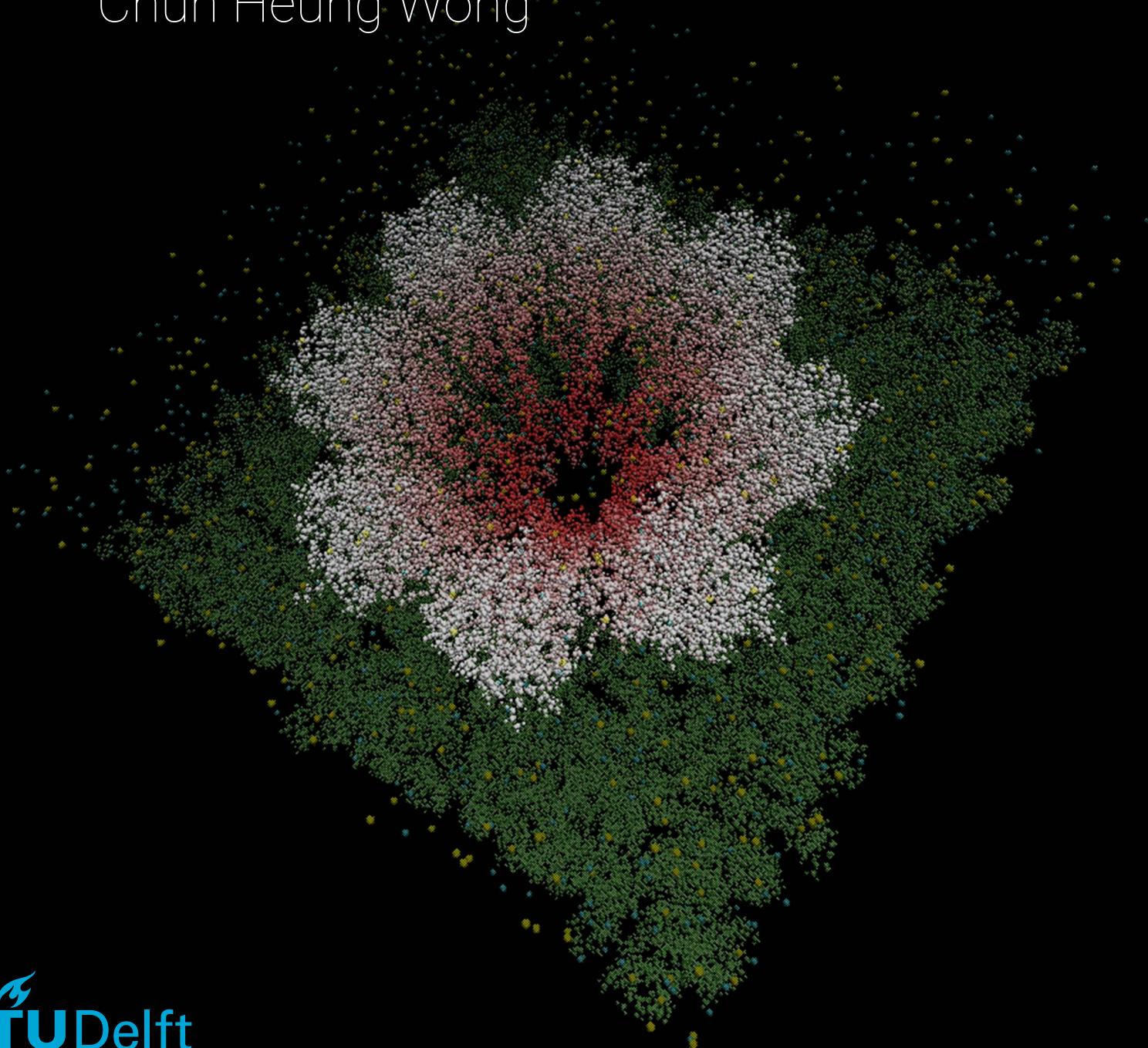


Detection of Post Translational Modifications using Biological Nanopores

Master of Science Thesis
Chun Heung Wong



Detection of Post Translational Modifications using Biological Nanopores

Master of Science Thesis

by

Chun Heung Wong

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday August 31st, 2018 at 9:30h,
in the Kronigzaal, TNW-Zuid.

