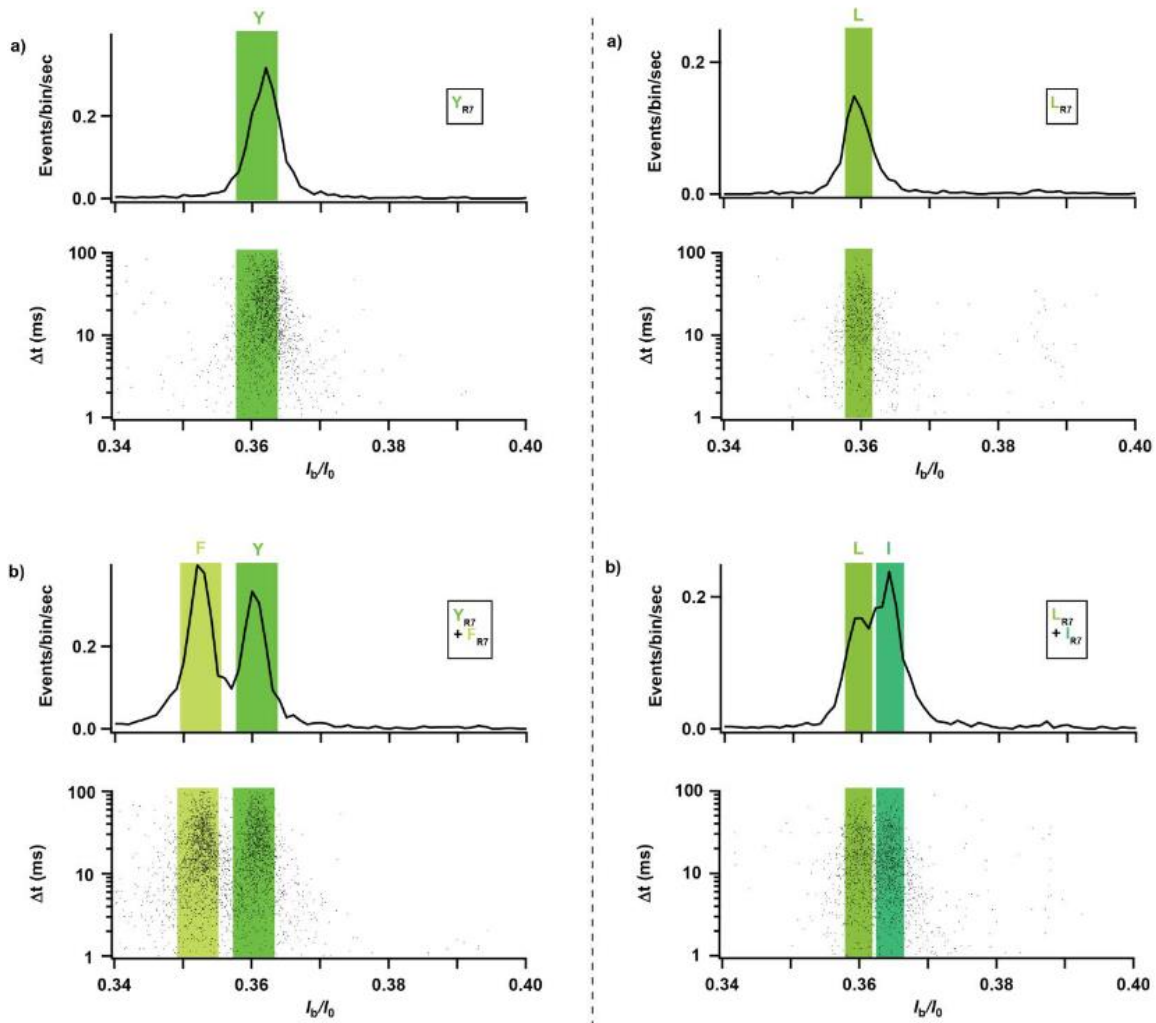


Figure 66 - Identification de peptides présentant des différences subtiles avec le nanopore d'aérolysine. Histogramme de I_b/I_0 (haut de chaque sous figure) et distribution des durées des

blocages du courant en fonction de I_b/I_0 (nuage de points, bas de chaque sous figure) obtenus dans le cas de l'analyse avec l'aérolysine : (gauche) d'un mélange équimolaire des peptides F_{R7} et Y_{R7} dont la différence correspond à une fonction hydroxyle et qui ont été ajoutés successivement en solution en commençant par le peptide Y_{R7} (a), puis F_{R7} (b) ; et, (droite) d'un mélange équimolaire des peptides L_{R7} et I_{R7} qui sont des isomères de structure de même masse molaire, et qui ont été ajoutés successivement en solution en commençant par le peptide Y_{R7} (a), puis F_{R7} (b). Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en Trans.

2.4 Dépendance en masse et en volume

De manière à explorer quels paramètres seraient à l'origine du lien existant entre les amplitudes des blocages du courant et les acides aminés, nous avons représenté la dépendance du courant relatif I_b/I_0 moyen en fonction de la masse moléculaire et en fonction du volume hydrodynamique [311] de l'acide aminé X du peptide X_{R7} (**Fig. 67**). Dans les deux cas, on constate que plus le volume ou la masse moléculaire augmente, plus le courant relatif I_b/I_0 diminue, c'est-à-dire que l'amplitude du blocage du courant engendré est plus importante. Ce résultat est cohérent avec ceux précédemment obtenus dans le cas de la discrimination en masse de polymères synthétiques de PEG [26] et d'homopeptides d'arginine [312]. Quelques exceptions ne suivent pas cette tendance comme c'est le cas pour les peptides F_{R7} et Y_{R7} . De façon générale, la distribution est en meilleure corrélation avec le volume hydrodynamique qu'avec la masse molaire. Par exemple, des valeurs du courant relatif résiduel I_b/I_0 distinctes sont observées dans le cas de deux acides aminés disposant de masses molaires très proches mais de deux volumes différents (*e.g* E_{R7} et K_{R7}), et *vice versa*, des amplitudes de blocages du courant superposables sont obtenues dans le cas de deux acides aminés disposant de volumes très proches mais de masses molaires différentes (*e.g* F_{R7} et Y_{R7}).

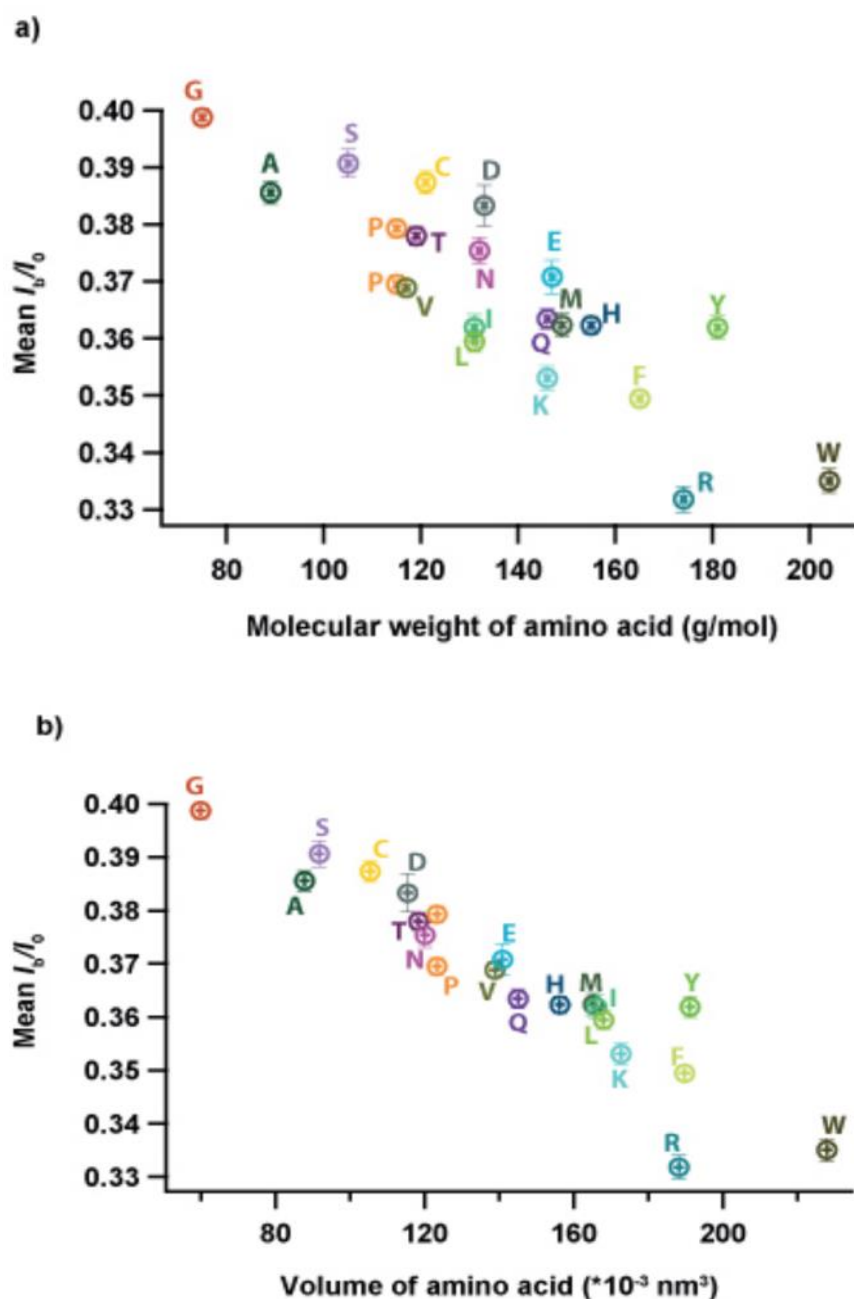


Figure 67 - Dépendance du courant relatif résiduel I_b/I_0 moyen à la masse molaire ou au volume hydrodynamique de l'acide aminé X. Distribution des valeurs moyennes de I_b/I_0 et de leurs déviations standard en fonction de la masse molaire (a), ou du volume hydrodynamique (b), de l'acide aminé X. Les expériences ont été réalisées dans une solution de KCl 4M, 25mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

2.5 Simulation

Des expériences de simulation de la dynamique moléculaire ont été réalisées dans cette étude, dans le cadre de notre collaboration avec le groupe du Pr. Aksimentiev de l'Université de l'Illinois Urbana-Champaign (USA).

Ces expériences de simulation ont consisté, d'une part, à déterminer le mécanisme moléculaire à l'origine de la détection de la variation de l'acide aminé X au sein du peptide X_{R7} et, d'autre part, à estimer, en établissant un modèle théorique, le potentiel du nanopore d'aérolysine à discriminer les acides aminés les plus difficiles à distinguer. Ce deuxième point sera décrit par la suite.

Pour cela, un modèle tout atome du dispositif expérimental a été réalisé en simulant l'insertion d'un nanopore d'aérolysine dans une membrane de lipides (**Fig. 68, a**). La reconstitution du nanopore et sa stabilisation dans la bicouche lipidique ont été réalisées grâce à l'utilisation de la structure cristallographique de l'aérolysine obtenu très récemment à une résolution de 7,4 Angström (Å) par l'utilisation de la cryo-microscopie électronique et de la simulation [77].

Dans un premier temps, une carte de la distribution du potentiel électrostatique a été établie pour l'aérolysine (**Fig. 68, b-c**). Cette représentation montre qu'il existe deux barrières d'énergie, dont l'une est localisée à l'entrée du canal du côté *Cis*, et l'autre, à la sortie du canal du côté *Trans*. Entre ces deux barrières d'énergie, le potentiel est homogène le long du canal interne. L'interprétation de ce résultat est qu'à une basse tension comme celle de -50 mV à laquelle sont réalisées nos expériences, le peptide transporté traverse une première barrière d'énergie, puis se retrouve coincé à l'intérieur du pore entre cette première barrière d'énergie et celle localisée à la sortie du pore. C'est ce qui serait à l'origine des durées de blocage du courant par la présence du peptide dans le nanopore que l'on mesure qui sont de l'ordre de la dizaine de millisecondes et qui sont assez long pour permettre l'obtention de blocages de courant d'amplitudes précises à l'origine de la distinction des acides aminés entre eux. Cette hypothèse est confortée par une étude de la dépendance de la discrimination à la tension électrique appliquée (**Fig. 69**) qui montre que plus l'amplitude de la tension électrique appliquée est importante, plus la séparation des populations est moindre jusqu'à être totalement perdue. Cela signifie qu'à des plus grandes amplitudes de tensions électriques (valeurs de tensions inférieures à -70 mV), les peptides traversent le nanopore en franchissant d'abord la première barrière d'énergie, puis visitent le pore mais pour une durée qui n'est significativement pas assez grande pour permettre d'avoir une signature électrique spécifique, et enfin traverse la

deuxième barrière d'énergie pour sortir du nanopore. Cela est cohérent car on voit que l'augmentation de l'amplitude de la tension électrique, en plus de la perte de discrimination, s'accompagne d'une diminution des durées des blocages du courant.

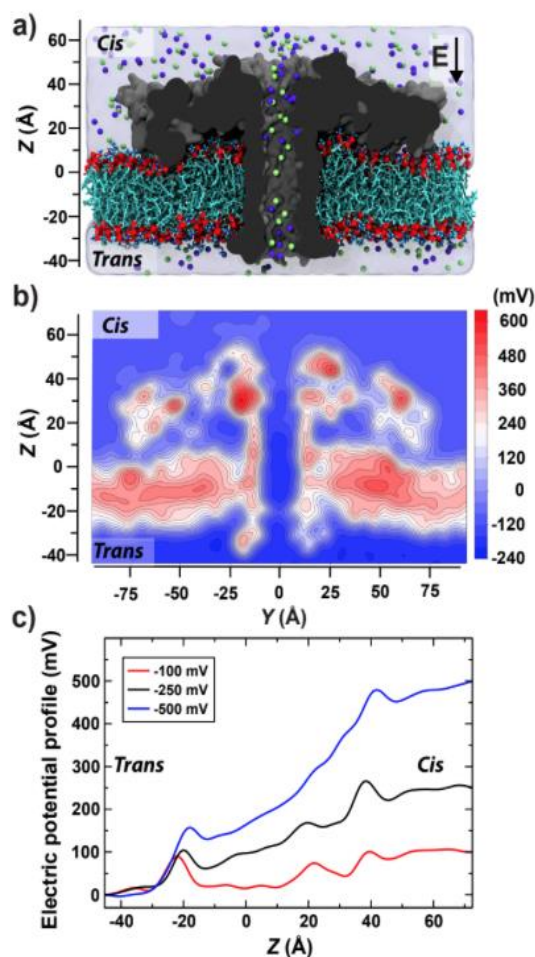


Figure 68 - Expériences de simulation dans l'aérolysine. Dispositif de simulation consistant en l'insertion d'un nanopore d'aérolysine dans une membrane de lipides, dans les mêmes conditions expérimentales que les expériences réalisées (a). Carte de la distribution du potentiel électrostatique de l'aérolysine, obtenue à -100 mV (b). Potentiel électrostatique moyen mesuré le long de l'axe z de l'aérolysine (c).

Par la suite, une simulation du transport d'un peptide X_{R7} a été réalisée par l'application d'une force mécanique permettant d'entraîner le peptide dans le pore à une vitesse constante. La résistance du peptide à cette force le long du canal interne de l'aérolysine a été mesurée (**Fig. 67, d**). Cette simulation révèle, de façon homogène à la répartition du potentiel électrostatique du pore décrite juste avant, que l'avancée du peptide présente le plus de résistance au niveau de l'entrée et de la sortie du canal où se trouvent les barrières d'énergie. Au contraire, c'est entre ces deux segments, à l'intérieur du canal, que le peptide présente le moins de résistance, ce qui

suggère que le peptide est piégé dans cette partie du pore et diffuse aisément à l'intérieur du canal.

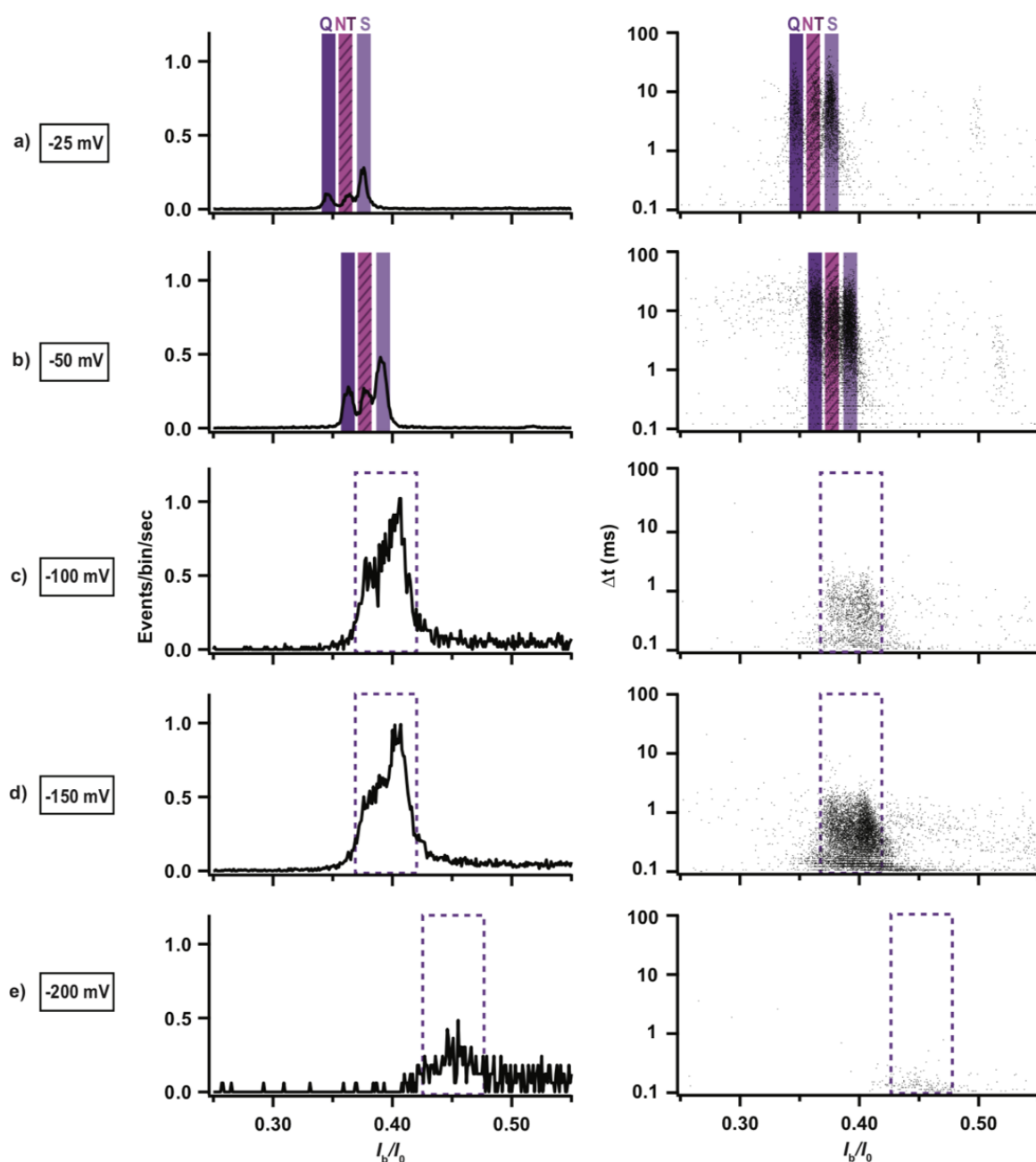


Figure 69 - Influence de la tension appliquée sur la discrimination des peptides X_{R7} . Histogramme de I_b/I_0 (gauche) et distribution des durées de blocage du courant en fonction de I_b/I_0 (nuage de points, droite) obtenus pour l'analyse d'un mélange équimolaire des peptides Q_{R7} , N_{R7} , T_{R7} et S_{R7} sous l'action d'une tension électrique de -25 mV (a), -50 mV (b), -100 mV (c), -150 mV (d) ou -200 mV (e), appliquée en trans. L'augmentation de la tension électrique est à l'origine d'une diminution puis d'une perte de la séparation des populations, d'un décalage des valeurs de I_b/I_0 , et d'une diminution des durées de blocage. Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C.