Student number: 4296141
Project duration: September 9, 2017 – August 31, 2018
Thesis committee:
Prof. dr. C. Dekker, TU Delft, supervisor, promotor
Assoc. Prof. dr. C. Joo, TU Delft, supervisor
Assoc. Prof. dr. R. Eelkema, TU Delft, committee member
L. Restrepo-Perez, MSc. TU Delft, daily supervisor



Contents

1	Introduction	1
2	Theoretical Background	3
2.1	Protein synthesis and PTMs	3
2.1.1	Acetylation.	4
2.1.2	Phosphorylation	5
2.1.3	Glycosylation	5
2.2	Nanopore sensing technique	7
2.3	Solid state and biological nanopores	8
2.3.1	PTM detection with Nanopores	8
2.3.2	Fragaceatoxin C nanopore	9
2.3.3	Noise.	10
2.4	Theoretical model of transport in Nanopores	11
2.4.1	Charge transport through nanopores	11
2.4.2	Surface charge effects	13
2.4.3	Pore assymetry: biased current and flow rectification	13
2.4.4	Application of theory on FraC	14
2.4.5	Capture rate of particles	17
2.4.6	Dwell-time distributions.	18
3	Material and Methods	19
3.1	Nanopore translocation.	19
3.1.1	Lipid bilayer formation	20
3.1.2	Nanopore insertion	20
3.1.3	Silver/Silver-Chloride electrodes.	20
3.1.4	Translocation measurements	21
3.1.5	Data acquisition and analysis	21
3.2	Spectrophotometer unfolding studies.	22
3.3	Sample preparation.	22
3.3.1	Peptide D	22
3.3.2	Arg-adaptor Ligation.	22
3.3.3	RNAse A, B and glycans	23
3.4	Materials	23
4	Model Peptide: Peptide D	25
4.1	Peptide D characterization	26
4.2	Distinguishing Peptide D from PTM'ed variants	28
4.3	Distinguishing O-GlcNAc and Phosphorylated variants.	30
5	Arginine-Adaptor: Proof-of-concept	33
5.1	Ligation and purification	34
5.2	Experimental effect of adaptor region.	36
6	N-linked Glycosylation: Ribonuclease A and B	39
6.1	RNAse A and B: Native and acidic state	41
6.2	Translocating Purified Glycans	45
6.3	Unfolding of Ribonuclease	46
6.4	Translocating Ribonuclease in denaturing conditions.	48

7 Discussion and outlook	53
8 Conclusion	55
9 Acknowledgements	57
A Translayzer Changes	59
B Supplementary experiments Peptide D	61
B.1 Repeat of Glycosylation, Phosphorylation and unmodified mixture.	61
B.2 High-low salt measurements PepD	63
C Ribonuclease A and B characterization	65
C.0.1 Mass spectra	65
C.0.2 Unfolding Ribonuclease A in Guanidinium HCl	66
Bibliography	67

1

Introduction

In the pursuit of a detailed understanding of the molecular mechanisms at work in living organisms, one quickly arrives at the study of proteins. These organic polymers can be seen as the molecular machines of living systems, which enable organisms to exhibit a spectacular array of characteristics and functionalities. From the copying of genetic material [5] to absorbing mechanical shocks [121], organisms employ protein for a host of different processes, including cell signalling, recognition, cell-differentiation, gene regulation, cell structure and many more [5, 44]. To have such a diverse set of functionalities, organisms need a way to create an equally diverse array of different molecular machines.

This study focusses on one of methods organisms employ to hypothetically expand the pool of unique proteins to billions: Post-Translational modifications (PTM) [2, 96]. This collective term refers to all modifications occurring after the translational synthesis of proteins, and are usually small chemical groups. Despite their size, these PTMs have been found to play vital roles in protein (in)activation, stability and recognition among other things [31, 33, 58, 70]. In the medical field in particular, PTMs have been found to play key roles in pathogenic pathways for cancer, Parkinsons disease, Alzheimers disease and diabetes [23, 27, 32].

As a result, it is pivotal to detect PTM variants of proteins with high sensitivity. The preferred method of PTM detection at the time of writing is mass spectrometry (MS) [69, 90]. This technique detects and separates molecules based on their respective mass to charge ratio (m/z) using electromagnetic fields. With the advent of soft ionization methods such as electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), it is possible to analyse full length proteins without (unintentionally) fractionating them at various positions during ionization [51]. This makes MS an efficient tool for studying complex mixtures of intact proteins. Modern automated data analysis software eases the intricate data analysis steps that can be challenging for newcomers and experts alike [14]. Some major achievements of the MS technique in PTM research is the successful analysis of 6000 full proteins and the detection of over 20,000 unique phosphorylation sites in human HeLa S3 cell-extracts [113].

Nevertheless, MS has its limitations, mainly brought about by the immense dynamic range in protein copy-numbers present *in-vivo*. These easily span over 10 orders of magnitude [2, 96]. Challenges arise when low-copy number proteoforms are of interest in a sea of others, especially because proteins not only have different kinds of PTMs, but also different quantities of the same PTMs. Extremely sparse proteins are practically invisible in the noise of a diverse mixture of abundant proteins [30]. In fact, many bio-markers for cancers and other diseases escape timely detection for this very reason [42, 100, 106]. To alleviate this problem, pre-purification is used to enrich samples of interest [2]. Examples include high-performance liquid-chromotagoraphy (HPLC) for purification based on hydrophobicity or titaniumdioxide columns and antibody binding of lectins that can specifically bind to a class of PTMed proteins [140]. Other approaches include the PTM-specific removal and subsequent chemical labelling of PTMs, enabling the researcher to detect it by characteristic mass-shifts in a mass spectrum [51, 84]. These enrichment and labelling methods are time-consuming and highly PTM specific, which discourages its use in a blanket analysis approach sought for in clinical research.

Not only the presence, but also the location of PTMs along the amino acid chain are of biological importance [71, 107, 121]. Reading the m/z value of a full PTMed protein can in principle shed light on the type and degree of a given PTM, but can in no way elucidate exactly which residue(s) was PTMed. Generally, the proteins are therefore broken down to pieces, small enough to recognize directly. The amino acid sequence

is then obtained from a bottom-up approach, giving the PTM's location [70, 113]. However, this inherently complicates the resulting mass spectra by spawning more species, thus generating more m/z peaks.

It is exactly for these reasons that we aim to employ a single-molecule technique for PTM detection, capable of individually probing each protein and thereby its PTM. This hypothetically enables us to overcome the dynamic range problem, as given enough time, all target proteins will be individually characterized. Starting relatively recently, existing nanopore sensing technology was applied for this purpose [34, 104]. Nanometer scale pores are used as extremely small ion channels, allowing pA currents to flow upon applying an electric field. The nanopore sensing technique then principally relies on the incredible sensitivity of the pore to any (partial) blocking of its cross-section [29, 86, 132]. Analytes are pulled near or through the pore as a result of the electric field, giving rise to characteristic ionic current modulations. In this way, it is possible to distinguish molecules based on volumetric size, but also charge and hydrophobicity/hydrophilicity [47, 86]. These features make nanopore sensing an excellent candidate for the detection small chemical moieties such as PTMs in a label-free manner [47, 119]. Major contributions of nanopores in the field of molecular biology to date are the direct sequencing of individual DNA molecules [11, 46, 132], probing protein-protein interactions [82], protein folding-unfolding [87, 119], and even chemical reaction rates [75], enabling researchers to answer biophysics questions on the molecular scale.