Pour ensuite déterminer quelle région du canal est plus précisément responsable de la détection des variations de volumes entre les peptides X_{R7} , une mesure de la distribution du courant relatif I_b/I_0 a été réalisée tout le long du canal. Pour cela, les valeurs de I_b/I_0 des homopeptides d'arginine précédemment analysés (R_{10-4}) ont été utilisées [312]. Le transport de ces peptides a été simulé le long du nanopore pour déterminer au niveau de quelle zone du canal interne sont retrouvées ces valeurs de I_b/I_0 (**Fig. 70, f**). Cette expérience montre que la valeur de I_b/I_0 est d'abord faible au niveau des portes d'entrée et de sortie du canal, et que les plus grandes valeurs de I_b/I_0 sont obtenues à l'intérieur du canal, de façon cohérente aux hypothèses précédentes. Les valeurs de I_b/I_0 obtenues lors de la simulation du transport des homopeptides sont les plus proches des valeurs expérimentales lorsque le peptide se retrouve à la position 22 Å sur l'axe du canal.

C'est donc précisément au niveau de cette zone que le nanopore d'aérollysine serait le plus sensible aux variations de volume entre les acides aminés, à l'origine de la discrimination des peptides X_{R7} . Pour confirmer cela, une expérience de simulation du transport de peptides X_{R7} a également été réalisée (**Fig. 70, h**). Les valeurs expérimentales de I_b/I_0 correspondant aux peptides dont le transport a été simulé, sont retrouvées lorsque les peptides se trouvent au niveau de la position 22 Å du segment du canal. Cela confirme donc que c'est cette partie du canal interne du nanopore d'aérollysine qui est la zone de sensibilité aux variations de volume.

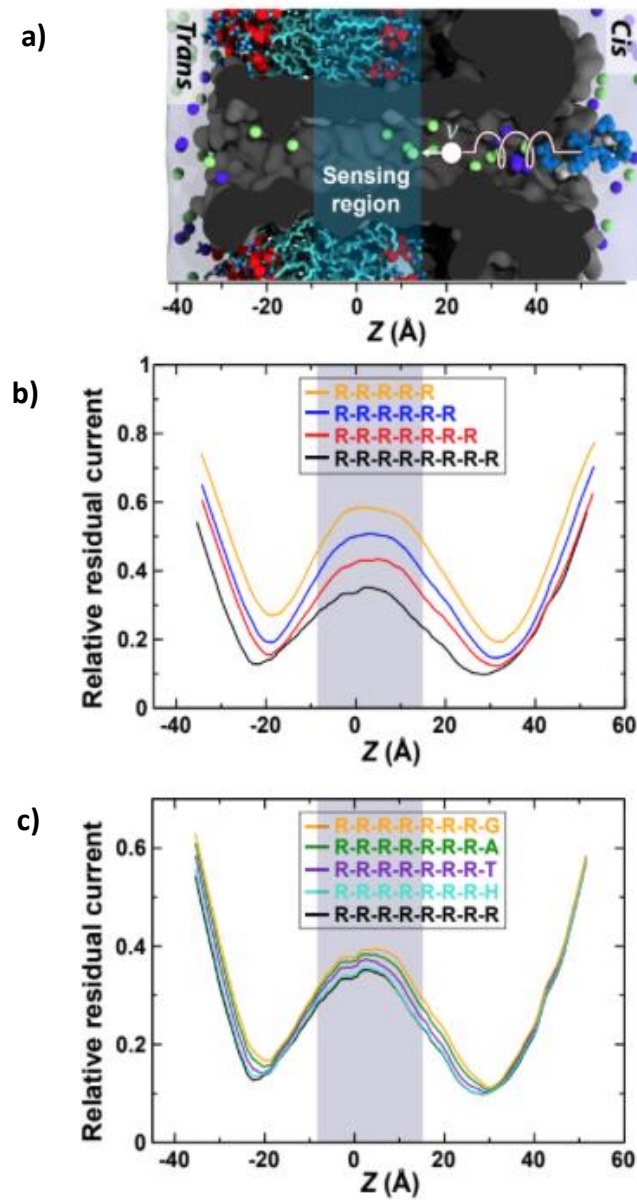


Figure 70 – Simulation du transport de peptides dans le pore d'aérollysine. (a) Schéma de la simulation de dynamique moléculaire du transport d'un peptide R_{R7} dans le pore d'aérollysine par l'action d'une force constante. La région de l'aérollysine identifiée comme sensible au volume est montrée en bleu. Distribution des courants relatifs moyens I_b/I_0 selon l'axe z de l'aérollysine dans le cas des homopeptides R_{R7} , R_{R6} , R_{R5} et R_{R4} (b), et dans le cas de cinq peptides X_{R7} (H_{R7} , T_{R7} , A_{R7} , G_{R7} , R_{R7}) (c).

2.6 Pistes d'améliorations pour une stratégie de séquençage

Dans le cadre de la révision de cet article, nous avons été amené à mettre en place des stratégies d'amélioration de nos résultats pour démontrer le potentiel de la méthode proposée pour le séquençage de protéines

2.6.1 Augmentation des temps de résidence

Dans un premier temps, une première stratégie proposée est d'allonger les temps de résidence des peptides dans le nanopore de façon à augmenter la précision d'identification des acides aminés. En effet, des blocages du courant de durées plus longues permettrait de réduire la dispersion autour de la moyenne et d'obtenir des valeurs de I_b/I_0 moyennes plus précises. Pour montrer que cette idée est intéressante, un modèle a été construit assumant qu'une mutation génétique réalisée à l'extrémité *trans* du nanopore d'aérolysine permettrait d'allonger les temps de résidence des peptides dans le pore, et cela sans modifier, ni le taux de capture des peptides dans le pore, ni le courant ionique circulant à travers le nanopore, ni les courants relatifs résiduels I_b/I_0 mesurés pour les vingt peptides X_{R7} . Ce modèle a permis de montrer qu'une augmentation des temps de résidence des peptides dans le pore jusqu'à 20 ms permettrait de distinguer 6 des 20 peptides analysés avec une précision de plus de 90 % (**Fig. 71**). Augmenter d'avantage les temps de résidence jusqu'à 200 ms serait bien plus intéressant et permettrait de distinguer 16 des 20 peptides analysés avec une précision de 90 % ou plus (**Fig. 71**).

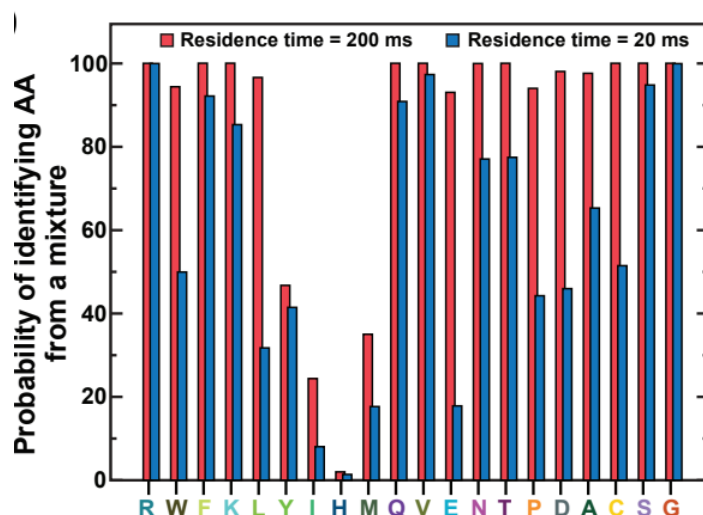
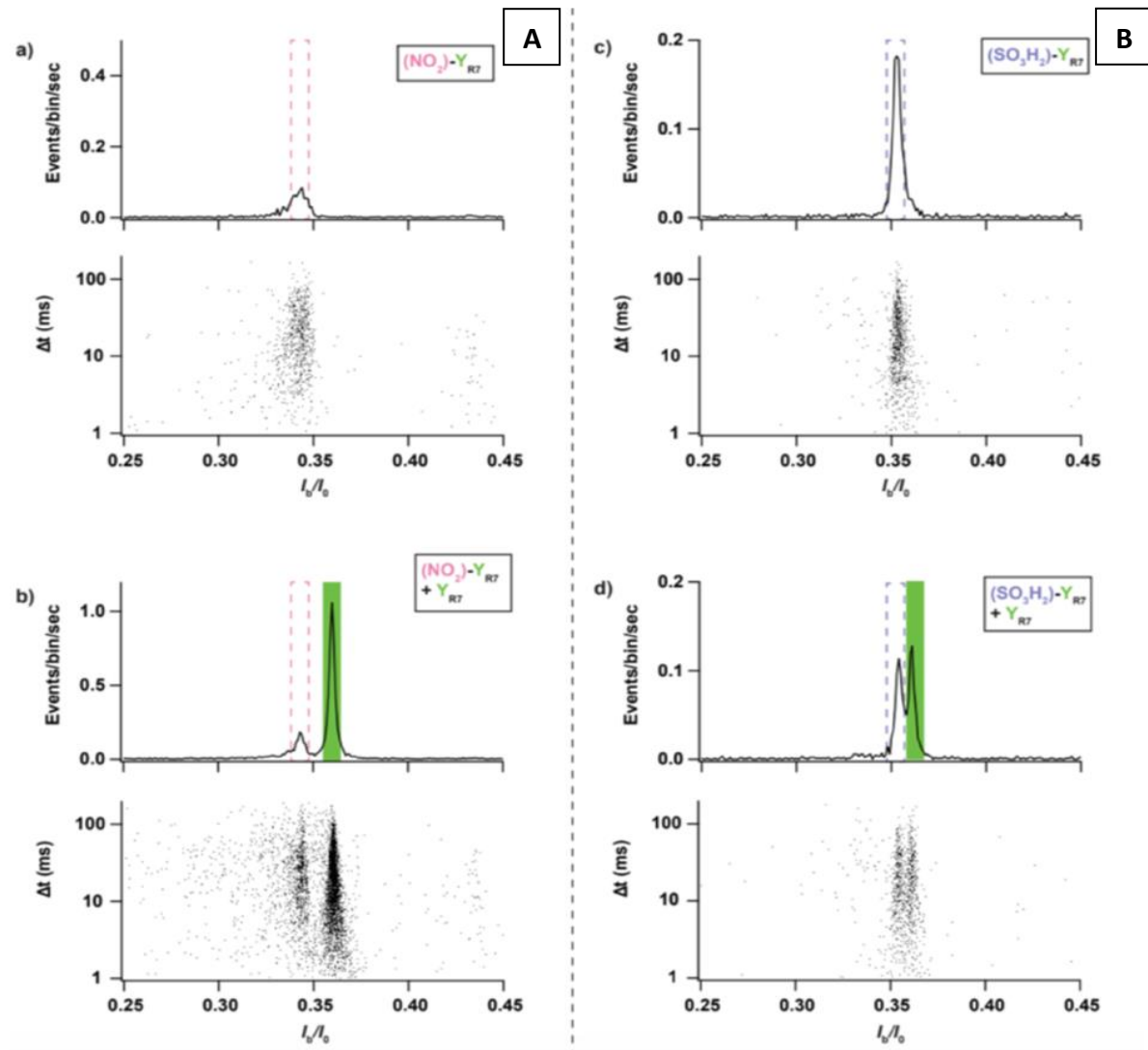


Figure 71 - Effet du temps de résidence sur la discrimination des peptides X_{R7} . Probabilité théorique de distinguer un peptide individuel X_{R7} dans un mélange des vingt peptides X_{R7} à partir d'un événement de blocage du courant unique d'une durée de 20 (bleu) ou 200 (rouge) ms.

2.6.2 Introduction de modifications chimiques

Dans un second temps, nous avons choisi d'introduire des modifications chimiques au niveau de l'acide aminé X de certains peptides X_{R7} qui sont plus particulièrement difficiles à distinguer à partir de leur valeur moyenne de I_b/I_0 . L'idée en réalisant cela est de moduler les valeurs du courant relatif I_b/I_0 pour les extraire de l'intervalle de $I_b/I_0 \approx 0,3-0,4$ dans lequel l'ensemble des vingt peptides sont regroupés. Pour cela, deux peptides dont les valeurs sont trop proches pour permettre leur distinction, T_{R7} et M_{R7} , ont été modifiés chimiquement. Le peptide Y_{R7} a subi une modification par l'ajout d'une fonction $(-NO_2)$, ou par l'ajout d'une phosphorylation $(-P)$, sur l'acide aminé Y. Le peptide M_{R7} a subi une modification par l'ajout d'un groupement chimique sulfoxide sur l'acide aminé M. Ces expériences ont montré que l'introduction des modifications chimiques ont permis de modifier les signatures électriques des acides aminés de telle sorte à les distinguer des autres peptides, mais également de distinguer ces modifications chimiques entre elles (**Fig. 72 et 73**).



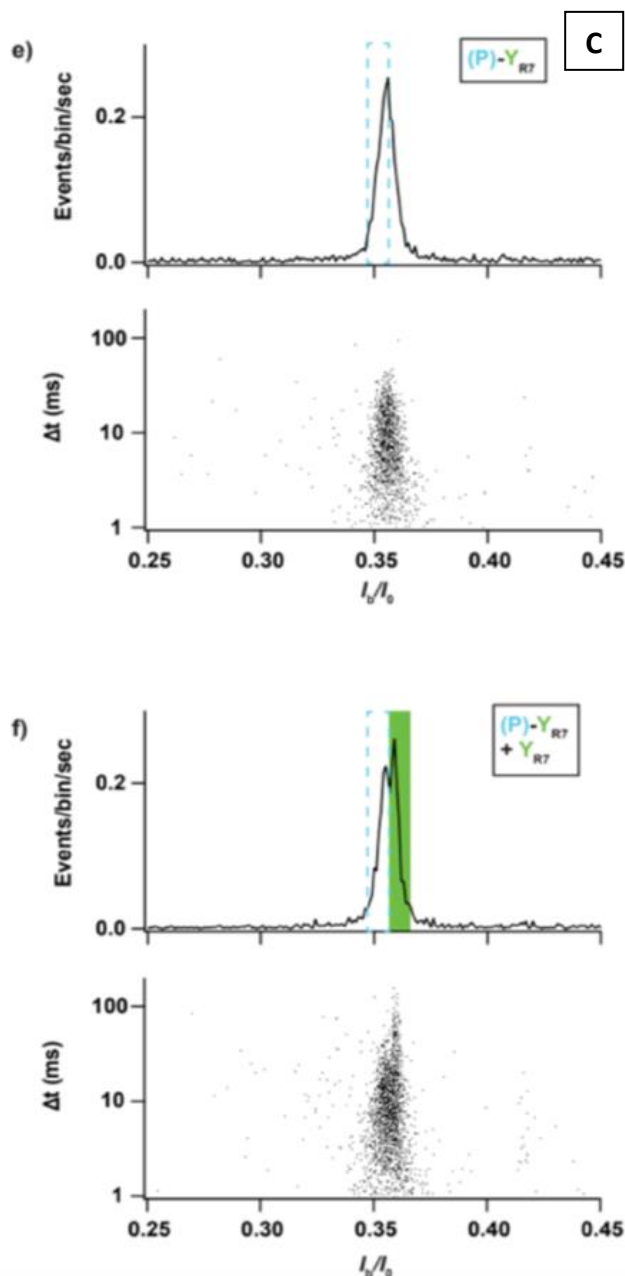


Figure 72 - Identification de peptides présentant des modifications chimiques avec le nanopore d'aérolysine. Histogrammes de I_b/I_0 et distribution des durées de blocages en fonction de I_b/I_0 (nuage de points) obtenus dans le cas de l'analyse avec l'aérolysine : (A) d'un mélange équimolaire des peptides Y_{R7} et (NO₂)-Y_{R7} et qui ont été ajoutés successivement en solution en commençant par le peptide (NO₂)-Y_{R7} (a), puis Y_{R7} (b) ; (B) d'un mélange équimolaire des peptides Y_{R7} et (SO₃H₂)-Y_{R7} et qui ont été ajoutés successivement en solution en commençant par le peptide (SO₃H₂)-Y_{R7} (c), puis Y_{R7} (d) ; (C) d'un mélange équimolaire des peptides Y_{R7} et (NO₂)-Y_{R7} et qui ont été ajoutés successivement en solution en commençant par le peptide (P)-Y_{R7} (e), puis Y_{R7} (f). Dans chaque cas, la modification chimique est à l'origine d'un décalage de la valeur de I_b/I_0 d'origine. Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

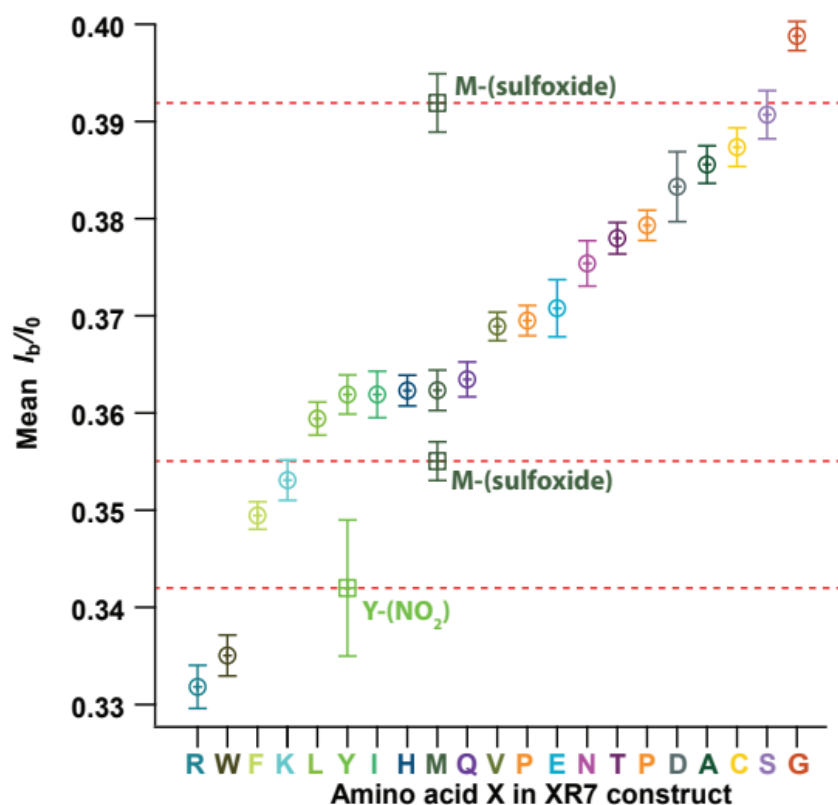


Figure 73 - Vue d'ensemble de la discrimination des acides aminés protéinogènes et modifiés chimiquement. Distribution des valeurs moyennes de I_b/I_0 et leurs déviations standards obtenues dans le cas de l'analyse individuelle des peptides X_{R7} et des peptides X_{R7} modifiés chimiquement : (NO_2) - Y_{R7} et (Met) - M_{R7} . Les expériences ont été réalisées dans une solution de KCl 4 M, 25 mM HEPES, pH = 7,5, à 20°C, et sous l'action d'une tension électrique de -50 mV appliquée en trans.

2.6.3 Variation des conditions expérimentales et d'enregistrement électriques

Enfin, une dernière stratégie a consisté à explorer les conditions expérimentales et à moduler les paramètres électroniques du dispositif expérimental de telle façon à améliorer la résolution. C'est dans le cadre d'une collaboration avec le groupe du Pr. Jan C. Behrends de l'Institut de Physiologie de Freiburg en Allemagne que ces expériences ont pu être réalisées. Les résultats obtenus montrent qu'il est possible d'améliorer la séparation de certains peptides comme L_{R7} et I_{R7} , en utilisant un dispositif à haute résolution permettant de réaliser des enregistrements du signal électrique avec une fréquence d'échantillonnage de 1 μ s et avec un bas bruit électrique, permet de reproduire les résultats obtenus avec le dispositif expérimental que l'on a utilisé pour l'ensemble des expériences présentées mais également, et en combinant ce dispositif à de nouvelles conditions expérimentales en électrolyte.

En effet, un mélange de sels de bromure de potassium (KNO_3^-) et de chlorure de potassium (KCl), selon un ratio 1:1, a été utilisé comme électrolyte. Les expériences réalisées en présence de cet électrolyte, ont permis d'améliorer la séparation des peptides L_{R7} et I_{R7} en abaissant leur taux d'erreur d'identification à moins de 5%, alors qu'il était de plus de 30%,. L'interprétation de ce phénomène serait liée à l'interaction électrostatique des ions NO_3^- avec les peptides X_{R7} , ce qui ralentirait leur dynamique et induirait en conséquence une diminution du bruit électrique, et serait à l'origine d'une meilleure précision de la mesure (**Fig. 74**)

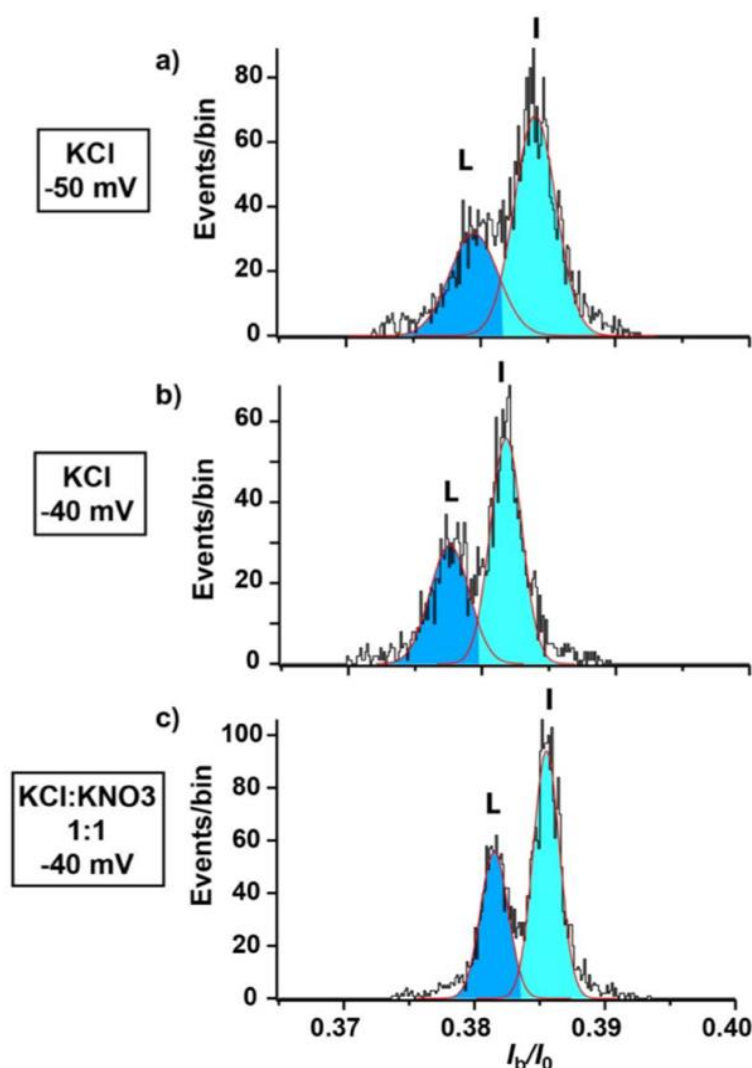


Figure 74 - Discrimination des peptides L_{R7} et I_{R7} dans des conditions expérimentales d'enregistrement optimisées. Histogramme de I_{b}/I_0 dans le cas de l'analyse avec le nanopore d'aérolysine d'un mélange équimolaire des peptides L_{R7} et I_{R7} ajouté dans le compartiment cis en présence : (a) d'une solution ionique de KCl à une concentration de 4 M, sous l'action d'une tension de -50 mV ou de (b) -40 mV ; ou (c) d'une solution ionique de 2 M KCl et 2 M KNO_3^- (ratio 1:1), sous l'action d'une tension électrique de -40 mV ; à chaque fois appliquée en trans.

III. Discussion & Conclusion

Nous avons donc démontré la possibilité de détecter et de caractériser les 20 acides aminés protéinogènes par des signatures électriques spécifiques obtenues avec le nanopore d'aérolysine sauvage. Parmi les 20 acides aminés analysés, l'approche que nous avons mis en place nous a permis d'identifier au moins 13 acides aminés dont R, W, F, K, L, N, T, P, D, A, C, S, et G, sans y introduire de modifications chimiques ; et d'étendre l'identification aux acides aminés Y, et M, en les modifiant chimiquement. Ce nanopore permet également de distinguer des différences aussi subtiles qu'une fonction hydroxyle ou que des isomères de structures de même masse molaire.

Nous avons également proposé des pistes expérimentales et théoriques d'amélioration de nos résultats, comme l'utilisation d'un mutant d'aérolysine, l'introduction de modifications chimiques dans les peptides, ou encore la variation des conditions expérimentales et du dispositif de mesure électrique. Les résultats obtenus confortent l'idée que notre stratégie pourrait être utilisée pour le séquençage de protéines par nanopore.

Les résultats obtenus sont un progrès supplémentaire par rapport aux autres techniques à l'échelle de la molécule unique, comme la détection par fingerprinting [16,18], par fluoroséquençage [17,19], ou par effet tunnel [20] ; et ouvre la possibilité vers le développement d'une méthode de séquençage de protéines par nanopore, résidus par résidus, à l'échelle de la molécule unique.

Les travaux de simulation ont permis d'identifier l'origine de la sensibilité atteinte avec cette approche qui a été attribuée au nanopore biologique que l'on utilise, qui agit comme une véritable trappe qui confine l'acide aminé lié à son amorce d'arginine polycationique dans la zone de sensibilité du nanopore.

Les durées de résidence allongées des peptides dans le nanopore d'aérolysine constituent l'originalité de nos travaux, en comparaison à d'autres approches comme celle qui consiste à tirer sur une protéine ou un peptide pour entraîner son dépliement, exposer ses domaines, et lire sa séquence durant sa translocation à travers le nanopore pendant un temps trop court pour être suffisant à la lecture de la séquence avec une précision d'un acide aminé.

Comme perspectives de ce travail, nous envisageons à plus long terme de proposer une stratégie de séquençage dans laquelle les acides aminés d'une protéine seraient clivés un par un, puis

seraient attachés à un homopeptide servant de locomotive, et seraient enfin analysés à travers le nanopore d'aérolysine. Ces étapes de réaction seraient réalisées en parallèle dans des volumes indépendants, isolés dans des compartiments différents. L'étape de clivage pourrait être réalisée à l'aide d'une enzyme immobilisée près du nanopore d'aérolysine.

Au cours de la réalisation de ces travaux, d'autres études ont été publiées proposant également l'utilisation des nanopores biologiques pour l'analyse de peptides à la recherche de la meilleure résolution, dans la perspective du séquençage de protéines, de la détection et l'identification d'acides aminés et de modifications post-traductionnelles.

Parmi ces travaux, on peut citer les travaux de Huang *et al.* nanopore FraC, qui ont proposé de moduler le diamètre du nanopore FraC en y introduisant des modifications génétiques produisant des pores de 1,1, 1,6, et 0,84 nm de diamètre interne pouvant accueillir des peptides d'une longueur de 4 à 22 acides aminés [91]. Un diamètre interne encore plus étroit que celui du nanopore FraC serait nécessaire pour détecter des peptides de plus petites longueurs. Le meilleur résultat montré dans cette étude est la discrimination entre une alanine et un glutamate respectivement entre les peptides angiotensine A et II, en utilisant le nanopore de diamètre interne 1,1 nm (W116S-FraC) et à pH 4,5, à -30 mV ; ce qui correspond à une différence de masse de 44 Da. Les auteurs ont voulu aller plus loin en montrant qu'il est possible d'identifier la masse de peptides à partir du courant ionique. Ils ont démontré l'existence d'un lien entre la masse de l'analyte et le courant ionique mesuré mais seulement à un pH de 3,8. Les résultats obtenus dans cette étude sont dépendants de conditions en pH très précises qui demeurent difficiles à contrôler dans une expérience en routine.

Les mêmes auteurs ont également récemment démontré la possibilité de détecter des molécules de glucose et d'asparagine issues de fluides biologiques grâce à l'utilisation de protéines de reconnaissance, spécifique de ces molécules ajoutées en solution en présence du nanopore ClyA [313]. Les mêmes auteurs ont étendu leurs études à l'utilisation de plusieurs autres protéines substrat de reconnaissance et ont montré que celles-ci peuvent être utilisées comme de vrais adaptateurs moléculaires avec le nanopore ClyA pour la détection d'une multitude de ligands, dont la leucine, l'asparagine, et la glutamine [314]. Ces travaux ont un intérêt plus particulier pour les analyses en métabolomique.

Le groupe de Y. T. Long a très récemment publié une étude montrant la détection d'acides aminés individuels de cystéine, sans marquage, avec le nanopore d'aérolysine, par des blocages du courant de l'ordre d'une dizaine de millisecondes atteints pour des hautes tensions de + 120

mV ou plus, dans des conditions en sels de KCl à une concentration de 1 M et à un pH d'environ 7,5-8. La même analyse de glutamine et d'asparagine n'a pas permis de détecter ces derniers acides aminés. La particularité de la détection de la Cystéine plus particulièrement a été attribué au fait que cet acide aminé peut créer un grand nombre de liaisons hydrogènes avec le nanopore d'aérolysine [315].

Enfin, plusieurs études ont été dédiées à l'analyse de modifications post-traductionnelles par un nanopore biologique. A ce titre, on peut citer les travaux de mesure en temps réel de la phosphorylation de peptides avec un nanopore d'aérolysine sauvage en présence d'une protéine kinase [294,316], de la détection de glycosylation et de phosphorylation d'hétéropeptides avec le nanopore FraC [317], ou encore de l'ajout de modifications chimiques comme des fluorophores (*e.g* Fluorescéine, Alexa 633, Texas Red) sur les mêmes peptides également avec le nanopore FraC [318].

Conclusion générale

En conclusion, nous avons présenté dans ce manuscrit les travaux de cette thèse portant sur le développement d'une stratégie de séquençage de protéines à l'échelle de la molécule unique basée sur l'utilisation d'un nanopore biologique recombinant sauvage, l'aérolysine.

La technologie de détection électrique à l'échelle de la molécule unique, basée sur l'utilisation de nanopores de différentes géométries, qu'ils soient biologiques, solides, ou hybrides, est utilisée depuis de nombreuses années pour l'analyse d'une large gamme de molécules comme les ions, les peptides, les protéines, les acides nucléiques, ou encore les polymères synthétiques. Une des démonstrations les plus spectaculaires ayant été réalisées avec la technologie des nanopores, concerne le séquençage de l'ADN à l'échelle de la molécule unique qui consiste à lire la séquence d'un simple brin d'ADN déplié à travers un nanopore au cours de sa translocation. Cependant, la même approche de séquençage par nanopore n'est pas parvenue à être transposée aux protéines en raison de leur complexité structurale et chimique.

L'objectif des travaux de thèse présentés dans ce manuscrit était de définir une stratégie de séquençage des protéines basée sur l'utilisation d'un nanopore stable et facilement modulable permettant d'atteindre une résolution de détection d'un acide aminé.

Pour cela, nous avons choisi l'aérolysine, une toxine bactérienne qui a la capacité de former des nanopores dans les membranes biologiques et qui a été utilisée dans de nombreuses études pour l'analyse à l'échelle de la molécule unique de peptides, de protéines, de polymères synthétiques, et de nucléotides. L'intérêt que nous avons porté à ce nanopore provient du fait qu'il a récemment été utilisé pour montrer la discrimination de polymères synthétiques de différentes longueurs, d'oligonucléotides de différentes longueurs, avec une résolution d'une unité monomérique ; mais également pour montrer la détection de la variation de la nature chimique d'une unité monomérique au sein d'un oligonucléotide.

En regard de cela, nous avons dans un premier temps utilisé le nanopore d'aérolysine pour analyser des homopeptides de différentes longueurs dans des conditions expérimentales identiques à celles ayant permis d'obtenir la séparation des polymères synthétiques et des oligonucléotides avec le même pore. Nous avons démontré que l'aérolysine, à l'état sauvage, permet la détection et la discrimination avec une résolution d'une unité monomérique de courts peptides d'arginine et de lysine d'une longueur de 5 à 10 acides aminés dans des conditions expérimentales à haute concentration en sel (KCl) et sous l'action d'une large gamme de

tensions électriques négatives. Nous avons également démontré que la détection peut être étendue à des homopeptides d'arginine d'une plus courte longueur de 3 et 4 acides aminés et d'une grande longueur de 12 acides aminés. Nous avons également démontré que des hétéropeptides de différentes séquences sont discriminables par bloc de 5 acides aminés. Enfin, nous avons montré dans ce travail que la sensibilité de ce nanopore nous permet de détecter des impuretés peptidiques présentes dans les échantillons analysés et comparable à une centaine de molécules. Ces impuretés avaient d'abord été détectées comme des populations secondaires présentes dans les analyses d'un homopeptide d'arginine synthétique d'une longueur de 9 acides aminés. Nous avons caractérisé ces impuretés comme étant des peptides de plus courtes longueurs que le peptide présent majoritairement en solution, en réalisant une analyse individuelle de chaque longueur et en comparant les valeurs d'amplitudes qui sont identiques.

Ces premiers résultats nous ont inspirés une stratégie pour identifier et distinguer les 20 acides aminés protéinogènes avec l'aérolysine.

Nous avons montré dans un deuxième temps, en utilisant une chaîne d'arginine d'une longueur de 7 acides aminés liée à un huitième acide aminé dont la nature chimique varie parmi les 20 acides aminés protéinogènes, que l'aérolysine à l'état sauvage, permet de détecter chacun des peptides et fournit un signal électrique caractéristique de chacun des 20 acides aminés. Dans ces résultats, nous avons démontré que l'analyse de cette construction avec ce nanopore nous permet de discriminer au moins 13 des 20 acides aminés protéinogènes, alors que d'autres présentent des signaux trop proches pour être facilement distinguables. Parmi ces acides aminés, nous avons plus particulièrement montré que l'aérolysine permet de distinguer, sur la base de l'amplitude de blocage du courant, des modifications chimiques subtiles telles que la différence d'un groupement hydroxyle entre les peptides portant une phénylalanine (F) et une tyrosine (Y), et telles que des isomères de structures de même masse moléculaire comme la leucine (L) et l'isoleucine (I). Ce résultat montre, d'une part, la haute sensibilité de ce nanopore pour la détection de différences aussi subtiles, ce qui peut se révéler être difficile avec des techniques plus conventionnelles ; d'autre part, ouvre une perspective vers la possibilité de l'utiliser pour la détection de modifications chimiques post-traductionnelles ; et enfin, ouvre le débat concernant la discrimination des molécules par l'aérolysine dans des conditions à haute concentration en sel qui était principalement attribuée à la différence de masse de ces molécules et avait donné naissance au concept de la spectrométrie de masse par nanopore qui est discuté aujourd'hui. En effet, nous avons démontré que la discrimination des acides aminés dans nos travaux est davantage liée à des différences de volumes qu'à des différences de masses.

Pour étendre la discrimination à plus d'acides aminés, nous avons introduit des modifications chimiques sur certains acides aminés difficiles à distinguer afin d'évaluer leur signal en présence des modifications chimiques. En réalisant cela, nous sommes parvenus à étendre la discrimination de 13 à 15 des 20 acides aminés protéinogènes. Nous avons également confirmé le potentiel du l'aérolysine pour la détection de modifications post-traductionnelles. Une autre approche d'amélioration de nos résultats a consisté à explorer de nouvelles conditions expérimentales favorisant la discrimination des peptides et l'utilisation de système d'acquisition plus performant. Nous avons démontré par cette approche qu'il est possible de réduire le taux d'erreur d'identification de deux acides aminés comme la leucine et l'isoleucine ; en modifiant les conditions expérimentales en électrolyte, et en utilisant un système d'acquisition du signal plus performant ; l'ensemble nous permettant de réduire le bruit électrique induit par la présence des peptides dans le pore.

Les travaux de simulation réalisés ont mis en lumière le mécanisme à l'origine de l'obtention d'une telle résolution de séparation avec le nanopore d'aérolysine. Les résultats obtenus ont été attribués aux propriétés de géométrie et de charge du pore d'aérolysine. En effet, l'intérieur du pore constitue un volume délimité par deux barrières d'énergie, localisées au niveau des deux constriction du nanopore, dans lequel les acides aminés liés à leur amorce d'arginine sont confinés pendant des durées de l'ordre de la dizaine de milliseconde.

Comme perspectives de ces travaux, nous avons proposé une stratégie de séquençage basée sur cette approche qui consisterait à séquencer des protéines résidus par résidus. Le principe serait de cliver les acides aminés d'une protéine ou d'un peptide d'intérêt, un par un, dans un premier compartiment, à l'aide de l'assistance d'une enzyme. L'acide aminé clivé serait ensuite lié dans un deuxième compartiment, à un homopeptide d'arginine constituant l'amorce peptidique, nécessaire à l'entrée de l'acide aminé dans le nanopore d'aérolysine inséré dans une membrane de lipides dans un troisième compartiment.

Dans la continuité de ces travaux de thèse, ces différentes étapes doivent être explorées pour démontrer de la faisabilité de l'approche de séquençage que nous proposons. Une des limitations qui concerne l'étape de ligation de l'acide aminé à l'amorce peptidique pourrait trouver comme alternative l'utilisation d'une molécule hôte capable de former un complexe avec les acides aminés qui seraient détectés dans le nanopore.

Dans ce sens, nous avons entamé au cours de ma thèse, des travaux portant sur l'analyse de molécules cavitaires de cucurbiturils qui sont connus pour présenter une forte affinité de liaison

avec les acides aminés, en comparaison à leurs homologues, les cyclodextrines. Nous avons déjà montré que des molécules de cucurbiturils de différents diamètres sont détectées avec le nanopore d'aérolysine, et dans différentes conditions expérimentales [319]. La prochaine étape serait de montrer la détection d'un complexe formé entre une molécule de cucurbituril et un acide aminé avec le nanopore d'aérolysine. Au-delà de l'intérêt de cette étude pour le séquençage de protéines, cette approche pourrait être utilisée pour détecter des acides aminés libre présents dans des fluides biologiques dans le contexte d'analyses métabolomiques.

Pour finir, la technologie des nanopores est en constant progrès pour l'analyse d'une large gamme d'analytes, et cela grâce à la possibilité d'utiliser des nanopores aux configurations et aux diamètres multiples comme des nanopores biologiques modifiés génétiquement, chimiquement, et/ou associés à des moteurs moléculaires, et/ou inséré dans des membranes solides ; des nanopores solides percés dans des matériaux de différentes natures qui peuvent être fonctionnalisés pour améliorer leur spécificité ; ou encore des nanopores hybrides combinant les avantages des nanopores biologiques.

Références

- [1] J.J. Kasianowicz, E. Brandin, D. Branton, D.W. Deamer, Characterization of individual polynucleotide molecules using a membrane channel, *Proceedings of the National Academy of Sciences*. 93 (1996) 13770–13773. <https://doi.org/10.1073/pnas.93.24.13770>.
- [2] H. Bayley, C.R. Martin, Resistive-Pulse Sensing From Microbes to Molecules, *Chem. Rev.* 100 (2000) 2575–2594. <https://doi.org/10.1021/cr980099g>.
- [3] M. Wanunu, Nanopores: A journey towards DNA sequencing, *Physics of Life Reviews*. 9 (2012) 125–158. <https://doi.org/10.1016/j.plrev.2012.05.010>.
- [4] J.E. Reiner, A. Balijepalli, J.W.F. Robertson, J. Campbell, J. Suehle, J.J. Kasianowicz, Disease Detection and Management via Single Nanopore-Based Sensors, *Chem. Rev.* 112 (2012) 6431–6451. <https://doi.org/10.1021/cr300381m>.
- [5] J.W.F. Robertson, J.E. Reiner, The Utility of Nanopore Technology for Protein and Peptide Sensing, *Proteomics*. 18 (2018) 1800026. <https://doi.org/10.1002/pmic.201800026>.
- [6] L. Restrepo-Pérez, C. Joo, C. Dekker, Paving the way to single-molecule protein sequencing, *Nature Nanotech.* 13 (2018) 786–796. <https://doi.org/10.1038/s41565-018-0236-6>.
- [7] N. Varongchayakul, J. Song, A. Meller, M.W. Grinstaff, Single-molecule protein sensing in a nanopore: a tutorial, *Chem. Soc. Rev.* 47 (2018) 8512–8524. <https://doi.org/10.1039/C8CS00106E>.
- [8] A. Asandei, G. Di Muccio, I. Schiopu, L. Mereuta, I.S. Dragomir, M. Chinappi, T. Luchian, Nanopore-Based Protein Sequencing Using Biopores: Current Achievements and Open Challenges, *Small Methods*. (2020) 1900595. <https://doi.org/10.1002/smtd.201900595>.
- [9] W. Timp, G. Timp, Beyond mass spectrometry, the next step in proteomics, *Sci. Adv.* 6 (2020) eaax8978. <https://doi.org/10.1126/sciadv.aax8978>.
- [10] B. Cressiot, L. Bacri, J. Pelta, The Promise of Nanopore Technology : Advances in the Discrimination of Protein Sequences and Chemical Modifications, *Small Methods*. 4 (2020) 2000090. <https://doi.org/10.1002/smtd.202000090>.
- [11] J.R. Yates Iii, A century of mass spectrometry: from atoms to proteomes, *Nature Methods*. 8 (2011) 633–637. <https://doi.org/10.1038/nmeth.1659>.
- [12] H. Steen, M. Mann, The abc's (and xyz's) of peptide sequencing, *Nat Rev Mol Cell Biol.* 5 (2004) 699–711. <https://doi.org/10.1038/nrm1468>.
- [13] J.R. Yates, III, Recent technical advances in proteomics, *F1000Res.* 8 (2019) 351. <https://doi.org/10.12688/f1000research.16987.1>.