In overview of this work, we aim to expand the current knowledge on the capability of biological nanopores in distinguishing peptides and proteins from their PTM'ed variants *in-vitro*. In Chapter 2, we discuss the theoretical background of this study. We first address the biological context of PTMs acetylation, phosphorylation and glycosylation, together with their biological relevance. We then move on to discuss the nanopore setup in general, and theoretical work done so far to understand it from the perspective of physical models and simulations. We finally apply this knowledge to the specific experimental conditions of this work. The materials and methods are detailed in Chapter 3. We then present the experimental results in three chapters, corresponding to the three different categories of protein and peptide samples studied. At the start of each section, we will give sample specific details and theory. We will also discuss the results with a focus on data-interpretation. We start with a model peptide in Chapter 4, specifically designed to increase measurement duration and regional sensitivity of the nanopore. We present the first reported successful attempt at distinguishing a phosphorylated and O-GlcNAc glycosylated proteoform using nanopores, which was shown to be relevant for DNA transcription *in-vivo* [121]. In Chapter 5 we present the successful application of a prototype adapter region, enabling an uncharged peptide to be analyzed by nanopores. In Chapter 6 we show our efforts on unfolding and translocating a full-length protein RNase A, with the goal of distinguishing it from its PTM'ed variant, Ribonuclease B. In Chapter 7 we discuss and compare our results with respect to the existing literature, and give a future outlook on possible expansions on this work. Finally, we give concluding remarks in Chapter 8.

2

Theoretical Background

In this chapter, an overview is given of the theoretical background relevant to the analysis of PTMs on protein and peptides using nanopores. We discuss the biological synthesis pathways of the three main PTMs modifications of this work, acetylation, phosphorylation and glycosylation. Their biological role is explained briefly cases of medical relevance are given. We then introduce the nanopore sensing technique, and discuss the advantages of biological and solid state nanopores respectively. We then give a brief literature review of the current state of nanopore-based PTM detection. Finally, we study the physical background of charge and analyte transport through nanopores, and present our efforts of applying the theory to our experimental setup.