- [14] N. Callahan, J. Tullman, Z. Kelman, J. Marino, Strategies for Development of a Next-Generation Protein Sequencing Platform, *Trends in Biochemical Sciences*. 45 (2020) 76–89. <https://doi.org/10.1016/j.tibs.2019.09.005>.
- [15] S. Schmid, C. Dekker, Nanopores -- a Versatile Tool to Study Protein Dynamics, *ArXiv:2010.07399 [q-Bio]*. (2020). <http://arxiv.org/abs/2010.07399> (accessed December 14, 2020).
- [16] Y. Yao, M. Docter, J. van Ginkel, D. de Ridder, C. Joo, Single-molecule protein sequencing through fingerprinting: computational assessment, *Phys. Biol.* 12 (2015) 055003. <https://doi.org/10.1088/1478-3975/12/5/055003>.
- [17] J. Swaminathan, A.A. Boulgakov, E.M. Marcotte, A Theoretical Justification for Single Molecule Peptide Sequencing, *PLoS Comput Biol.* 11 (2015) e1004080. <https://doi.org/10.1371/journal.pcbi.1004080>.
- [18] J. van Ginkel, M. Filius, M. Szczepaniak, P. Tulinski, A.S. Meyer, C. Joo, Single-molecule peptide fingerprinting, *Proc Natl Acad Sci USA*. 115 (2018) 3338–3343. <https://doi.org/10.1073/pnas.1707207115>.
- [19] J. Swaminathan, A.A. Boulgakov, E.T. Hernandez, A.M. Bardo, J.L. Bachman, J. Marotta, A.M. Johnson, E.V. Anslyn, E.M. Marcotte, Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures, *Nat Biotechnol.* 36 (2018) 1076–1082. <https://doi.org/10.1038/nbt.4278>.
- [20] Y. Zhao, B. Ashcroft, P. Zhang, H. Liu, S. Sen, W. Song, J. Im, B. Gyrfas, S. Manna, S. Biswas, C. Borges, S. Lindsay, Single-molecule spectroscopy of amino acids and peptides by recognition tunnelling, *Nature Nanotech.* 9 (2014) 466–473. <https://doi.org/10.1038/nnano.2014.54>.
- [21] Z. Dong, E. Kennedy, M. Hokmabadi, G. Timp, Discriminating Residue Substitutions in a Single Protein Molecule Using a Sub-nanopore, *ACS Nano*. 11 (2017) 5440–5452. <https://doi.org/10.1021/acsnano.6b08452>.
- [22] L. Restrepo-Pérez, S. John, A. Aksimentiev, C. Joo, C. Dekker, SDS-assisted protein transport through solid-state nanopores, *Nanoscale*. 9 (2017) 11685–11693. <https://doi.org/10.1039/C7NR02450A>.
- [23] M.W. Parker, F.G. van der Goot, J.T. Buckley, Aerolysin - the ins and outs of a model channel-forming toxin, *Mol Microbiol.* 19 (1996) 205–212. <https://doi.org/10.1046/j.1365-2958.1996.355887.x>.
- [24] M. Pastoriza-Gallego, L. Rabah, G. Gibrat, B. Thiebot, F.G. van der Goot, L. Auvray, J.-M. Betton, J. Pelta, Dynamics of Unfolded Protein Transport through an Aerolysin Pore, *J. Am. Chem. Soc.* 133 (2011) 2923–2931. <https://doi.org/10.1021/ja1073245>.
- [25] B. Cressiot, H. Ouldali, M. Pastoriza-Gallego, L. Bacri, F.G. Van der Goot, J. Pelta, Aerolysin, a Powerful Protein Sensor for Fundamental Studies and Development of Upcoming Applications, *ACS Sens.* 4 (2019) 530–548. <https://doi.org/10.1021/acssensors.8b01636>.

- [26] G. Baaken, I. Halimeh, L. Bacri, J. Pelta, A. Oukhaled, J.C. Behrends, High-Resolution Size-Discrimination of Single Nonionic Synthetic Polymers with a Highly Charged Biological Nanopore, *ACS Nano*. 9 (2015) 6443–6449. <https://doi.org/10.1021/acsnano.5b02096>.
- [27] C. Cao, Y.-L. Ying, Z.-L. Hu, D.-F. Liao, H. Tian, Y.-T. Long, Discrimination of oligonucleotides of different lengths with a wild-type aerolysin nanopore, *Nature Nanotech*. 11 (2016) 713–718. <https://doi.org/10.1038/nnano.2016.66>.
- [28] W.H. Coulter, Means for counting particles suspended in a fluid. US Patent, 2656508, United States Patent Office. Patentiert Am. 20 (1953) 1953.
- [29] Y. Song, J. Zhang, D. Li, Microfluidic and Nanofluidic Resistive Pulse Sensing: A Review, *Micromachines*. 8 (2017) 204. <https://doi.org/10.3390/mi8070204>.
- [30] G. Brecher, M. Schneiderman, G.Z. Williams, Evaluation of Electronic Red Blood Cell Counter, *Am J Clin Pathol*. 26 (1956) 1439–1449. <https://doi.org/10.1093/ajcp/26.12.1439>.
- [31] C.F.T. Mattern, F.S. Brackett, B.J. Olson, *Determination of Number and Size of Particles by Electrical Gating: Blood Cells*, *Journal of Applied Physiology*. 10 (1957) 56–70. <https://doi.org/10.1152/jappl.1957.10.1.56>.
- [32] B.S. Bull, M.A. Schneiderman, G. Brecher, Platelet Counts with the Coulter Counter, *Am J Clin Pathol*. 44 (1965) 678–688. <https://doi.org/10.1093/ajcp/44.6.678>.
- [33] R.M. Smeets, U.F. Keyser, D. Krapf, M.-Y. Wu, N.H. Dekker, C. Dekker, Salt dependence of ion transport and DNA translocation through solid-state nanopores, *Nano Letters*. 6 (2006) 89–95. <https://doi.org/10.1021/nl052107w>.
- [34] S.W. Kowalczyk, C. Dekker, Measurement of the Docking Time of a DNA Molecule onto a Solid-State Nanopore, *Nano Lett*. 12 (2012) 4159–4163. <https://doi.org/10.1021/nl301719a>.
- [35] J. Roman, N. Jarroux, G. Patriarche, O. Français, J. Pelta, B. Le Pioufle, L. Bacri, Functionalized Solid-State Nanopore Integrated in a Reusable Microfluidic Device for a Better Stability and Nanoparticle Detection, *ACS Appl. Mater. Interfaces*. 9 (2017) 41634–41640. <https://doi.org/10.1021/acsaami.7b14717>.
- [36] H.E. Kubitschek, Electronic counting and sizing of bacteria, *Nature*. 182 (1958) 234–235. <https://doi.org/10.1038/182234a0>.
- [37] R.W. DeBlois, C.P. Bean, Counting and Sizing of Submicron Particles by the Resistive Pulse Technique, *Review of Scientific Instruments*. 41 (1970) 909–916. <https://doi.org/10.1063/1.1684724>.
- [38] R.W. DeBlois, C.P. Bean, R.K.A. Wesley, Electrokinetic measurements with submicron particles and pores by the resistive pulse technique, *Journal of Colloid and Interface Science*. 61 (1977) 323–335. [https://doi.org/10.1016/0021-9797\(77\)90395-2](https://doi.org/10.1016/0021-9797(77)90395-2).

- [39] R.W. DeBlois, E.E. Uzgiris, D.H. Cluxton, H.M. Mazzone, Comparative measurements of size and polydispersity of several insect viruses, *Analytical Biochemistry*. 90 (1978) 273–288. [https://doi.org/10.1016/0003-2697\(78\)90032-5](https://doi.org/10.1016/0003-2697(78)90032-5).
- [40] J. Zimmerberg, V.A. Parsegian, Polymer inaccessible volume changes during opening and closing of a voltage-dependent ionic channel, *Nature*. 323 (1986) 36–39. <https://doi.org/10.1038/323036a0>.
- [41] P. Mueller, D.O. Rudin, H.T. Tien, W.C. Wescott, Reconstitution of Excitable Cell Membrane Structure in Vitro, *Circulation*. 26 (1962) 1167–1171. <https://doi.org/10.1161/01.CIR.26.5.1167>.
- [42] P. Mueller, D.O. Rudin, H. Ti Tien, W.C. Wescott, Reconstitution of Cell Membrane Structure in vitro and its Transformation into an Excitable System, *Nature*. 194 (1962) 979–980. <https://doi.org/10.1038/194979a0>.
- [43] P. Mueller, D.O. Rudin, H.T. Tien, W.C. Wescott, Methods for the formation of single bimolecular lipid membranes in aqueous solution, *The Journal of Physical Chemistry*. 67 (1963) 534–535.
- [44] E. Neher, B. Sakmann, Single-channel currents recorded from membrane of denervated frog muscle fibres, *Nature*. 260 (1976) 799–802. <https://doi.org/10.1038/260799a0>.
- [45] K.S. Cole, Dynamic electrical characteristics of the squid axon membrane, *Archives Des Sciences Physiologiques*. 3 (1949) 253–258.
- [46] A.L. Hodgkin, A.F. Huxley, B. Katz, Measurement of current-voltage relations in the membrane of the giant axon of *Loligo*, *The Journal of Physiology*. 116 (1952) 424–448. <https://doi.org/10.1113/jphysiol.1952.sp004716>.
- [47] K.S. Cole, J.W. Moore, Ionic Current Measurements in the Squid Giant Axon Membrane, *Journal of General Physiology*. 44 (1960) 123–167. <https://doi.org/10.1085/jgp.44.1.123>.
- [48] H. Mamad-Hemouch, H. Ramoul, M. Abou Taha, L. Bacri, C. Huin, C. Przybylski, A. Oukhaled, B. Thiébot, G. Patriarche, N. Jarroux, J. Pelta, Biomimetic Nanotubes Based on Cyclodextrins for Ion-Channel Applications, *Nano Lett*. 15 (2015) 7748–7754. <https://doi.org/10.1021/acs.nanolett.5b03938>.
- [49] L. Liu, C. Yang, K. Zhao, J. Li, H.-C. Wu, Ultrashort single-walled carbon nanotubes in a lipid bilayer as a new nanopore sensor, *Nature Communications*. 4 (2013) 2989. <https://doi.org/10.1038/ncomms3989>.
- [50] N.A.W. Bell, C.R. Engst, Ma. Ablay, G. Divitini, C. Ducati, im Liedl, U.F. Keyser, DNA Origami Nanopores, *Nano Lett*. 12 (2012) 512–517. <https://doi.org/10.1021/nl204098n>.
- [51] H. Furukawa, K.E. Cordova, M. O’Keeffe, O.M. Yaghi, The Chemistry and Applications of Metal-Organic Frameworks, *Science*. 341 (2013). <https://doi.org/10.1126/science.1230444>.

- [52] Y. Jiang, W. Ma, Y. Qiao, Y. Xue, J. Lu, J. Gao, N. Liu, F. Wu, P. Yu, L. Jiang, L. Mao, Metal–Organic Framework Membrane Nanopores as Biomimetic Photoresponsive Ion Channels and Photodriven Ion Pumps, *Angewandte Chemie*. 132 (2020) 12895–12899. <https://doi.org/10.1002/ange.202005084>.
- [53] S.M. Bezrukov, Ion Channels as Molecular Coulter Counters to Probe Metabolite Transport, *Journal of Membrane Biology*. 174 (2000) 1–13. <https://doi.org/10.1007/s002320001026>.
- [54] H. Bayley, O. Braha, L.-Q. Gu, Stochastic Sensing with Protein Pores, *Advanced Materials*. 12 (2000) 139–142. [https://doi.org/10.1002/\(SICI\)1521-4095\(200001\)12:2<139::AID-ADMA139>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1521-4095(200001)12:2<139::AID-ADMA139>3.0.CO;2-Q).
- [55] B. Hille, *Ion channels of excitable membranes*, Sinauer, Sunderland, MA, 2001.
- [56] S. Simon M., G. Blobel, Mechanisms of Translocation of Proteins across Membranes, in: *Endoplasmic Reticulum*, Springer, Boston, MA, 1993. https://doi.org/10.1007/978-1-4615-2912-5_1.
- [57] S.M. Simon, G. Blobel, A protein-conducting channel in the endoplasmic reticulum, *Cell*. 65 (1991) 371–380. [https://doi.org/10.1016/0092-8674\(91\)90455-8](https://doi.org/10.1016/0092-8674(91)90455-8).
- [58] B. Eisenberg, Ionic channels in biological membranes: natural nanotubes, *Accounts of Chemical Research*. 31 (1998) 117–123. [https://doi.org/10.1016/0092-8674\(91\)90455-8](https://doi.org/10.1016/0092-8674(91)90455-8)
- [59] R.S. Eisenberg, From Structure to Function in Open Ionic Channels, *Journal of Membrane Biology*. 171 (1999) 1–24. <https://doi.org/10.1007/s002329900554>.
- [60] B. Hille, Ionic channels in nerve membranes, *Progress in Biophysics and Molecular Biology*. 21 (1970) 1–32. [https://doi.org/10.1016/0079-6107\(70\)90022-2](https://doi.org/10.1016/0079-6107(70)90022-2).
- [61] M.S. Shapiro, J. Gomez, S.E. Hamilton, B. Hille, M.D. Loose, N.M. Nathanson, J.P. Roche, J. Wess, Identification of subtypes of muscarinic receptors that regulate Ca²⁺ and K⁺ channel activity in sympathetic neurons, *Life Sciences*. 68 (2001) 2481–2487. [https://doi.org/10.1016/S0024-3205\(01\)01042-6](https://doi.org/10.1016/S0024-3205(01)01042-6).
- [62] G. Heimlich, C.D. Bortner, J.A. Cidlowski, Apoptosis and Cell Volume Regulation, in: P.K. Lauf, N.C. Adragna (Eds.), *Cell Volume and Signaling*, Springer US, 2005: pp. 189–203. https://doi.org/10.1007/0-387-23752-6_18.
- [63] M.J. Welsh, A.E. Smith, Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis, *Cell*. 73 (1993) 1251–1254. [https://doi.org/10.1016/0092-8674\(93\)90353-R](https://doi.org/10.1016/0092-8674(93)90353-R).
- [64] G. Diamond, N. Beckloff, A. Weinberg, K. Kisich, The Roles of Antimicrobial Peptides in Innate Host Defense, *CPD*. 15 (2009) 2377–2392. <https://doi.org/10.2174/138161209788682325>.
- [65] G.A. Woolley, B.A. Wallace, Model ion channels: gramicidin and alamethicin, *The Journal of Membrane Biology*. 129 (1992) 109–136. <https://doi.org/10.1007/BF00219508>.

- [66] S.M. Bezrukov, I. Vodyanoy, Noise-induced enhancement of signal transduction across voltage-dependent ion channels, *Nature*. 378 (1995) 362–364. <https://doi.org/10.1038/378362a0>.
- [67] M.W. Parker, S.C. Feil, Pore-forming protein toxins: from structure to function, *Progress in Biophysics and Molecular Biology*. 88 (2005) 91–142. <https://doi.org/10.1016/j.pbiomolbio.2004.01.009>.
- [68] B. Geny, M.R. Popoff, Bacterial protein toxins and lipids: pore formation or toxin entry into cells, *Biology of the Cell*. 98 (2006) 667–678. <https://doi.org/10.1042/BC20050082>.
- [69] C.J. Rosado, A.M. Buckle, R.H.P. Law, R.E. Butcher, W.-T. Kan, C.H. Bird, K. Ung, K.A. Browne, K. Baran, T.A. Bashtannyk-Puhalovich, N.G. Faux, W. Wong, C.J. Porter, R.N. Pike, A.M. Ellisdon, M.C. Pearce, S.P. Bottomley, J. Emsley, A.I. Smith, J. Rossjohn, E.L. Hartland, I. Voskoboinik, J.A. Trapani, P.I. Bird, M.A. Dunstone, J.C. Whisstock, A Common Fold Mediates Vertebrate Defense and Bacterial Attack, *Science*. 317 (2007) 1548–1551. <https://doi.org/10.1126/science.1144706>.
- [70] M.D. Peraro, F.G. van der Goot, Pore-forming toxins: ancient, but never really out of fashion, *Nature Reviews Microbiology*. 14 (2016) 77–92. <https://doi.org/10.1038/nrmicro.2015.3>.
- [71] I. Iacovache, F.G. van der Goot, L. Pernot, Pore formation: An ancient yet complex form of attack, *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 1778 (2008) 1611–1623. <https://doi.org/10.1016/j.bbamem.2008.01.026>.
- [72] J.M. Janda, S.L. Abbott, Evolving Concepts Regarding the Genus *Aeromonas*: An Expanding Panorama of Species, Disease Presentations, and Unanswered Questions, *Clinical Infectious Diseases*. 27 (1998) 332–344. <https://doi.org/10.1086/514652>.
- [73] S.M. Kirov, L.A. O'Donovan, K. Sanderson, Functional Characterization of Type IV Pili Expressed on Diarrhea-Associated Isolates of *Aeromonas* species, *Infect. Immun.* 67 (1999) 5447–5454. <https://doi.org/10.1128/IAI.67.10.5447-5454.1999>.
- [74] K. Tanaka, J.M.M. Caaveiro, K. Morante, J.M. González-Mañas, K. Tsumoto, Structural basis for self-assembly of a cytolytic pore lined by protein and lipid, *Nat Commun.* 6 (2015) 6337. <https://doi.org/10.1038/ncomms7337>.
- [75] W. Peng, M. de Souza Santos, Y. Li, D.R. Tomchick, K. Orth, High-resolution cryo-EM structures of the *E. coli* hemolysin ClyA oligomers, *PLoS ONE*. 14 (2019) e0213423. <https://doi.org/10.1371/journal.pone.0213423>.
- [76] L. Song, M.R. Hobaugh, C. Shustak, S. Cheley, H. Bayley, J.E. Gouaux, Structure of Staphylococcal α -Hemolysin, a Heptameric Transmembrane Pore, *Science*. 274 (1996) 1859–1865. <https://doi.org/10.1126/science.274.5294.1859>.
- [77] I. Iacovache, S. De Carlo, N. Cirauqui, M. Dal Peraro, F.G. van der Goot, B. Zuber, Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process, *Nat Commun.* 7 (2016) 12062. <https://doi.org/10.1038/ncomms12062>.

- [78] M. Podobnik, P. Savory, N. Rojko, M. Kisovec, N. Wood, R. Hambley, J. Pugh, E.J. Wallace, L. McNeill, M. Bruce, I. Liko, T.M. Allison, S. Mehmood, N. Yilmaz, T. Kobayashi, R.J.C. Gilbert, C.V. Robinson, L. Jayasinghe, G. Anderluh, Crystal structure of an invertebrate cytolysin pore reveals unique properties and mechanism of assembly, *Nat Commun.* 7 (2016) 11598. <https://doi.org/10.1038/ncomms11598>.
- [79] M. Bokori-Brown, T.G. Martin, C.E. Naylor, A.K. Basak, R.W. Titball, C.G. Savva, Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein, *Nat Commun.* 7 (2016) 11293. <https://doi.org/10.1038/ncomms11293>.
- [80] B.J. Berube, J.B. Wardenburg, *Staphylococcus aureus* α -Toxin: Nearly a Century of Intrigue, *Toxins.* 5 (2013) 1140–1166. <https://doi.org/10.3390/toxins5061140>.
- [81] S.E. Henrickson, M. Misakian, B. Robertson, J.J. Kasianowicz, Driven DNA Transport into an Asymmetric Nanometer-Scale Pore, *Phys. Rev. Lett.* 85 (2000) 3057–3060. <https://doi.org/10.1103/PhysRevLett.85.3057>.
- [82] L. Brun, M. Pastoriza-Gallego, G. Oukhaled, J. Mathé, L. Bacri, L. Auvray, J. Pelta, Dynamics of Polyelectrolyte Transport through a Protein Channel as a Function of Applied Voltage, *Phys. Rev. Lett.* 100 (2008) 158302. <https://doi.org/10.1103/PhysRevLett.100.158302>.
- [83] G. Oukhaled, L. Bacri, J. Mathé, J. Pelta, L. Auvray, Effect of screening on the transport of polyelectrolytes through nanopores, *EPL.* 82 (2008) 48003. <https://doi.org/10.1209/0295-5075/82/48003>.
- [84] A. Aksimentiev, K. Schulten, Imaging α -Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability, and the Electrostatic Potential Map, *Biophysical Journal.* 88 (2005) 3745–3761. <https://doi.org/10.1529/biophysj.104.058727>.
- [85] O. Braha, L.-Q. Gu, L. Zhou, X. Lu, S. Cheley, H. Bayley, Simultaneous stochastic sensing of divalent metal ions, *Nat Biotechnol.* 18 (2000) 1005–1007. <https://doi.org/10.1038/79275>.
- [86] Y.E. Korchev, G.M. Alder, A. Bakhramov, C.L. Bashford, B.S. Joomun, E.V. Sviderskaya, P.N.R. Usherwood, C.A. Pasternak, *Staphylococcus aureus* α -toxin-induced pores: Channel-like behavior in lipid bilayers and patch clamped cells, *J. Membran Biol.* 143 (1995) 143–151. <https://doi.org/10.1007/BF00234660>.
- [87] X. Kang, L.-Q. Gu, S. Cheley, H. Bayley, Single Protein Pores Containing Molecular Adapters at High Temperatures, *Angew. Chem. Int. Ed.* 44 (2005) 1495–1499. <https://doi.org/10.1002/anie.200461885>.
- [88] M. Pastoriza-Gallego, G. Oukhaled, J. Mathé, B. Thiebot, J.-M. Betton, L. Auvray, J. Pelta, Urea denaturation of α -hemolysin pore inserted in planar lipid bilayer detected by single nanopore recording: Loss of structural asymmetry, *FEBS Letters.* 581 (2007) 3371–3376. <https://doi.org/10.1016/j.febslet.2007.06.036>.
- [89] M. Soskine, A. Biesemans, B. Moeyaert, S. Cheley, H. Bayley, G. Maglia, An Engineered ClyA Nanopore Detects Folded Target Proteins by Selective External

- Association and Pore Entry, *Nano Lett.* 12 (2012) 4895–4900. <https://doi.org/10.1021/nl3024438>.
- [90] G. Huang, K. Willems, M. Soskine, C. Wloka, G. Maglia, Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores, *Nat Commun.* 8 (2017) 935. <https://doi.org/10.1038/s41467-017-01006-4>.
- [91] G. Huang, A. Voet, G. Maglia, FraC nanopores with adjustable diameter identify the mass of opposite-charge peptides with 44 dalton resolution, *Nat Commun.* 10 (2019) 835. <https://doi.org/10.1038/s41467-019-08761-6>.
- [92] C. Wloka, N.L. Mutter, M. Soskine, G. Maglia, Alpha-Helical Fragaceatoxin C Nanopore Engineered for Double-Stranded and Single-Stranded Nucleic Acid Analysis, *Angew. Chem. Int. Ed.* 55 (2016) 12494–12498. <https://doi.org/10.1002/anie.201606742>.
- [93] R. Benz, *Bacterial and Eukaryotic Porins: Structure, Function, Mechanism*, John Wiley & Sons, 2006.
- [94] H. Nikaido, Molecular Basis of Bacterial Outer Membrane Permeability Revisited, *MMBR.* 67 (2003) 593–656. <https://doi.org/10.1128/MMBR.67.4.593-656.2003>.
- [95] R. Koebnik, K.P. Locher, P.V. Gelder, Structure and function of bacterial outer membrane proteins: barrels in a nutshell, *Molecular Microbiology.* 37 (2000) 239–253. <https://doi.org/10.1046/j.1365-2958.2000.01983.x>.
- [96] J. Vergalli, I.V. Bodrenko, M. Masi, L. Moynié, S. Acosta-Gutiérrez, J.H. Naismith, A. Davin-Regli, M. Ceccarelli, B. van den Berg, M. Winterhalter, J.-M. Pagès, Porins and small-molecule translocation across the outer membrane of Gram-negative bacteria, *Nat Rev Microbiol.* 18 (2020) 164–176. <https://doi.org/10.1038/s41579-019-0294-2>.
- [97] T. Schirmer, General and Specific Porins from Bacterial Outer Membranes, *Journal of Structural Biology.* 121 (1998) 101–109. <https://doi.org/10.1006/jsbi.1997.3946>.
- [98] E.M. Nestorovich, C. Danelon, M. Winterhalter, S.M. Bezrukov, Designed to penetrate: Time-resolved interaction of single antibiotic molecules with bacterial pores, *PNAS.* 99 (2002) 9789–9794. <https://doi.org/10.1073/pnas.152206799>.
- [99] C. Danelon, E.M. Nestorovich, M. Winterhalter, M. Ceccarelli, S.M. Bezrukov, Interaction of Zwitterionic Penicillins with the OmpF Channel Facilitates Their Translocation, *Biophysical Journal.* 90 (2006) 1617–1627. <https://doi.org/10.1529/biophysj.105.075192>.
- [100] K.R. Mahendran, C. Chimere, T. Mach, M. Winterhalter, Antibiotic translocation through membrane channels: temperature-dependent ion current fluctuation for catching the fast events, *Eur Biophys J.* 38 (2009) 1141–1145. <https://doi.org/10.1007/s00249-009-0495-0>.
- [101] M.R. Kozhinjampara, M. Kreir, N. Fertig, J.-M. Pages, E. Hajjar, M. Ceccarelli, M. Winterhalter, Permeation of Antibiotics through Bacterial Porins: Screening for Influx on a Single Molecular Level, *Biophysical Journal.* 98 (2010) 375a.

- [102] J.A. Bafna, E. Sans-Serramitjana, S. Acosta-Gutiérrez, I.V. Bodrenko, D. Hörömpöli, A. Berscheid, H. Brötz-Oesterhelt, M. Winterhalter, M. Ceccarelli, Kanamycin Uptake into *Escherichia coli* Is Facilitated by OmpF and OmpC Porin Channels Located in the Outer Membrane, *ACS Infect. Dis.* 6 (2020) 1855–1865. <https://doi.org/10.1021/acsinfecdis.0c00102>.
- [103] J. Wang, J.D. Prajapati, U. Kleinekathöfer, M. Winterhalter, Dynamic interaction of fluoroquinolones with magnesium ions monitored using bacterial outer membrane nanopores, *Chem. Sci.* 11 (2020) 10344–10353. <https://doi.org/10.1039/D0SC03486J>.
- [104] I. Ghai, M. Winterhalter, R. Wagner, Probing transport of charged β -lactamase inhibitors through OmpC, a membrane channel from *E. coli*, *Biochemical and Biophysical Research Communications*. 484 (2017) 51–55. <https://doi.org/10.1016/j.bbrc.2017.01.076>.
- [105] J.-M. Pagès, C.E. James, M. Winterhalter, The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria, *Nature Reviews Microbiology*. 6 (2008) 893–903. <https://doi.org/10.1038/nrmicro1994>.
- [106] M. Fahie, C. Chisholm, M. Chen, Resolved Single-Molecule Detection of Individual Species within a Mixture of *anti* -Biotin Antibodies Using an Engineered Monomeric Nanopore, *ACS Nano*. 9 (2015) 1089–1098. <https://doi.org/10.1021/nn506606e>.
- [107] C. Schmitt, J.A. Bafna, B. Schmid, S. Klingl, S. Baier, B. Hemmis, R. Wagner, M. Winterhalter, L.M. Voll, Manipulation of charge distribution in the arginine and glutamate clusters of the OmpG pore alters sugar specificity and ion selectivity, *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 1861 (2019) 183021. <https://doi.org/10.1016/j.bbamem.2019.07.009>.
- [108] A.K. Thakur, M.G. Larimi, K. Gooden, L. Movileanu, Aberrantly Large Single-Channel Conductance of Polyhistidine Arm-Containing Protein Nanopores, *Biochemistry*. 56 (2017) 4895–4905. <https://doi.org/10.1021/acs.biochem.7b00577>.
- [109] E. Berkane, F. Orlik, A. Charbit, C. Danelon, D. Fournier, R. Benz, M. Winterhalter, Nanopores: maltoporin channel as a sensor for maltodextrin and lambda-phage, *Journal of Nanobiotechnology*. 3 (2005) 3. <https://doi.org/10.1186/1477-3155-3-3>
- [110] P.R. Singh, I. Bárcena-Uribarri, N. Modi, U. Kleinekathöfer, R. Benz, M. Winterhalter, K.R. Mahendran, Pulling Peptides across Nanochannels: Resolving Peptide Binding and Translocation through the Hetero-oligomeric Channel from *Nocardia farcinica*, *ACS Nano*. 6 (2012) 10699–10707. <https://doi.org/10.1021/nn303900y>.
- [111] T.Z. Butler, M. Pavlenok, I.M. Derrington, M. Niederweis, J.H. Gundlach, Single-molecule DNA detection with an engineered MspA protein nanopore, *Proceedings of the National Academy of Sciences*. 105 (2008) 20647–20652. <https://doi.org/10.1073/pnas.0807514106>.
- [112] D. Jacobs, D.P. Hoogerheide, A. Rovini, Z. Jiang, J.C. Lee, T.K. Rostovtseva, S.M. Bezrukov, Probing Membrane Association of α -Synuclein Domains with VDAC Nanopore Reveals Unexpected Binding Pattern, *Scientific Reports*. 9 (2019) 4580. <https://doi.org/10.1038/s41598-019-40979-8>.

- [113] T.K. Rostovtseva, A. Komarov, S.M. Bezrukov, M. Colombini, VDAC Channels Differentiate between Natural Metabolites and Synthetic Molecules, *Journal of Membrane Biology*. 187 (2002) 147–156. <https://doi.org/10.1007/s00232-001-0159-1>.
- [114] D.P. Hoogerheide, P.A. Gurnev, T.K. Rostovtseva, S.M. Bezrukov, Real-Time Nanopore-Based Recognition of Protein Translocation Success, *Biophysical Journal*. 114 (2018) 772–776. <https://doi.org/10.1016/j.bpj.2017.12.019>.
- [115] M. Colombini, VDAC: The channel at the interface between mitochondria and the cytosol, *Mol Cell Biochem*. 256 (2004) 107–115. <https://doi.org/10.1023/B:MCBI.0000009862.17396.8d>.
- [116] P. Goyal, P.V. Krasteva, N. Van Gerven, F. Gubellini, I. Van den Broeck, A. Trounion-Tsailaki, W. Jonckheere, G. Péhau-Arnaudet, J.S. Pinkner, M.R. Chapman, S.J. Hultgren, S. Howorka, R. Fronzes, H. Remaut, Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG, *Nature*. 516 (2014) 250–253. <https://doi.org/10.1038/nature13768>.
- [117] L.S. Robinson, E.M. Ashman, S.J. Hultgren, M.R. Chapman, Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein, *Mol Microbiol*. 59 (2006) 870–881. <https://doi.org/10.1111/j.1365-2958.2005.04997.x>.
- [118] N. Van Gerven, S.E. Van der Verren, D.M. Reiter, H. Remaut, The Role of Functional Amyloids in Bacterial Virulence, *Journal of Molecular Biology*. 430 (2018) 3657–3684. <https://doi.org/10.1016/j.jmb.2018.07.010>.
- [119] S.E. Van der Verren, N. Van Gerven, W. Jonckheere, R. Hambley, P. Singh, J. Kilgour, M. Jordan, E.J. Wallace, L. Jayasinghe, H. Remaut, A dual-constriction biological nanopore resolves homonucleotide sequences with high fidelity, *Nat Biotechnol*. 38 (2020) 1415–1420. <https://doi.org/10.1038/s41587-020-0570-8>.
- [120] C.G. Brown, J. Clarke, Nanopore development at Oxford Nanopore, *Nat Biotechnol*. 34 (2016) 810–811. <https://doi.org/10.1038/nbt.3622>.
- [121] J.-M. Carter, S. Hussain, Robust long-read native DNA sequencing using the ONT CsgG Nanopore system, *Wellcome Open Res*. 2 (2018). <https://doi.org/10.12688/wellcomeopenres.11246.3>.
- [122] P.S. Phale, T. Schirmer, A. Prilipov, K.-L. Lou, A. Hardmeyer, J.P. Rosenbusch, Voltage gating of Escherichia coli porin channels: Role of the constriction loop, *PNAS*. 94 (1997) 6741–6745. <https://doi.org/10.1073/pnas.94.13.6741>.
- [123] M. Chen, Q.-H. Li, H. Bayley, Orientation of the Monomeric Porin OmpG in Planar Lipid Bilayers, *ChemBioChem*. 9 (2008) 3029–3036. <https://doi.org/10.1002/cbic.200800444>.
- [124] N. Liu, A.H. Delcour, The spontaneous gating activity of OmpC porin is affected by mutations of a putative hydrogen bond network or of a salt bridge between the L3 loop and the barrel, *Protein Engineering Design and Selection*. 11 (1998) 797–802. <https://doi.org/10.1093/protein/11.9.797>.

- [125] L.W. Black, DNA PACKAGING IN dsDNA BACTERIOPHAGES, *Annual Review of Microbiology*. 43 (1989) 267–292. <https://doi.org/10.1146/annurev.mi.43.100189.001411>.
- [126] C.-S. Lee, P. Guo, Sequential Interactions of Structural Proteins in Phage Φ 29 Procapsid Assembly, *J. VIROL.* 69 (1995) 9. <https://doi.org/10.1128/jvi.69.8.5024-5032.1995>.
- [127] P. Guo, T.J. Lee, Viral nanomotors for packaging of dsDNA and dsRNA, *Mol Microbiol.* 64 (2007) 886–903. <https://doi.org/10.1111/j.1365-2958.2007.05706.x>.
- [128] V.B. Rao, M. Feiss, The Bacteriophage DNA Packaging Motor, *Annu. Rev. Genet.* 42 (2008) 647–681. <https://doi.org/10.1146/annurev.genet.42.110807.091545>.
- [129] A.A. Simpson, Y. Tao, M.O. Badasso, P.J. Jardine, Structure determination of the head±tail connector of bacteriophage ϕ 29, (2001) 10. <https://doi.org/10.1107/s0907444901010435>
- [130] A. Guasch, J. Pous, B. Ibarra, F.X. Gomis-Rüth, J.M. Valpuesta, N. Sousa, J.L. Carrascosa, M. Coll, Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage ϕ 29 connector particle11Edited by R. Huber, *Journal of Molecular Biology*. 315 (2002) 663–676. <https://doi.org/10.1006/jmbi.2001.5278>.
- [131] X. Agirrezabala, J. Martín-Benito, M. Valle, J.M. González, A. Valencia, J.M. Valpuesta, J.L. Carrascosa, Structure of the Connector of Bacteriophage T7 at 8Å Resolution: Structural Homologies of a Basic Component of a DNA Translocating Machinery, *Journal of Molecular Biology*. 347 (2005) 895–902. <https://doi.org/10.1016/j.jmb.2005.02.005>.
- [132] D. Wendell, P. Jing, J. Geng, V. Subramaniam, T.J. Lee, C. Montemagno, P. Guo, Translocation of double-stranded DNA through membrane-adapted ϕ 29 motor protein nanopores, *Nature Nanotech.* 4 (2009) 765–772. <https://doi.org/10.1038/nnano.2009.259>.
- [133] P. Jing, F. Haque, A.P. Vonderheide, C. Montemagno, P. Guo, Robust properties of membrane-embedded connector channel of bacterial virus ϕ 29 DNA packaging motor, *Mol. BioSyst.* 6 (2010) 1844. <https://doi.org/10.1039/c003010d>.
- [134] J. Geng, S. Wang, H. Fang, P. Guo, Channel Size Conversion of Φ 29 DNA-Packaging Nanomotor for Discrimination of Single- and Double-Stranded Nucleic Acids, *ACS Nano*. 7 (2013) 3315–3323. <https://doi.org/10.1021/nn400020z>.
- [135] J. Geng, H. Fang, F. Haque, L. Zhang, P. Guo, Three reversible and controllable discrete steps of channel gating of a viral DNA packaging motor, *Biomaterials*. 32 (2011) 8234–8242. <https://doi.org/10.1016/j.biomaterials.2011.07.034>.
- [136] F. Haque, J. Geng, C. Montemagno, P. Guo, Incorporation of a viral DNA-packaging motor channel in lipid bilayers for real-time, single-molecule sensing of chemicals and double-stranded DNA, *Nat Protoc.* 8 (2013) 373–392. <https://doi.org/10.1038/nprot.2013.001>.

- [137] F. Haque, S. Wang, C. Stites, L. Chen, C. Wang, P. Guo, Single pore translocation of folded, double-stranded, and tetra-stranded DNA through channel of bacteriophage phi29 DNA packaging motor, *Biomaterials*. 53 (2015) 744–752. <https://doi.org/10.1016/j.biomaterials.2015.02.104>.
- [138] B. Cressiot, S.J. Greive, W. Si, T.C. Pascoa, M. Mojtabavi, M. Chechik, H.T. Jenkins, X. Lu, K. Zhang, A. Aksimentiev, A.A. Antson, M. Wanunu, Porphyrin-Assisted Docking of a Thermophage Portal Protein into Lipid Bilayers: Nanopore Engineering and Characterization, *ACS Nano*. 11 (2017) 11931–11945. <https://doi.org/10.1021/acsnano.7b06980>.
- [139] S. Wang, Z. Zhao, F. Haque, P. Guo, Engineering of protein nanopores for sequencing, chemical or protein sensing and disease diagnosis, *Current Opinion in Biotechnology*. 51 (2018) 80–89. <https://doi.org/10.1016/j.copbio.2017.11.006>.
- [140] M. Lepoitevin, T. Ma, M. Bechelany, J.-M. Janot, S. Balme, Functionalization of single solid state nanopores to mimic biological ion channels: A review, *Advances in Colloid and Interface Science*. 250 (2017) 195–213. <https://doi.org/10.1016/j.cis.2017.09.001>.
- [141] L. Movileanu, S. Howorka, O. Braha, H. Bayley, Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore, *Nat Biotechnol*. 18 (2000) 1091–1095. <https://doi.org/10.1038/80295>.
- [142] S. Howorka, L. Movileanu, X. Lu, M. Magnon, S. Cheley, O. Braha, H. Bayley, A Protein Pore with a Single Polymer Chain Tethered within the Lumen, *J. Am. Chem. Soc*. 122 (2000) 2411–2416. <https://doi.org/10.1021/ja993221h>.
- [143] S. Howorka, J. Nam, H. Bayley, D. Kahne, Stochastic Detection of Monovalent and Bivalent Protein–Ligand Interactions, *Angew. Chem*. 116 (2004) 860–864. <https://doi.org/10.1002/ange.200352614>.
- [144] H. Xie, O. Braha, L.-Q. Gu, S. Cheley, H. Bayley, Single-Molecule Observation of the Catalytic Subunit of cAMP-Dependent Protein Kinase Binding to an Inhibitor Peptide, *Chemistry & Biology*. 12 (2005) 109–120. <https://doi.org/10.1016/j.chembiol.2004.11.013>.
- [145] D. Rotem, L. Jayasinghe, M. Salichou, H. Bayley, Protein Detection by Nanopores Equipped with Aptamers, *J. Am. Chem. Soc*. 134 (2012) 2781–2787. <https://doi.org/10.1021/ja2105653>.
- [146] J. Li, D. Stein, C. McMullan, D. Branton, M.J. Aziz, J.A. Golovchenko, Ion-beam sculpting at nanometre length scales, *Nature*. 412 (2001) 166–169. <https://doi.org/10.1038/35084037>.
- [147] A.J. Storm, J.H. Chen, X.S. Ling, H.W. Zandbergen, C. Dekker, Fabrication of solid-state nanopores with single-nanometre precision, *Nature Mater*. 2 (2003) 537–540. <https://doi.org/10.1038/nmat941>.
- [148] E. Beamish, H. Kwok, V. Tabard-Cossa, M. Godin, Precise control of the size and noise of solid-state nanopores using high electric fields, *Nanotechnology*. 23 (2012) 405301. <https://doi.org/10.1088/0957-4484/23/40/405301>.