2.1. Protein synthesis and PTMs

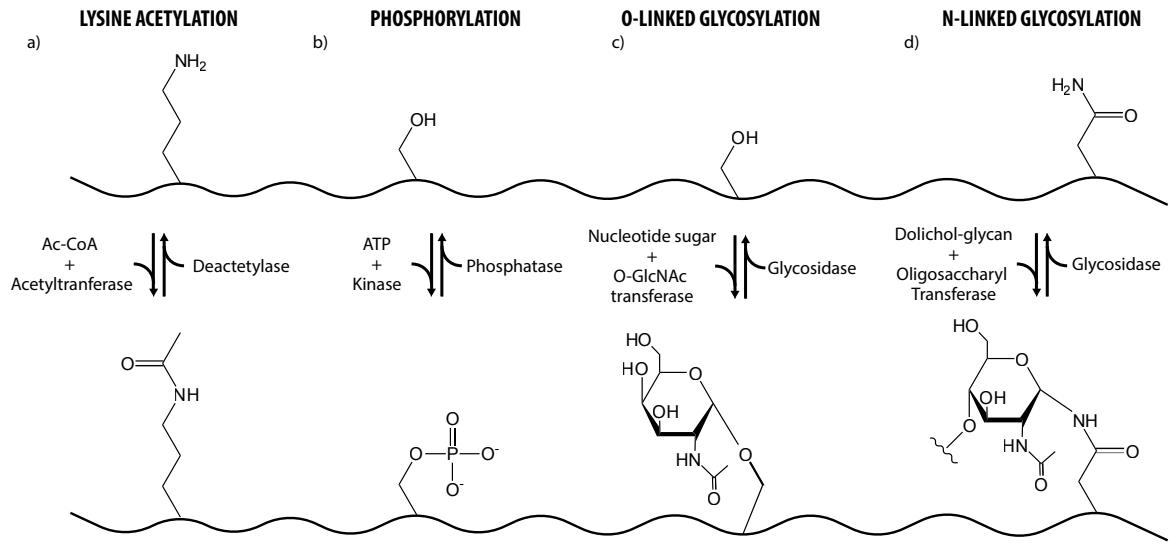


Figure 2.1: Schematic overview of the PTMs relevant to this study. The jagged line represents a random peptides. (Top) peptide with unmodified residues. (Bottom) PTM'ed peptide. The double arrows include the corresponding enzymes responsible for the modifications.

Proteins are organic polymers built from a set of twenty different amino acids arranged in a linear order. The sequential arrangement of the amino acids is then the primary factor in deciding the 3D folded state of the molecule, as it moves in an energy landscape determined mainly by electrostatic, hydrophilic/hydrophobic and Vanderwaals interactions between the various amino acids[44, 49]. Ultimately, the 3D structure will then give the molecule its functionality, as the chemical groups of the amino acids often need to be spatially co-localized to perform their function [44].

Given that proteins of up to 30,000 amino acids have been found, it should be remarked that the number of unique amino acid permutations seems implausibly high. Naturally, many of these random sequences

will not be biologically useful, necessitating a way for the organism to remember which are. The information required for this specificity is stored in the genetic code or DNA of the organism. Using numerous protein complexes, the DNA molecules can be pried apart, read and translated into an amino acid sequence. It is known that a set of three DNA nucleobases can unique encode for one amino acid, effectively making DNA molecules function as protein templates [5].

Human genomes have been found to contain roughly 20,000 protein-encoding genes [2, 96], which might lead one to conclude that the human cells can express an equally large amount of unique protein sequences. Surprisingly however, the actual number of unique protein variants, or proteoforms, exceeds this number by several orders of magnitude [2, 96]. The diversification of the set of proteoforms, or proteome, is achieved by what is collectively known as post-translational modifications (PTMs)¹. As the name suggests, these modifications are applied after the translational synthesis of protein in the ribosome, enabling the organism to adapt and react to e.g. environmental queues on a much shorter time scale compared to synthesizing completely new protein [38, 58, 90]. For the most part, PTMs are small chemical groups bound to the protein, which can serve as reversible activators, recognition sites, markers of degradation, precursors for protein folding and even stabilizing agents [58, 69, 70]. Including all PTM variants, the hypothetically possible size of the proteoform is in the billions [2, 96]. It is still an ongoing debate as to whether organisms are able to strictly regulate the PTM processes enzymatically to reduce the actual number of proteoforms, or that we are simply blissfully unaware of most proteoforms due to technical limitations of our detection methods [2].

Therefore, much work is still to be done in developing detection strategies. An extensive meta-analysis of the current state of research in the PTM field shows that contemporary studies are focussed on acetylation, glycosylation, and phosphorylation [57, 90], visualized in Figure 2.1 a-c respectively. These PTMs have been found to be widespread in both bacterial and eukaryotic cells, and play vital roles in the biological function of the proteins [38, 58, 90]. As a result, we aim to detect these PTMs in this study using nanopore sensing. Below, we will individually discuss the biological synthesis pathways of these PTMs and their biological relevance. Although these three PTMs are highly important for cell viability, we stress at this point that many more PTMs play vital roles in cellular functions. The reader is referred to Knorre et al. [58] for a more complete review.

2.1.1. Acetylation

Acetylation refers to the addition of an acetyl group to the N-terminal α -amine or the ϵ -amine on a lysine side-chain of a protein or peptide. It is a major PTM, present at least once in over 80% of all human protein [17]. From a physical perspective, acetylation removes the positive charge of the amine group, and replaces it with a more hydrophobic moiety. This can significantly alter the protein structure by redistribution of charges and hydrophilic/hydrophobic interactions [8]. Many protein domains have been identified that specifically interact with acetylated lysines, such as the bromodomain, enabling the cell to recognize and react to acetylation events [32].