- [149] C. Dekker, Solid-state nanopores, *Nature Nanotechnology*. 2 (2007) 209–215. <https://doi.org/10.1038/nnano.2007.27>.
- [150] A. de la Escosura-Muñiz, A. Merkoçi, Nanochannels Preparation and Application in Biosensing, *ACS Nano*. 6 (2012) 7556–7583. <https://doi.org/10.1021/nn301368z>.
- [151] L. Xue, H. Yamazaki, R. Ren, M. Wanunu, A.P. Ivanov, J.B. Edel, Solid-state nanopore sensors, *Nat Rev Mater*. 5 (2020) 931–951. <https://doi.org/10.1038/s41578-020-0229-6>.
- [152] B.M. Venkatesan, A.B. Shah, J.-M. Zuo, R. Bashir, DNA Sensing Using Nanocrystalline Surface-Enhanced Al₂O₃ Nanopore Sensors, *Adv. Funct. Mater*. 20 (2010) 1266–1275. <https://doi.org/10.1002/adfm.200902128>.
- [153] K. Liu, J. Feng, A. Kis, A. Radenovic, Atomically Thin Molybdenum Disulfide Nanopores with High Sensitivity for DNA Translocation, *ACS Nano*. 8 (2014) 2504–2511. <https://doi.org/10.1021/nn406102h>.
- [154] S. Garaj, W. Hubbard, A. Reina, J. Kong, D. Branton, J.A. Golovchenko, Graphene as a subnanometre trans-electrode membrane, *Nature*. 467 (2010) 190–193. <https://doi.org/10.1038/nature09379>.
- [155] T. Jain, B.C. Raseria, R.J.S. Guerrero, M.S.H. Boutilier, S.C. O’Hern, J.-C. Idrobo, R. Karnik, Heterogeneous sub-continuum ionic transport in statistically isolated graphene nanopores, *Nature Nanotech*. 10 (2015) 1053–1057. <https://doi.org/10.1038/nnano.2015.222>.
- [156] P. Apel, Track etching technique in membrane technology, *Radiation Measurements*. 34 (2001) 559–566. [https://doi.org/10.1016/S1350-4487\(01\)00228-1](https://doi.org/10.1016/S1350-4487(01)00228-1).
- [157] X. Hou, W. Guo, L. Jiang, Biomimetic smart nanopores and nanochannels, *Chem. Soc. Rev*. 40 (2011) 2385. <https://doi.org/10.1039/c0cs00053a>.
- [158] M. Pevarnik, K. Healy, M.E. Toimil-Molares, A. Morrison, S.E. Létant, Z.S. Siwy, Polystyrene Particles Reveal Pore Substructure As They Translocate, *ACS Nano*. 6 (2012) 7295–7302. <https://doi.org/10.1021/nn302413u>.
- [159] S. Cabello-Aguilar, S. Balme, A.A. Chaaya, M. Bechelany, E. Balanzat, J.-M. Janot, C. Pochat-Bohatier, P. Miele, P. Dejardin, Slow translocation of polynucleotides and their discrimination by α -hemolysin inside a single track-etched nanopore designed by atomic layer deposition, *Nanoscale*. 5 (2013) 9582. <https://doi.org/10.1039/c3nr03683a>.
- [160] A. Oukhaled, B. Cressiot, L. Bacri, M. Pastoriza-Gallego, J.-M. Betton, E. Bourhis, R. Jede, J. Gierak, L. Auvray, J. Pelta, Dynamics of Completely Unfolded and Native Proteins through Solid-State Nanopores as a Function of Electric Driving Force, *ACS Nano*. 5 (2011) 3628–3638. <https://doi.org/10.1021/nn1034795>.
- [161] S.M. Iqbal, D. Akin, R. Bashir, Solid-state nanopore channels with DNA selectivity, *Nature Nanotech*. 2 (2007) 243–248. <https://doi.org/10.1038/nnano.2007.78>.

- [162] E.C. Yusko, Y.N. Billeh, J. Yang, M. Mayer, Nanopore Recordings to Quantify Activity-Related Properties of Proteins, in: S.M. Iqbal, R. Bashir (Eds.), *Nanopores*, Springer US, Boston, MA, 2011: pp. 203–225. https://doi.org/10.1007/978-1-4419-8252-0_9.
- [163] B.M. Venkatesan, J. Polans, J. Comer, S. Sridhar, D. Wendell, A. Aksimentiev, R. Bashir, Lipid bilayer coated Al₂O₃ nanopore sensors: towards a hybrid biological solid-state nanopore, *Biomed Microdevices*. 13 (2011) 671–682. <https://doi.org/10.1007/s10544-011-9537-3>.
- [164] J. Roman, O. Français, N. Jarroux, G. Patriarche, J. Pelta, L. Bacri, B. Le Pioufle, Solid-State Nanopore Easy Chip Integration in a Cheap and Reusable Microfluidic Device for Ion Transport and Polymer Conformation Sensing, *ACS Sens.* 3 (2018) 2129–2137. <https://doi.org/10.1021/acssensors.8b00700>.
- [165] A.R. Hall, A. Scott, D. Rotem, K.K. Mehta, H. Bayley, C. Dekker, Hybrid pore formation by directed insertion of α -haemolysin into solid-state nanopores, *Nature Nanotech.* 5 (2010) 874–877. <https://doi.org/10.1038/nnano.2010.237>.
- [166] S. Balme, J.-M. Janot, L. Berardo, F. Henn, D. Bonhenry, S. Kraszewski, F. Picaud, C. Ramseyer, New Bioinspired Membrane Made of a Biological Ion Channel Confined into the Cylindrical Nanopore of a Solid-State Polymer., *Nano Lett.* 11 (2011) 712–716. <https://doi.org/10.1021/nl103841m>.
- [167] S. Balme, F. Picaud, S. Kraszewski, P. Déjardin, J.M. Janot, M. Lepoitevin, J. Capomanes, C. Ramseyer, F. Henn, Controlling potassium selectivity and proton blocking in a hybrid biological/solid-state polymer nanoporous membrane, *Nanoscale*. 5 (2013) 3961–3968. <https://doi.org/10.1039/C3NR00564J>.
- [168] B. Cressiot, S.J. Greive, M. Mojtavavi, A.A. Antson, M. Wanunu, Thermostable virus portal proteins as reprogrammable adapters for solid-state nanopore sensors, *Nat Commun.* 9 (2018) 4652. <https://doi.org/10.1038/s41467-018-07116-x>.
- [169] M. Winterhalter, Black lipid membranes, *Current Opinion in Colloid & Interface Science*. 5 (2000) 250–255. [https://doi.org/10.1016/S1359-0294\(00\)00063-7](https://doi.org/10.1016/S1359-0294(00)00063-7).
- [170] S.M. Bezrukov, I. Vodyanoy, R.A. Brutyan, J.J. Kasianowicz, Dynamics and Free Energy of Polymers Partitioning into a Nanoscale Pore, *Macromolecules*. 29 (1996) 8517–8522. <https://doi.org/10.1021/ma960841j>.
- [171] O.V. Krasilnikov, R.Z. Sabirov, V.I. Ternovsky, P.G. Merzliak, J.N. Muratkhodjaev, A simple method for the determination of the pore radius of ion channels in planar lipid bilayer membranes, *FEMS Microbiology Letters*. 105 (1992) 93–100. <https://doi.org/10.1111/j.1574-6968.1992.tb05891.x>.
- [172] I. Vodyanoy, S.M. Bezrukov, Sizing of an ion pore by access resistance measurements, *Biophysical Journal*. 62 (1992) 10–11. [https://doi.org/10.1016/S0006-3495\(92\)81762-9](https://doi.org/10.1016/S0006-3495(92)81762-9).

- [173] S.M. Bezrukov, I. Vodyanoy, Probing alamethicin channels with water-soluble polymers. Effect on conductance of channel states, *Biophysical Journal*. 64 (1993) 16–25. [https://doi.org/10.1016/S0006-3495\(93\)81336-5](https://doi.org/10.1016/S0006-3495(93)81336-5).
- [174] P.G. Merzlyak, L.N. Yuldasheva, C.G. Rodrigues, C.M.M. Carneiro, O.V. Krasilnikov, S.M. Bezrukov, Polymeric Nonelectrolytes to Probe Pore Geometry: Application to the α -Toxin Transmembrane Channel, *Biophysical Journal*. 77 (1999) 3023–3033. [https://doi.org/10.1016/S0006-3495\(99\)77133-X](https://doi.org/10.1016/S0006-3495(99)77133-X).
- [175] J.W.F. Robertson, C.G. Rodrigues, V.M. Stanford, K.A. Robinson, O.V. Krasilnikov, J.J. Kasianowicz, Single-molecule mass spectrometry in solution using a solitary nanopore, *PNAS*. 104 (2007) 8207–8211. <https://doi.org/10.1073/pnas.0611085104>.
- [176] O.V. Krasilnikov, C.G. Rodrigues, S.M. Bezrukov, Single Polymer Molecules in a Protein Nanopore in the Limit of a Strong Polymer-Pore Attraction, *Phys. Rev. Lett.* 97 (2006) 018301. <https://doi.org/10.1103/PhysRevLett.97.018301>.
- [177] F. Piguet, H. Ouldali, F. Discala, M.-F. Breton, J.C. Behrends, J. Pelta, A. Oukhaled, High Temperature Extends the Range of Size Discrimination of Nonionic Polymers by a Biological Nanopore, *Sci Rep*. 6 (2016) 38675. <https://doi.org/10.1038/srep38675>.
- [178] W. Vercountere, S. Winters-Hilt, H. Olsen, D. Deamer, D. Haussler, M. Akeson, Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel, *Nat Biotechnol*. 19 (2001) 248–252. <https://doi.org/10.1038/85696>.
- [179] M. Akeson, D. Branton, J.J. Kasianowicz, E. Brandin, D.W. Deamer, Microsecond Time-Scale Discrimination Among Polycytidylic Acid, Polyadenylic Acid, and Polyuridylic Acid as Homopolymers or as Segments Within Single RNA Molecules, *Biophysical Journal*. 77 (1999) 3227–3233. [https://doi.org/10.1016/S0006-3495\(99\)77153-5](https://doi.org/10.1016/S0006-3495(99)77153-5).
- [180] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, D. Branton, Rapid nanopore discrimination between single polynucleotide molecules, *Proceedings of the National Academy of Sciences*. 97 (2000) 1079–1084. <https://doi.org/10.1073/pnas.97.3.1079>.
- [181] R. Kawano, A.E.P. Schibel, C. Cauley, H.S. White, Controlling the Translocation of Single-Stranded DNA through α -Hemolysin Ion Channels Using Viscosity, *Langmuir*. 25 (2009) 1233–1237. <https://doi.org/10.1021/la803556p>.
- [182] D. Stoddart, A.J. Heron, E. Mikhailova, G. Maglia, H. Bayley, Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore, *Proceedings of the National Academy of Sciences*. 106 (2009) 7702–7707. <https://doi.org/10.1073/pnas.0901054106>.
- [183] M. Rincon-Restrepo, E. Mikhailova, H. Bayley, G. Maglia, Controlled Translocation of Individual DNA Molecules through Protein Nanopores with Engineered Molecular Brakes, *Nano Lett.* 11 (2011) 746–750. <https://doi.org/10.1021/nl1038874>.

- [184] L.-Q. Gu, O. Braha, S. Conlan, S. Cheley, H. Bayley, Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter, *Nature*. 398 (1999) 686–690. <https://doi.org/10.1038/19491>.
- [185] J. Clarke, H.-C. Wu, L. Jayasinghe, A. Patel, S. Reid, H. Bayley, Continuous base identification for single-molecule nanopore DNA sequencing, *Nature Nanotech.* 4 (2009) 265–270. <https://doi.org/10.1038/nnano.2009.12>.
- [186] S. Howorka, S. Cheley, H. Bayley, Sequence-specific detection of individual DNA strands using engineered nanopores, *Nat Biotechnol.* 19 (2001) 636–639. <https://doi.org/10.1038/90236>.
- [187] B. Gyarfas, F. Olasagasti, S. Benner, D. Garalde, K.R. Lieberman, M. Akeson, Mapping the Position of DNA Polymerase-Bound DNA Templates in a Nanopore at 5 Å Resolution, *ACS Nano*. 3 (2009) 1457–1466. <https://doi.org/10.1021/nm900303g>.
- [188] E.A. Manrao, I.M. Derrington, A.H. Laszlo, K.W. Langford, M.K. Hopper, N. Gillgren, M. Pavlenok, M. Niederweis, J.H. Gundlach, Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase, *Nat Biotechnol.* 30 (2012) 349–353. <https://doi.org/10.1038/nbt.2171>.
- [189] A.H. Laszlo, I.M. Derrington, B.C. Ross, H. Brinkerhoff, A. Adey, I.C. Nova, J.M. Craig, K.W. Langford, J.M. Samson, R. Daza, K. Doering, J. Shendure, J.H. Gundlach, Decoding long nanopore sequencing reads of natural DNA, *Nat Biotechnol.* 32 (2014) 829–833. <https://doi.org/10.1038/nbt.2950>.
- [190] F. Olasagasti, K.R. Lieberman, S. Benner, G.M. Cherf, J.M. Dahl, D.W. Deamer, M. Akeson, Replication of individual DNA molecules under electronic control using a protein nanopore, *Nature Nanotech.* 5 (2010) 798–806. <https://doi.org/10.1038/nnano.2010.177>.
- [191] K.R. Lieberman, G.M. Cherf, M.J. Doody, F. Olasagasti, Y. Kolodji, M. Akeson, Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase, *Journal of the American Chemical Society*. 132 (2010) 17961–17972. <https://doi.org/10.1021/ja1087612>.
- [192] G.M. Cherf, K.R. Lieberman, H. Rashid, C.E. Lam, K. Karplus, M. Akeson, Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision, *Nature Biotechnology*. 30 (2012) 344–348. <https://doi.org/10.1038/nbt.2147>.
- [193] A. Meller, L. Nivon, D. Branton, Voltage-Driven DNA Translocations through a Nanopore, *Phys. Rev. Lett.* 86 (2001) 3435–3438. <https://doi.org/10.1103/PhysRevLett.86.3435>.
- [194] M. Faller, The Structure of a Mycobacterial Outer-Membrane Channel, *Science*. 303 (2004) 1189–1192. <https://doi.org/10.1126/science.1094114>.
- [195] D. Deamer, M. Akeson, D. Branton, Three decades of nanopore sequencing, *Nature Biotechnology*. 34 (2016) 518–524. <https://doi.org/10.1038/nbt.3423>.

- [196] I.M. Derrington, T.Z. Butler, M.D. Collins, E. Manrao, M. Pavlenok, M. Niederweis, J.H. Gundlach, Nanopore DNA sequencing with MspA, *Proceedings of the National Academy of Sciences*. 107 (2010) 16060–16065. <https://doi.org/10.1073/pnas.1001831107>.
- [197] Y. Wang, Q. Yang, Z. Wang, The evolution of nanopore sequencing, *Front. Genet.* 5 (2015). <https://doi.org/10.3389/fgene.2014.00449>.
- [198] Y. Astier, O. Braha, H. Bayley, Toward Single Molecule DNA Sequencing: Direct Identification of Ribonucleoside and Deoxyribonucleoside 5'-Monophosphates by Using an Engineered Protein Nanopore Equipped with a Molecular Adapter, *Journal of the American Chemical Society*, 128(5), 1705-1710. <https://doi.org/10.1021/ja057123+>
- [199] H.-C. Wu, Y. Astier, G. Maglia, E. Mikhailova, H. Bayley, Protein Nanopores with Covalently Attached Molecular Adapters, *J. Am. Chem. Soc.* 129 (2007) 16142–16148. <https://doi.org/10.1021/ja0761840>.
- [200] A.P. Demchenko, *Introduction to Fluorescence Sensing*, Springer Science & Business Media, 2008.
- [201] C. Formoso, The interaction of β -cyclodextrin with nucleic acid monomer units, *Biochemical and Biophysical Research Communications*. 50 (1973) 999–1005. [https://doi.org/10.1016/0006-291X\(73\)91505-2](https://doi.org/10.1016/0006-291X(73)91505-2).
- [202] S. Kumar, C. Tao, M. Chien, B. Hellner, A. Balijepalli, J.W.F. Robertson, Z. Li, J.J. Russo, J.E. Reiner, J.J. Kasianowicz, J. Ju, PEG-Labeled Nucleotides and Nanopore Detection for Single Molecule DNA Sequencing by Synthesis, *Sci Rep.* 2 (2012) 684. <https://doi.org/10.1038/srep00684>.
- [203] J. Shendure, Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome, *Science*. 309 (2005) 1728–1732. <https://doi.org/10.1126/science.1117389>.
- [204] E.R. Mardis, DNA sequencing technologies: 2006–2016, *Nature Protocols*. 12 (2017) 213–218. <https://doi.org/10.1038/nprot.2016.182>.
- [205] E.N. Ervin, G.A. Barrall, P. Pal, M.K. Bean, A.E.P. Schibel, A.D. Hibbs, Creating a Single Sensing Zone Within an Alpha-Hemolysin Pore via Site-Directed Mutagenesis, *BioNanoSci.* 4 (2014) 78–84. <https://doi.org/10.1007/s12668-013-0119-0>.
- [206] D. Stoddart, A.J. Heron, J. Klingelhoefer, E. Mikhailova, G. Maglia, H. Bayley, Nucleobase Recognition in ssDNA at the Central Constriction of the α -Hemolysin Pore, *Nano Lett.* 10 (2010) 3633–3637. <https://doi.org/10.1021/nl101955a>.
- [207] F. Sanger, Chemistry of insulin, *British Medical Bulletin*. 16 (1960) 183–188. <https://doi.org/10.1093/oxfordjournals.bmb.a069832>
- [208] P. Edman, E. Högfeldt, L.G. Sillén, Method for determination of the amino acid sequence in peptides, *Acta Chem. Scand.* 4 (1950) 283–293.

- [209] P. Edman, Preparation of phenyl thiohydantoin from some natural amino acids, *Acta Chem. Scand.* 4 (1950) 277–282.
- [210] P. Edman, Mechanism of the phenyl isothiocyanate degradation of peptides, *Nature*. 177 (1956) 667–668.
- [211] P. Edman, G. Begg, A Protein Sequenator, in: C. Liébecq (Ed.), *European Journal of Biochemistry*, Springer, Berlin, Heidelberg, 1967: pp. 80–91. https://doi.org/10.1007/978-3-662-25813-2_14.
- [212] P. Edman, Sequence Determination, in: S.B. Needleman (Ed.), *Protein Sequence Determination: A Sourcebook of Methods and Techniques*, Springer, Berlin, Heidelberg, 1970: pp. 211–255. https://doi.org/10.1007/978-3-662-12834-3_8.
- [213] R.M. Twyman, Proteomics, in: *Encyclopedia of Applied Ethics*, Elsevier, 2012: pp. 642–649. <https://doi.org/10.1016/B978-0-12-373932-2.00047-8>.
- [214] E. Pennisi, ENCODE Project Writes Eulogy for Junk DNA, *Science*. 337 (2012) 1159–1161. <https://doi.org/10.1126/science.337.6099.1159>.
- [215] J. Muñoz, A.J.R. Heck, From the Human Genome to the Human Proteome, *Angew. Chem. Int. Ed.* 53 (2014) 10864–10866. <https://doi.org/10.1002/anie.201406545>.
- [216] B.T. Chait, Mass Spectrometry: Bottom-Up or Top-Down?, *Science*. 314 (2006) 65–66. <https://doi.org/10.1126/science.1133987>.
- [217] L. Payet, M. Martinho, M. Pastoriza-Gallego, J.-M. Betton, L. Auvray, J. Pelta, J. Mathé, Thermal Unfolding of Proteins Probed at the Single Molecule Level Using Nanopores, *Anal. Chem.* 84 (2012) 4071–4076. <https://doi.org/10.1021/ac300129e>.
- [218] D. Rodriguez-Larrea, H. Bayley, Multistep protein unfolding during nanopore translocation, *Nature Nanotech.* 8 (2013) 288–295. <https://doi.org/10.1038/nnano.2013.22>.
- [219] J. Nivala, D.B. Marks, M. Akeson, Unfoldase-mediated protein translocation through an α -hemolysin nanopore, *Nature Biotechnology*. 31 (2013) 247–250. <https://doi.org/10.1038/nbt.2503>.
- [220] M. Pastoriza-Gallego, M.-F. Breton, F. Discala, L. Auvray, J.-M. Betton, J. Pelta, Evidence of Unfolded Protein Translocation through a Protein Nanopore, *ACS Nano*. 8 (2014) 11350–11360. <https://doi.org/10.1021/nn5042398>.
- [221] B. Cressiot, A. Oukhaled, G. Patriarche, M. Pastoriza-Gallego, J.-M. Betton, M. Muthukumar, L. Bacri, J. Pelta, Protein Transport through a Narrow Solid-State Nanopore at High Voltage: Experiments and Theory | *ACS Nano*, *ACS Nano*. 6 (2012) 6236–6243. <https://doi.org/10.1021/nn301672g>.
- [222] D.S. Talaga, J. Li, Single-Molecule Protein Unfolding in Solid State Nanopores, *J. Am. Chem. Soc.* 131 (2009) 9287–9297. <https://doi.org/10.1021/ja901088b>.

- [223] K.J. Freedman, S.R. Haq, J.B. Edel, P. Jemth, M.J. Kim, Single molecule unfolding and stretching of protein domains inside a solid-state nanopore by electric field, *Sci Rep.* 3 (2013) 1638. <https://doi.org/10.1038/srep01638>.
- [224] K.J. Freedman, M. Jürgens, A. Prabhu, C.W. Ahn, P. Jemth, J.B. Edel, M.J. Kim, Chemical, Thermal, and Electric Field Induced Unfolding of Single Protein Molecules Studied Using Nanopores, *Anal. Chem.* 83 (2011) 5137–5144. <https://doi.org/10.1021/ac2001725>.
- [225] C. Merstorf, B. Cressiot, M. Pastoriza-Gallego, A. Oukhaled, J.-M. Betton, L. Auvray, J. Pelta, Wild Type, Mutant Protein Unfolding and Phase Transition Detected by Single-Nanopore Recording, *ACS Chem. Biol.* 7 (2012) 652–658. <https://doi.org/10.1021/cb2004737>.
- [226] D. Japrun, J. Dogan, K.J. Freedman, A. Nadzeyka, S. Bauerdick, T. Albrecht, M.J. Kim, P. Jemth, J.B. Edel, Single-Molecule Studies of Intrinsically Disordered Proteins Using Solid-State Nanopores, *Anal. Chem.* 85 (2013) 2449–2456. <https://doi.org/10.1021/ac3035025>.
- [227] N.N. Jetha, V. Semchenko, D.S. Wishart, N.R. Cashman, A. Marziali, Nanopore Analysis of Wild-Type and Mutant Prion Protein (PrPC): Single Molecule Discrimination and PrPC Kinetics, *PLOS ONE.* 8 (2013) e54982. <https://doi.org/10.1371/journal.pone.0054982>.
- [228] E.C. Yusko, P. Prangkio, D. Sept, R.C. Rollings, J. Li, M. Mayer, Single-particle characterization of A β oligomers in solution, *ACS Nano.* 6 (2012) 5909–5919. <https://doi.org/10.1021/nn300542q>
- [229] E.C. Yusko, R. An, M. Mayer, Fluidic Diodes in Nano-and Microscale Pores to Detect Drug Aggregation, *Biophysical Journal.* 96 (2009) 646a.
- [230] N. Giamblanco, J.-M. Janot, A. Gubbiotti, M. Chinappi, S. Balme, Characterization of Food Amyloid Protein Digestion by Conical Nanopore, *Small Methods.* 4 (2020) 1900703. <https://doi.org/10.1002/smt.201900703>.
- [231] L.T. Sexton, L.P. Horne, S.A. Sherrill, G.W. Bishop, L.A. Baker, C.R. Martin, Resistive-Pulse Studies of Proteins and Protein/Antibody Complexes Using a Conical Nanotube Sensor, *J. Am. Chem. Soc.* 129 (2007) 13144–13152. <https://doi.org/10.1021/ja0739943>.
- [232] V. Van Meervelt, M. Soskine, S. Singh, G.K. Schuurman-Wolters, H.J. Wijma, B. Poolman, G. Maglia, Real-Time Conformational Changes and Controlled Orientation of Native Proteins Inside a Protein Nanoreactor, *J. Am. Chem. Soc.* 139 (2017) 18640–18646. <https://doi.org/10.1021/jacs.7b10106>.
- [233] P. Waduge, R. Hu, P. Bandarkar, H. Yamazaki, B. Cressiot, Q. Zhao, P.C. Whitford, M. Wanunu, Nanopore-Based Measurements of Protein Size, Fluctuations, and Conformational Changes, *ACS Nano.* 11 (2017) 5706–5716. <https://doi.org/10.1021/acsnano.7b01212>.

- [234] D. Rodriguez-Larrea, H. Bayley, Protein co-translocational unfolding depends on the direction of pulling, *Nat Commun.* 5 (2014) 4841. <https://doi.org/10.1038/ncomms5841>.
- [235] D. Fologea, B. Ledden, D.S. McNabb, J. Li, Electrical characterization of protein molecules by a solid-state nanopore, *Appl. Phys. Lett.* 91 (2007) 053901. <https://doi.org/10.1063/1.2767206>.
- [236] C. Plesa, S.W. Kowalczyk, R. Zinsmeister, A.Y. Grosberg, Y. Rabin, C. Dekker, Fast Translocation of Proteins through Solid State Nanopores, *Nano Lett.* 13 (2013) 658–663. <https://doi.org/10.1021/nl3042678>.
- [237] E. Kennedy, Z. Dong, C. Tennant, G. Timp, Reading the primary structure of a protein with 0.07 nm³ resolution using a subnanometre-diameter pore, *Nature Nanotech.* 11 (2016) 968–976. <https://doi.org/10.1038/nnano.2016.120>.
- [238] M. Kolmogorov, E. Kennedy, Z. Dong, G. Timp, P.A. Pevzner, Single-molecule protein identification by sub-nanopore sensors, *PLoS Comput Biol.* 13 (2017) e1005356. <https://doi.org/10.1371/journal.pcbi.1005356>.
- [239] W. Li, N.A.W. Bell, S. Hernández-Ainsa, V.V. Thacker, A.M. Thackray, R. Bujdoso, U.F. Keyser, Single Protein Molecule Detection by Glass Nanopores, *ACS Nano.* 7 (2013) 4129–4134. <https://doi.org/10.1021/nn4004567>.
- [240] L.J. Steinbock, S. Krishnan, R.D. Bulushev, S. Borgeaud, M. Blokesch, L. Feletti, A. Radenovic, Probing the size of proteins with glass nanopores, *Nanoscale.* 6 (2014) 14380–14387. <https://doi.org/10.1039/C4NR05001K>.
- [241] I. Nir, D. Huttner, A. Meller, Direct Sensing and Discrimination among Ubiquitin and Ubiquitin Chains Using Solid-State Nanopores, *Biophysical Journal.* 108 (2015) 2340–2349. <https://doi.org/10.1016/j.bpj.2015.03.025>.
- [242] R. Stefureac, Y. Long, H.-B. Kraatz, P. Howard, J.S. Lee, Transport of α -Helical Peptides through α -Hemolysin and Aerolysin Pores[†], *Biochemistry.* 45 (2006) 9172–9179. <https://doi.org/10.1021/bi0604835>.
- [243] C.P. Goodrich, S. Kirmizialtin, B.M. Huyghues-Despointes, A. Zhu, J.M. Scholtz, D.E. Makarov, L. Movileanu, Single-Molecule Electrophoresis of β -Hairpin Peptides by Electrical Recordings and Langevin Dynamics Simulations, *J. Phys. Chem. B.* 111 (2007) 3332–3335. <https://doi.org/10.1021/jp071364h>.
- [244] T.C. Sutherland, Y.-T. Long, R.-I. Stefureac, I. Bediako-Amoa, H.-B. Kraatz, J.S. Lee, Structure of Peptides Investigated by Nanopore Analysis, *Nano Lett.* 4 (2004) 1273–1277. <https://doi.org/10.1021/nl049413e>.
- [245] L. Movileanu, J.P. Schmittschmitt, J. Martin Scholtz, H. Bayley, Interactions of Peptides with a Protein Pore, *Biophysical Journal.* 89 (2005) 1030–1045. <https://doi.org/10.1529/biophysj.104.057406>.
- [246] R. Stefureac, L. Waldner, P. Howard, J.S. Lee, Nanopore Analysis of a Small 86-Residue Protein, *Small.* 4 (2008) 59–63. <https://doi.org/10.1002/smll.200700402>.

- [247] A. Asandei, A.E. Rossini, M. Chinappi, Y. Park, T. Luchian, Protein Nanopore-Based Discrimination between Selected Neutral Amino Acids from Polypeptides, *Langmuir*. 33 (2017) 14451–14459. <https://doi.org/10.1021/acs.langmuir.7b03163>.
- [248] C.B. Rosen, D. Rodriguez-Larrea, H. Bayley, Single-molecule site-specific detection of protein phosphorylation with a nanopore, *Nat Biotechnol*. 32 (2014) 179–181. <https://doi.org/10.1038/nbt.2799>.
- [249] A.E. Chavis, K.T. Brady, G.A. Hatmaker, C.E. Angevine, N. Kothalawala, A. Dass, J.W.F. Robertson, J.E. Reiner, Single Molecule Nanopore Spectrometry for Peptide Detection, *ACS Sens*. 2 (2017) 1319–1328. <https://doi.org/10.1021/acssensors.7b00362>.
- [250] N. Shrestha, S.L. Bryant, C. Thomas, D. Richtsmeier, X. Pu, J. Tinker, D. Fologea, Stochastic sensing of Angiotensin II with lysenin channels, *Sci Rep*. 7 (2017) 2448. <https://doi.org/10.1038/s41598-017-02438-0>.
- [251] A.J. Boersma, H. Bayley, Continuous Stochastic Detection of Amino Acid Enantiomers with a Protein Nanopore, *Angew. Chem*. 124 (2012) 9744–9747. <https://doi.org/10.1002/ange.201205687>.
- [252] A.J. Boersma, K.L. Brain, H. Bayley, Real-Time Stochastic Detection of Multiple Neurotransmitters with a Protein Nanopore, *ACS Nano*. 6 (2012) 5304–5308. <https://doi.org/10.1021/nn301125y>.
- [253] C. Wloka, V. Van Meervelt, D. van Gelder, N. Danda, N. Jager, C.P. Williams, G. Maglia, Label-Free and Real-Time Detection of Protein Ubiquitination with a Biological Nanopore, *ACS Nano*. 11 (2017) 4387–4394. <https://doi.org/10.1021/acsnano.6b07760>.
- [254] S. Lindsay, J. He, O. Sankey, P. Hapala, P. Jelinek, P. Zhang, S. Chang, S. Huang, Recognition tunneling, *Nanotechnology*. 21 (2010) 262001. <https://doi.org/10.1088/0957-4484/21/26/262001>.
- [255] T. Ohshiro, M. Tsutsui, K. Yokota, M. Furuhashi, M. Taniguchi, T. Kawai, Detection of post-translational modifications in single peptides using electron tunnelling currents, *Nature Nanotech*. 9 (2014) 835–840. <https://doi.org/10.1038/nnano.2014.193>.
- [256] C.L. Galindo, J. Sha, D.A. Ribardo, A.A. Fadl, L. Pillai, A.K. Chopra, Identification of *Aeromonas hydrophila* Cytotoxic Enterotoxin-induced Genes in Macrophages Using Microarrays, *J. Biol. Chem*. 278 (2003) 40198–40212. <https://doi.org/10.1074/jbc.M305788200>.
- [257] F.H. Caselitz, R. Guenther, Hemolysin studies with *Aeromonas* strains, *Zentralbl Bakteriol*. 180 (1960) 30–38.
- [258] A.W. Bernheimer, L.S. Avigad, Partial Characterization of Aerolysin, a Lytic Exotoxin from *Aeromonas hydrophila*, *Infection and Immunity*. 9 (1974) 1016–1021. <https://doi.org/10.1128/iai.9.6.1016-1021.1974>

- [259] J.T. Buckley, L.N. Halasa, K.D. Lund, S. MacIntyre, Purification and some properties of the hemolytic toxin aerolysin, *Can. J. Biochem.* 59 (1981) 430–435. <https://doi.org/10.1139/o81-059>.
- [260] S.P. Howard, J.T. Buckley, Molecular cloning and expression in *Escherichia coli* of the structural gene for the hemolytic toxin aerolysin from *Aeromonas hydrophila*, *Molec. Gen. Genet.* 204 (1986) 289–295. <https://doi.org/10.1007/BF00425512>.
- [261] S.P. Howard, W.J. Garland, M.J. Green, J.T. Buckley, Nucleotide sequence of the gene for the hole-forming toxin aerolysin of *Aeromonas hydrophila*, *Journal of Bacteriology*. 169 (1987) 2869–2871. <https://doi.org/10.1128/JB.169.6.2869-2871.1987>.
- [262] M.J. Green, J.T. Buckley, Site-directed mutagenesis of the hole-forming toxin aerolysin: studies on the roles of histidines in receptor binding and oligomerization of the monomer, *Biochemistry*. 29 (1990) 2177–2180.
- [263] J.T. Buckley, Purification of cloned proaerolysin released by a low protease mutant of *Aeromonas salmonicida*, *Biochemistry and Cell Biology*. 68 (1990) 221–224. <https://doi.org/10.1139/o90-029>.
- [264] S.P. Howard, J.T. Buckley, Protein export by a gram-negative bacterium: production of aerolysin by *Aeromonas hydrophila*, *Journal of Bacteriology*. 161 (1985) 1118–1124. <https://doi.org/10.1128/JB.161.3.1118-1124.1985>.
- [265] S.P. Howard, J.T. Buckley, Activation of the hole-forming toxin aerolysin by extracellular processing., *Journal of Bacteriology*. 163 (1985) 336–340. <https://doi.org/10.1128/JB.163.1.336-340.1985>.
- [266] Y. Wang, L.-Q. Gu, K. Tian, The aerolysin nanopore: from peptidomic to genomic applications, *Nanoscale*. 10 (2018) 13857–13866. <https://doi.org/10.1039/C8NR04255A>.
- [267] M.W. Parker, J.T. Buckley, J.P.M. Postma, A.D. Tucker, K. Leonard, F. Pattus, D. Tsernoglou, Structure of the *Aeromonas* toxin proaerolysin in its water-soluble and membrane-channel states, *Nature*. 367 (1994) 292–295. <https://doi.org/10.1038/367292a0>.
- [268] F.G. van der Goot, J. Ausio, K.R. Wong, F. Pattus, J.T. Buckley, Dimerization stabilizes the pore-forming toxin aerolysin in solution., *J. Biol. Chem.* 268 (1993) 18272–18279.
- [269] C. Lesieur, S. Frutiger, G. Hughes, R. Kellner, F. Pattus, F.G. van der Goot, Increased Stability upon Heptamerization of the Pore-forming Toxin Aerolysin, *J. Biol. Chem.* 274 (1999) 36722–36728. <https://doi.org/10.1074/jbc.274.51.36722>.
- [270] J. Rossjohn, J.T. Buckley, B. Hazes, A.G. Murzin, R.J. Read, M.W. Parker, Aerolysin and pertussis toxin share a common receptor-binding domain, *The EMBO Journal*. 16 (1997) 3426–3434. <https://doi.org/10.1093/emboj/16.12.3426>.
- [271] C.R. MacKenzie, T. Hiram, J.T. Buckley, Analysis of Receptor Binding by the Channel-forming Toxin Aerolysin Using Surface Plasmon Resonance, *J. Biol. Chem.* 274 (1999) 22604–22609. <https://doi.org/10.1074/jbc.274.32.22604>.

- [272] M.T. Degiacomi, I. Iacovache, L. Pernot, M. Chami, M. Kudryashev, H. Stahlberg, F.G. van der Goot, M. Dal Peraro, Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism, *Nat Chem Biol.* 9 (2013) 623–629. <https://doi.org/10.1038/nchembio.1312>.
- [273] F.G. Van der Goot, F. Pattus, K.R. Wong, J.T. Buckley, Oligomerization of the channel-forming toxin aerolysin precedes insertion into lipid bilayers, *Biochemistry.* 32 (1993) 2636–2642. <https://doi.org/10.1021/bi00061a023>.
- [274] J.T. Buckley, U. Wilmsen, C. Lesieur, A. Schulze, F. Pattus, M.W. Parker, F.G. van der Goot, Protonation of Histidine-132 Promotes Oligomerization of the Channel-Forming Toxin Aerolysin, *Biochemistry.* 34 (1995) 16450–16455. <https://doi.org/10.1021/bi00050a028>.
- [275] J.E. Alouf, D. Ladant, M.R. Popoff, *The comprehensive sourcebook of bacterial protein toxins*, Elsevier, 2005.
- [276] K.R. Wong, J.T. Buckley, Site-directed mutagenesis of a single tryptophan near the middle of the channel-forming toxin aerolysin inhibits its transfer across the outer membrane of *Aeromonas salmonicida*., *J. Biol. Chem.* 266 (1991) 14451–14456.
- [277] I. Iacovache, M.T. Degiacomi, L. Pernot, S. Ho, M. Schiltz, M. Dal Peraro, F.G. van der Goot, Dual Chaperone Role of the C-Terminal Propeptide in Folding and Oligomerization of the Pore-Forming Toxin Aerolysin, *PLoS Pathog.* 7 (2011) e1002135. <https://doi.org/10.1371/journal.ppat.1002135>.
- [278] I. Iacovache, P. Paumard, H. Scheib, C. Lesieur, N. Sakai, S. Matile, M.W. Parker, F.G. van der Goot, A rivet model for channel formation by aerolysin-like pore-forming toxins, *EMBO J.* 25 (2006) 457–466. <https://doi.org/10.1038/sj.emboj.7600959>.
- [279] C. Cao, M.-Y. Li, N. Cirauqui, Y.-Q. Wang, M. Dal Peraro, H. Tian, Y.-T. Long, Mapping the sensing spots of aerolysin for single oligonucleotides analysis, *Nat Commun.* 9 (2018) 2823. <https://doi.org/10.1038/s41467-018-05108-5>.
- [280] R. G. Landes Company, M.W. Parker, J.T. Buckley, F.G. van der Goot, D. Tsernoglou, Structure and Assembly of the Channel-Forming *Aeromonas* Toxin Aerolysin, in: *Protein Toxin Structure*, Springer-Verlag, Berlin, Heidelberg, 1996: pp. 79–95. https://doi.org/10.1007/978-3-662-22352-9_5.
- [281] H.U. Wilmsen, F. Pattus, J.T. Buckley, Aerolysin, a hemolysin from *Aeromonas hydrophila*, forms voltage-gated channels in planar lipid bilayers, *J. Membrin Biol.* 115 (1990) 71–81. <https://doi.org/10.1007/BF01869107>.
- [282] T. Chakraborty, A. Schmid, S. Notermans, R. Benz, Aerolysin of *Aeromonas sobria*: evidence for formation of ion-permeable channels and comparison with alpha-toxin of *Staphylococcus aureus*., *Infection and Immunity.* 58 (1990) 2127–2132.
- [283] M. Boukhet, F. Piguet, H. Ouldali, M. Pastoriza-Gallego, J. Pelta, A. Oukhaled, Probing driving forces in aerolysin and α -hemolysin biological nanopores: electrophoresis versus electroosmosis, *Nanoscale.* 8 (2016) 18352–18359. <https://doi.org/10.1039/C6NR06936C>.