In this study, we aim to detect lysine acetylation. This process is mediated by lysineacetyltransferases *in vivo*, which are capable of transferring acetyl groups from acetyl-coenzyme A (acetyl-CoA) to ϵ -amino groups on lysines, see Figure 2.1. Their counterpart, lysine deacetylases, can in turn remove the acetyl group. Dynamic activity of both acetyltransferases and deacetylases regulate among other things metabolic processes and initiation of gene transcription [32]. It is worth mentioning that non-enzymatic acetylation also occurs at high pH, where deprotonated amine groups can directly attack the terminal carbonyl group of acetyl-CoA.

Medically relevant is the acetylation of Tau protein, present in human neuron cells. Acetylation of the Lys174 slows cellular degradation of Tau, implicated in cognitive malfunction in mice models of Alzheimer's disease [8]. Acetylation of Lys280 has been implicated in the aggregation of Tau protein, and is often found in tissues of patients with Alzheimer's [8, 25]. Furthermore, acetylation has been shown to disrupt recognition motifs for other PTMs, such as phosphorylation, which has also been identified as a key problem in the development of Alzheimer's [32, 59].

¹Note that there are more ways to enlarge the number of proteoforms, such as alternative splicing of RNA and recombination of genetic information in somatic cells [5]. However, with the scope of this study in mind and to maintain clarity, these have been omitted.

2.1.2. Phosphorylation

Phosphorylation refers to the transfer of a phosphoryl group from adenosinetriphosphate to the hydroxyl group of a Ser, Thr, His or Tyr, without about 90% of the cases involving Ser and Thr [130]. The enzymes that mediate this process *in-vivo* are referred to as kinases. Functionally, it is often likened to flipping an "on-switch" for a given protein, as the phosphoryl group is able to activate protein activity by changing their conformation[58]. With over 500 human kinases identified more than a decade ago, astonishing levels of specificity can be obtained through each kinase variant's specific functionality and recognition site [130]. It is therefore not surprising that the diverse set of chemical mechanisms behind phosphoryl transfer and target site recognition is still an active area of research [58]. Complementary to the kinases are phosphatase families which are able to remove the phosphoryl and restore the original hydroxyl groups.

One of the pivotal roles of protein phosphorylation is in a signalling pathways of the cell [24]. One of the most well studied examples is the Mitogen-activated protein (MAP) kinase signalling cascade, which have been found to play a role in virtually all cellular processes [24]. By sequential phosphorylation of enzymes, signals from e.g. sensory protein are connected to the enzymatic functionality further downstream of the signalling cascade [40]. This enables the cell to respond to e.g. hormonal and environmental signals appropriately. These often involve very complex positive and negative autocrine feedback loops, and are heavily interconnected with a host of other cellular processes. It is therefore vital to have accurate kinase functionality [24, 40].

Mutations resulting in hyperphosphorylation within this cascade has been implicated in over 60% of human melanoma cases [27]. Specifically, the mutation of Val599Asp in B-RAF kinase was found to be particularly prevalent.

2.1.3. Glycosylation

Glycosylation refers the attachment of sugars to peptides/protein, creating glycopeptide/protein. These have been divided into N-linked and O-linked glycosylations, referring to the attachment site of the sugar to the carboxamide group of an Asp (N) or the hydroxyl group of a Ser/Thr (O) respectively. The attached sugar group, referred to as a glycan, is a complex entity, both in terms of composition and structure. It can range from a single sugar monomer to an intricate, branched structure of 10-20 subunits, built from a diverse set of over 10 sugar monomers [121]. There is no template determining the glycan structures, as there is for protein. Instead it is often the result of a dynamic interplay between transferase and glycosidase enzymes adding and removing sugar groups incrementally. This leads to a large degree of heterogeneity, even for glycopeptides with identical amino-acid sequences [121]. However, the cell can still indirectly influence the glycan structures by balancing the availability of functional enzymes. From a biological standpoint, glycan heterogeneity has been hypothesized to play a role in cellular immunity to viral attacks. Specificity in glycan structure recognition allows viruses to be highly selective of target cells. Cell-to-cell heterogeneity might then passively increase immunity to these attacks. We now focus on N-linked and O-linked glycopeptides related to this study.