- [284] B. Cressiot, E. Braselmann, A. Oukhaled, A.H. Elcock, J. Pelta, P.L. Clark, Dynamics and Energy Contributions for Transport of Unfolded Pertactin through a Protein Nanopore, *ACS Nano*. 9 (2015) 9050–9061. <https://doi.org/10.1021/acsnano.5b03053>.
- [285] M. Pastoriza-Gallego, B. Thiébot, L. Bacri, L. Auvray, J. Pelta, Dynamics of a polyelectrolyte through aerolysin channel as a function of applied voltage and concentration*, *Eur. Phys. J. E*. 41 (2018) 58. <https://doi.org/10.1140/epje/i2018-11661-4>.
- [286] L. Payet, M. Martinho, C. Merstorf, M. Pastoriza-Gallego, J. Pelta, V. Viasnoff, L. Auvray, M. Muthukumar, J. Mathé, Temperature Effect on Ionic Current and ssDNA Transport through Nanopores, *Biophysical Journal*. 109 (2015) 1600–1607. <https://doi.org/10.1016/j.bpj.2015.08.043>.
- [287] Z.-L. Hu, Z.-Y. Li, Y.-L. Ying, J. Zhang, C. Cao, Y.-T. Long, H. Tian, Real-Time and Accurate Identification of Single Oligonucleotide Photoisomers via an Aerolysin Nanopore, *Anal. Chem.* 90 (2018) 4268–4272. <https://doi.org/10.1021/acs.analchem.8b00096>.
- [288] C. Cao, J. Yu, Y.-Q. Wang, Y.-L. Ying, Y.-T. Long, Driven Translocation of Polynucleotides Through an Aerolysin Nanopore, *Anal. Chem.* 88 (2016) 5046–5049. <https://doi.org/10.1021/acs.analchem.6b01514>.
- [289] C.G. Rodrigues, D.C. Machado, S.F. Chevtchenko, O.V. Krasilnikov, Mechanism of KCl Enhancement in Detection of Nonionic Polymers by Nanopore Sensors, *Biophysical Journal*. 95 (2008) 5186–5192. <https://doi.org/10.1529/biophysj.108.140814>.
- [290] M. Boukhet, N.F. König, A.A. Ouahabi, G. Baaken, J.-F. Lutz, J.C. Behrends, Translocation of Precision Polymers through Biological Nanopores, *Macromol. Rapid Commun.* 38 (2017) 1700680. <https://doi.org/10.1002/marc.201700680>.
- [291] A. Fennouri, C. Przybylski, M. Pastoriza-Gallego, L. Bacri, L. Auvray, R. Daniel, Single Molecule Detection of Glycosaminoglycan Hyaluronic Acid Oligosaccharides and Depolymerization Enzyme Activity Using a Protein Nanopore, *ACS Nano*. 6 (2012) 9672–9678. <https://doi.org/10.1021/nn3031047>.
- [292] C. Cao, J. Yu, M.-Y. Li, Y.-Q. Wang, H. Tian, Y.-T. Long, Direct Readout of Single Nucleobase Variations in an Oligonucleotide, *Small*. 13 (2017) 1702011. <https://doi.org/10.1002/smll.201702011>.
- [293] S. Li, C. Cao, J. Yang, Y.-T. Long, Detection of peptides with different charges and lengths by using the aerolysin nanopore, *ChemElectroChem*. 6 (2019) 126–129. <https://doi.org/10.1002/celec.201800288>
- [294] Y.-L. Ying, J. Yang, F.-N. Meng, S. Li, M.-Y. Li, Y.-T. Long, A Nanopore Phosphorylation Sensor for Single Oligonucleotides and Peptides, *Research*. 2019 (2019) 1–8. <https://doi.org/10.34133/2019/1050735>.
- [295] J. Wang, M.-Y. Li, J. Yang, Y.-Q. Wang, X.-Y. Wu, J. Huang, Y.-L. Ying, Y.-T. Long, Direct Quantification of Damaged Nucleotides in Oligonucleotides Using an Aerolysin

- Single Molecule Interface, ACS Cent. Sci. 6 (2020) 76–82. <https://doi.org/10.1021/acscentsci.9b01129>.
- [296] N.A. Shumilova, G.V. Zhutaeva, Silver, in: Encyclopedia of Electrochemistry of the Elements, Marcel Dekker, New York, 1978: pp. 106–110.
- [297] B.F.E. Matarèse, P.L.C. Feyen, A. Falco, F. Benfenati, P. Lugli, J.C. deMello, Use of SU8 as a stable and biocompatible adhesion layer for gold bioelectrodes, Sci Rep. 8 (2018) 5560. <https://doi.org/10.1038/s41598-018-21755-6>.
- [298] D. Bitounis, R. Fanciullino, A. Iliadis, J. Ciccolini, Optimizing Druggability through Liposomal Formulations: New Approaches to an Old Concept, ISRN Pharmaceuticals. 2012 (2012) e738432. <https://doi.org/10.5402/2012/738432>.
- [299] I. Langmuir, The constitution and fundamental properties of solids and liquids. II. Liquids., Journal of the American Chemical Society. 39 (1917) 1848–1906.
- [300] E. Gorter, F. Grendel, On bimolecular layers of lipoids on the chromocytes of the blood, J Exp Med. 41 (1925) 439–443. <https://doi.org/10.1084/jem.41.4.439>
- [301] D. Marquardt, F.A. Heberle, T. Miti, B. Eicher, E. London, J. Katsaras, G. Pabst, ¹H NMR Shows Slow Phospholipid Flip-Flop in Gel and Fluid Bilayers, Langmuir. 33 (2017) 3731–3741. <https://doi.org/10.1021/acs.langmuir.6b04485>.
- [302] F.A. Henn, T.E. Thompson, Synthetic Lipid Bilayer Membranes, Annu. Rev. Biochem. 38 (1969) 241–262. <https://doi.org/10.1146/annurev.bi.38.070169.001325>.
- [303] K. Sharp, A. Nicholls, R. Fine, B. Honig, Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects, Science. 252 (1991) 106–109. <https://doi.org/10.1126/science.2011744>.
- [304] H.T. Tien, A. Ottova-Leitmannova, Planar Lipid Bilayers (BLM's) and Their Applications, Elsevier, 2003.
- [305] R. Guidelli, L. Becucci, Model Lipid Bilayers at Electrode Surfaces, in: Advances in Electrochemical Science and Engineering, John Wiley & Sons, Ltd, n.d.: pp. 189–227. <https://doi.org/10.1002/9783527644117.ch4>.
- [306] S. Kara, S. Afonin, O. Babii, A.N. Tkachenko, I.V. Komarov, A.S. Ulrich, Diphytanoyl lipids as model systems for studying membrane-active peptides, Biochimica et Biophysica Acta (BBA) - Biomembranes. 1859 (2017) 1828–1837. <https://doi.org/10.1016/j.bbamem.2017.06.003>.
- [307] D.C. Golibersuch, Observation of Aspherical Particle Rotation in Poiseuille Flow via the Resistance Pulse Technique: I. Application to Human Erythrocytes, Biophysical Journal. 13 (1973) 265–280. [https://doi.org/10.1016/S0006-3495\(73\)85984-3](https://doi.org/10.1016/S0006-3495(73)85984-3).
- [308] J.C. Maxwell, A Treatise on Electricity and Magnetism, Clarendon Press, 1873.

- [309] A. Oukhaled, L. Bacri, M. Pastoriza-Gallego, J.-M. Betton, J. Pelta, Sensing Proteins through Nanopores: Fundamental to Applications, *ACS Chem. Biol.* 7 (2012) 1935–1949. <https://doi.org/10.1021/cb300449t>.
- [310] T.Z. Butler, J.H. Gundlach, M.A. Troll, Determination of RNA Orientation during Translocation through a Biological Nanopore, *Biophysical Journal*. 90 (2006) 190–199. <https://doi.org/10.1529/biophysj.105.068957>.
- [311] S.J. Perkins, Protein volumes and hydration effects, *European Journal of Biochemistry*. 157 (1986) 169–180. <https://doi.org/10.1111/j.1432-1033.1986.tb09653.x>.
- [312] F. Piguet, H. Ouldali, M. Pastoriza-Gallego, P. Manivet, J. Pelta, A. Oukhaled, Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore, *Nat Commun.* 9 (2018) 966. <https://doi.org/10.1038/s41467-018-03418-2>.
- [313] N.S. Galenkamp, M. Soskine, J. Hermans, C. Wloka, G. Maglia, Direct electrical quantification of glucose and asparagine from bodily fluids using nanopores, *Nat Commun.* 9 (2018) 4085. <https://doi.org/10.1038/s41467-018-06534-1>.
- [314] S. Zernia, N.J. van der Heide, N.S. Galenkamp, G. Gouridis, G. Maglia, Current Blockades of Proteins inside Nanopores for Real-Time Metabolome Analysis, *ACS Nano*. 14 (2020) 2296–2307. <https://doi.org/10.1021/acsnano.9b09434>.
- [315] B. Yuan, S. Li, Y.-L. Ying, Y.-T. Long, The analysis of single cysteine molecules with an aerolysin nanopore, *Analyst*. 145 (2020) 1179–1183. <https://doi.org/10.1039/C9AN01965K>.
- [316] F.-N. Meng, Y.-L. Ying, J. Yang, Y.-T. Long, A Wild-Type Nanopore Sensor for Protein Kinase Activity, *Anal. Chem.* 91 (2019) 9910–9915. <https://doi.org/10.1021/acs.analchem.9b01570>.
- [317] L. Restrepo-Pérez, C.H. Wong, G. Maglia, C. Dekker, C. Joo, Label-Free Detection of Post-translational Modifications with a Nanopore, *Nano Lett.* 19 (2019) 7957–7964. <https://doi.org/10.1021/acs.nanolett.9b03134>.
- [318] L. Restrepo-Pérez, G. Huang, P.R. Bohländer, N. Worp, R. Eelkema, G. Maglia, C. Joo, C. Dekker, Resolving Chemical Modifications to a Single Amino Acid within a Peptide Using a Biological Nanopore, *ACS Nano*. 13 (2019) 13668–13676. <https://doi.org/10.1021/acsnano.9b05156>.
- [319] H. Ouldali, A. Oukhaled, Interaction of Cucurbituril Molecular Containers with the Aerolysin Nanopore for Molecular Recognition, *Biophysical Journal*. 118 (2020) 473a–474a.

Glossaire des abréviations

ADN : acide désoxyribonucléique

ARN : acide ribonucléique

ARNm : acide ribonucléique messenger

ATP : Adénosine triphosphate

Al₂O₃ : Oxyde d'aluminium

A1 : angiotensine I

A2 : angiotensine II

Ag : Argent

Å : Angstrom

BtuB : Binding to vitamin B12 porin

BN : Nitrure de bore

BLM : Black lipid membrane

CFTR : Cystic Fibrosis Transmembrane Regulator

ClpX : Caseinolytic Mitochondrial Matrix Peptidase Chaperone Subunit X

ClpP : caseinolytic protease proteolytic subunit

Cy3 : Cyanine 3

CTP : C-terminal peptide

cryo-EM : microscopie cryo-électronique atomique

ClyA : Cytolysine A

CsgG : Curli production assembly/transport component

°C : degré Celsius

dNTP : désoxynucléoside triphosphate

dATP : désoxyadénosine triphosphate

dCTP : désoxycytidine triphosphate

dGTP : désoxyguanosine triphosphate

dTTP : désoxythymidine triphosphate

Da : Dalton

DPhPC : 1,2-diphytanoyl-sn-glycero-3-phosphocholine

EpCAM : Epithelial Cell Adhesion Molecule

EGF : Epidermal Growth Factor

ET-1 : Endothéline 1

ET-2 : Endothéline 2

FRET : **Fluorescence** resonance energy transfer

FraC : Fragaceatoxin

FhuA : Ferrichrome porin A

gp10 : gene product 10

g/mol : gramme par mole

Gdn-HCl : Guanidium Hydrochloride

GPI : glycosyl-phosphatidylinositol

HPLC : chromatographie liquide à haute performance

HEPES : acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique

HCl : acide chlorhydrique

IPTG : isopropyl β -D-1-thiogalactopyranoside

KCl : chlorure de potassium

KF : Klenow fragment

kb : kilobase

kDa : kiloDalton

KOH : hydroxyde de potassium

kHz : kiloHertz

KNO_3^- : bromure de potassium

LiCl : chlorure de lithium

LamB : bacteriophage lambda receptor porin

LE : encéphalite

LC-MS : chromatographie liquide couplé à la spectrométrie de masse

M : mol/L ou moles par litre

MOF : Metal Organic Framework

MspA : Mycobacterium segmatis protein A

MoS_2 : Disulfure de molybdène

mV : millivolt

MALDI : Matrix-assisted laser desorption

mg : milligramme

MECA : Multi-Electrode Channel Array

mg/mL : milligramme par millilitre

μm : micromètre

Np-SMMS : Nanopore-based Single Molecule Mass Spectrometry

Np-SMSD : Nanopore-based Single-Molecule Size- Discrimination

NanoTag-SBS : NanoTag-based real- time Sequencing By Synthesis

NT : neurotensine

NaCl : chlorure de sodium

NO_2 : dioxyde d'azote

nm : nanomètre

NfpA/B : Nocardia farcinica porin A/B

Omp : Outer membrane protein

PCR : Polymerase Chain Reaction

pN : picoNewton

pS : picoSiemens

PTFE : polytétrafluoroéthylène

pF : picoFarad

pHi : potentiel hydrogène isoélectrique

PET : polyéthylène téréphtalate

PEG : polyéthylène glycol

pA : picoAmpère

PFT : pore-forming toxin

pH : potentiel hydrogène

PhoE : phosphoprotein E

% : pour cent

QBP1 : polyglutamine binding peptide

RPS : Resistive Pulse Sensing

SiN : Nitrure de silicium

TAT : transactivator of transcription

Tris : tris(hydroxyméthyl) aminométhane

TOF : Time of flight

Ti : temps inter-événements

Ub : Ubiquitine

VDAC : Voltage-dependent anionic channel

V : Volt

ZnO : Oxyde de zinc

Table des figures

Figure 1 - Schéma représentatif du principe du compteur Coulter.	8
Figure 2 - Schéma représentatif des étapes de formation d'un pore par une PFT.	13
Figure 3 - Représentation de quelques nanopores biologiques communément utilisés ...	18
Figure 4 - Représentation de quelques exemples de nanopores solides et hybrides	21
Figure 5 - Schéma représentatif d'un montage expérimental de détection à l'échelle moléculaire par un nanopore biologique inséré dans une membrane de lipides.	22
Figure 6 - Discrimination en taille de polymères synthétiques avec le nanopore d' α -hémolysine.	24
Figure 7 - Discrimination de molécules d'ADN en épingles à cheveux (DNA hairpin) avec une résolution d'une paire de base, obtenue avec le nanopore d' α -hémolysine.....	26
Figure 8 - Détection du transport d'homopolymères simple brin d'ADN avec le nanopore d' α -hémolysine.	27
Figure 9 - Discrimination d'homopolymères d'ADN et d'ARN avec le nanopore d' α -hémolysine.	29
Figure 10 - Représentation schématique du principe de lecture séquentielle d'une chaîne d'ADN dépliée à travers un nanopore.	30
Figure 11 - Schéma représentatif des nanopores biologiques MspA et α -hémolysine et de leurs zones de sensibilité.	32
Figure 12 - Représentation schématique du principe de contrôle du transport d'une chaîne d'ADN par nanopore assisté par un moteur moléculaire.....	33
Figure 13 - Représentation schématique du principe de lecture séquentielle d'une chaîne d'ADN clivée base par base à l'entrée d'un nanopore.....	34
Figure 14 - Approche de discrimination des bases nucléotidiques par l'utilisation d'une β -cyclodextrine comme adaptateur moléculaire associé au nanopore d' α -hémolysine.	35
Figure 15 - Approche de séquençage d'ADN par marquage et synthèse en temps réel. .	37
Figure 16 - Stratégie de liaison d'un oligonucléotide à une protéine pour son transport unidirectionnel à travers le nanopore d' α -hémolysine.	44
Figure 17 - Stratégie de contrôle de la translocation d'une protéine à travers un nanopore par association d'un moteur moléculaire au nanopore	46

Figure 18 - Discrimination de peptides biologiques par le nanopore d'α-hémolysine en présence ou non de nanoparticules d'or.....	49
Figure 19 - Discrimination de peptides par le nanopore FraC sauvage à pH 4,5. (b) Séquence en acides aminés des peptides endothéline 1 (ET-1) et endothéline 2 (ET-2)	50
Figure 20 - Stratégie d'identification des acides aminés par effet courant tunnel. Schéma représentatif de la mesure par effet tunnel.....	55
Figure 21 - Stratégie d'identification d'une protéine par la méthode FRET fingerprinting.	57
Figure 22 - Stratégie de séquençage des protéines par la méthode de fluoroséquençage basée sur la dégradation d'Edman.	58
Figure 23 - Représentation de la structure et de la séquence de la pro-aérolysine de <i>Aeromonas hydrophila</i>.	63
Figure 24 - Représentation schématique d'un monomère de pro-aérolysine lié à son récepteur membranaire à GPI par ses différents domaines..	64
Figure 25 - Représentation schématique des différentes étapes du mécanisme d'insertion du pore d'aérolysine dans une membrane lipidique	67
Figure 26 - Représentation du nanopore d'aérolysine produit par <i>Aeromonas hydrophila</i>	68
Figure 27 - Modèle quasi-atomique de la structure du pore d'aérolysine sauvage montrant les résidus impliqués dans l'ancrage du pore dans la membrane et les résidus chargés pointant vers la lumière du pore	69
Figure 28 - Discrimination de molécules de PEG avec une résolution de séparation d'un monomère, obtenue avec le nanopore d'aérolysine sauvage.	73
Figure 29 - Discrimination de courts oligonucléotides avec une résolution de séparation d'un nucléotide avec le nanopore d'aérolysine sauvage.	75
Figure 30 - Discrimination d'oligonucléotides d'adénine présentant une substitution d'un nucleotide individuel	76
Figure 31 - Schéma d'un montage expérimental classique d'enregistrement électrique par nanopore.....	82
Figure 32 - Schéma d'une micropuce électronique à multiples cavités, utilisée pour l'enregistrement électrique simultané de signaux provenant de multiples nanopores.	85
Figure 33 - Représentation schématique et structure chimique d'un phospholipide et de son arrangement dans une bicouche de lipides.	86

Figure 34 - Modes d'organisation des lipides. Schéma représentatif d'une bicouche de lipides sphérique (a), en micelle (b), ou plane (c)	87
Figure 35 - Schéma représentatif des mouvements de diffusion des lipides dans une membrane	88
Figure 36 - Formation d'une bicouche lipidique plane par la méthode Montal-Mueller.	89
Figure 37 - Rectification du courant dans le nanopore d'aérolysine.....	92
Figure 38 - Schéma représentatif du principe de la résistance électrique d'un cylindre en présence d'une bille isolante	93
Figure 39 - Méthode d'analyse utilisée pour la détection d'événements de blocage du courant à partir d'un enregistrement électrique.....	99
Figure 40 - Principe de l'analyse d'une population polydisperse par un nanopore.	101
Figure 41 – Exemple d'une analyse statistique des durées et des fréquences des blocages du courant.	103
Figure 42 - Schéma illustratif du dispositif expérimental général.....	108
Figure 43 - Détection d'homopeptides d'arginine d'une longueur de 9 acides aminés avec l'aérolysine	109
Figure 44 - Analyse statistique de la trace de courant électrique obtenue pour l'analyse d'homopeptides d'une longueur de 9 acides aminés avec l'aérolysine.....	110
Figure 45 - Analyse d'homopeptides d'une longueur de 9 acides avec le nanopore d'aérolysine	112
Figure 46 - Analyse individuelle d'homopeptides d'une longueur de 5 à 10 acides aminés avec le nanopore d'aérolysine	115
Figure 47 - Détection d'homopeptides d'arginine avec le nanopore d'aérolysine en fonction du côté du nanopore dans lequel sont ajoutés les peptides, et en fonction du signe de la tension électrique appliquée.	116
Figure 48 - Détection d'un mélange d'homopeptides d'une longueur de 5 à 10 acides aminés avec l'aérolysine.....	117
Figure 49 - Analyse d'un mélange d'homopeptides d'arginine d'une longueur de 5 à 10 acides aminés avec l'aérolysine	118
Figure 50 - Reproductibilité des résultats expérimentaux.	119

Figure 51 - Effet de la tension électrique sur l'interaction du nanopore d'aérollysine avec un mélange d'homopeptides d'arginine de différentes longueurs	121
Figure 52 - Effet du voltage sur la discrimination en taille d'homopeptides d'arginine de différentes longueurs.....	122
Figure 53 - Expérience de suivi en temps réel de la dégradation enzymatique par la trypsine d'un échantillon d'homopeptides d'arginine avec le nanopore d'aérollysine..	124
Figure 54 - Détection et identification d'homopeptides d'arginine d'une longueur de 3 et 12 acides aminés..	125
Figure 55 - Analyse d'homopeptides de lysine d'une longueur de 10 acides aminés avec le nanopore biologique d'aérollysine.	128
Figure 56 - Analyse d'hétéropeptides avec le nanopore d'aérollysine	130
Figure 57 - Discrimination de peptides d'une longueur de 10 acides aminés de séquences différentes avec l'aérollysine	131
Figure 58 - Effet de la concentration en peptides sur l'interaction du nanopore d'aérollysine avec des homopeptides d'arginine de différentes longueurs.	133
Figure 59 - Schéma illustratif du dispositif expérimental	136
Figure 60 - Détection et analyse d'un peptide X_{R7} avec l'aérollysine	137
Figure 61 - Analyse des peptides X_{R7} avec l'aérollysine.	138
Figure 62 - Vue d'ensemble des valeurs du courant relatif résiduel I_b/I_0 pour les peptides X_{R7} analysés avec le nanopore d'aérollysine..	140
Figure 63 - Analyse des peptides C_{R7} en présence ou en absence d'un agent réducteur de ponts disulfures.	141
Figure 64 - Discrimination de peptides individuels au sein d'un mélange avec le nanopore d'aérollysine.	144
Figure 65 - Identification de peptides individuels au sein d'un mélange avec le nanopore d'aérollysine.	145
Figure 66 - Identification de peptides présentant des différences subtiles avec le nanopore d'aérollysine	146
Figure 67 - Dépendance du courant relatif résiduel I_b/I_0 moyen à la masse molaire ou au volume hydrodynamique de l'acide aminé X	148
Figure 68 - Expériences de simulation dans l'aérollysine	150

Figure 69 - Influence de la tension appliquée sur la discrimination des peptides X_{R7}...	151
Figure 70 – Simulation du transport de peptides dans le pore d’aérolysine	153
Figure 71 - Effet du temps de résidence sur la discrimination des peptides X_{R7}.....	154
Figure 72 - Identification de peptides présentant des modifications chimiques avec le nanopore d’aérolysine.....	156
Figure 73 - Vue d'ensemble de la discrimination des acides aminés protéinogènes et modifiés chimiquement.	157
Figure 74 - Discrimination des peptides L_{R7} et I_{R7} dans des conditions expérimentales d'enregistrement optimisées.	158

Annexes