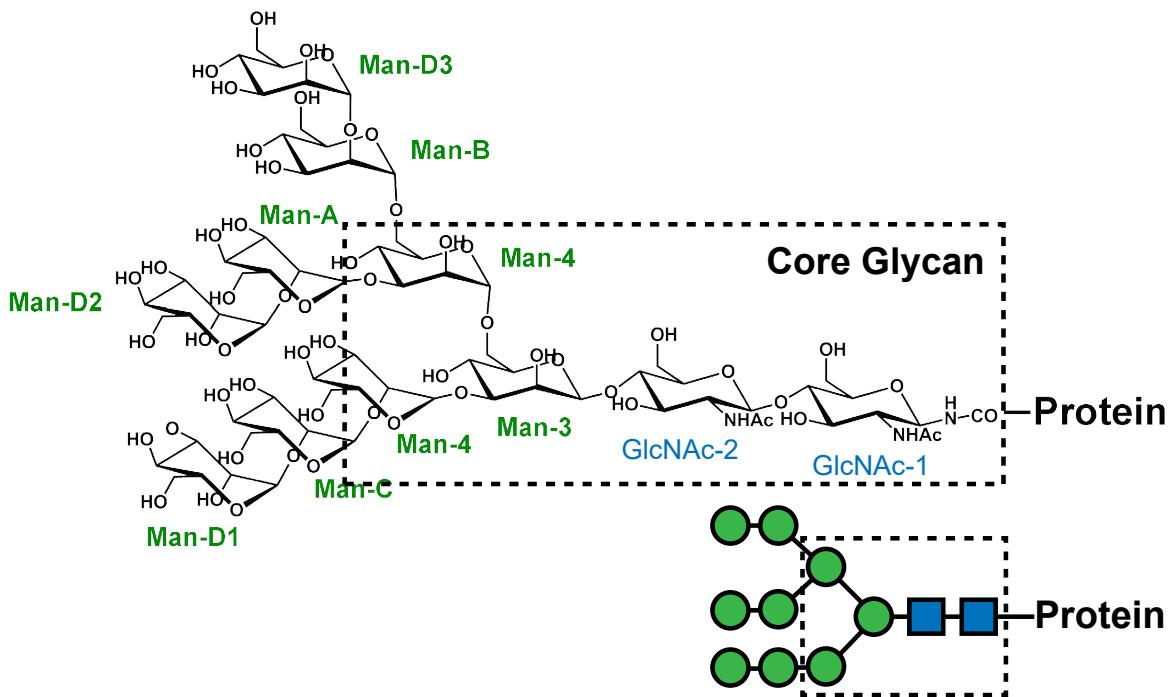


Figure 2.2: Structure of an 9-mannose N-Glycan. The core glycan structure is indicated by the dashed box. A simplified representation is often used consisting of geometrical shapes and colours , as is visible in the bottom. *Glycan structure was taken from Satoh et al. [109].*

N-linked glycosylation is performed co-translationally by a single oligosaccharyltransferase as the protein is synthesized in the ribosome [60]. Whilst the being processed, the protein inevitably passes the enzyme, which recognizes the sequence Asn-X-Ser/Thr, with X being any residue. A large 12-subunit glycan is then transferred from adjacent dolichol-lipids to the protein, giving the preliminary glycoprotein. Tail regions of the glycan are then quickly deconstructed further down the endoplasmatic reticulum by glycosidases, leaving the *core glycan*, consisting of two N-acetylglucosamine (GlcNAc) and three mannose (MAN) subunits [60]. An example 9mannose glycan is visualized in Figure 6.5, where the core glycan is boxed. A variety of enzymes then specifically recognize glycan structures and a step-wise addition of sugar subunits ensues, until this process is terminated by the addition of a negatively charged N-acetylneuraminic acid-galactose (NeuAc-Gal) subunit.

The most familiar form of N-linked glycans are those present on human red blood cells, where the terminal sugars determine the individuals blood type [121]. This showcases one of the many functionalities of glycoprotein, namely as specific receptor sites. A less familiar example that is more relevant to this study, is Ribonuclease A and its N-linked glycosylated variant Ribonuclease B. A single glycan is attached in the B variant at Asn34, structurally similar to the core glycan with 2-5 additional mannose subunits [134]. High resolution X-Ray diffraction techniques reveal that the glycan has nearly no influence on the crystal structure of the protein itself [134], even though glycans do have a strong effect on the structure in general [48]. Despite this, RNase B has a decreased degree of solvent exposure of residues near Asn34 [52] as well as increased thermodynamic stability of the protein as a whole [112]. Both of these phenomena have been attributed to steric hindrance provided by the glycan [52, 112]. Interestingly, the degree of solvent exposure is only weakly correlated with the size of the glycan [52]. This can be understood by the fact that the glycan points radially outward into the solution, with only the closest sugar groups interacting strongly with the protein [134].

O-linked glycosylations have less uniformity in glycan structure, as there is no universal core-glycan. Nevertheless, researchers have been able to define subclasses within the O-linked glycans, with the most prominent one being mucins [124]. Glycoprotein belonging to the mucin family do have core structures analogous to their N-linked counterparts. In contrast however, they are comprised of N-acetyl-galactoseamine (GalNAc) and galactose (Gal) subunits [121].

For our study, we used a mucin type glycopeptide with a single GalNAc subunit as a model substrate, namely MUC5AC. Mucin-type glycoprotein excel at holding water, as they are often heavily and densely glycosylated [124]. Hydrogen bonds formed between the sugar and water molecules ensure that evaporation is energetically unfavourable, giving surfaces exposed to air efficient moisture retention capabilities [124]. In

some cases, mucins can even form gel-like layers, such as those found in the nasal pathway. These can trap pathogens long enough for them to be removed, preventing them from infection. MUC5AC for example serves exactly these purposes [124]. It has also been speculated that mucin-like regions in other glycoprotein are used as structural supports. Steric hindrance due to the glycan can effectively make protein stiffer [121]. Examples include the hinge region between the Fab and Fac domain of Immunoglobulin A (IgA) which is heavily glycosylated and might otherwise be too flexible. Other structural functions of O-glycans include the linearisation of protein; heavily glycosylated mucins often cannot fold, simply due to the sheer abundance and volume of glycans encasing the protein [127]. This can force membrane glycoprotein to extend further outward, which again may play a role in cell-to-cell communication or protection against pathogens.

Finally we note that not all glycosylations involve large, complex glycans. Increasingly more protein in the cytoplasm and nucleus have been found to be dynamically glycosylated with a single O-GlcNAc subunit [121]. What is more, it was found that the target sites of O-GlcNAc transferase have a strong overlap with some protein kinase target sites [127]. Notably, transferases and kinases are in a dynamic competition at the C-terminus of RNA polymerase II, where phosphorylation signals the conversion of transcription initiation factors to mRNA elongation complexes [18]. It is hypothesized that O-GlcNAc glycosylations then allow for a second layer of control, as now both a kinase as well as a glycosidase is required to allow phosphorylation on a previously glycosylated target site. Knock-out studies of the O-GlcNAc transferase have shown that cells cannot survive a single cell-cycle without it, suggesting that it is vital for proper cell function [121]. O-GlcNAc-transferase is also implicated in the development of insulin resistance in patients with type 2 diabetes [18]. In view of this, we aimed to distinguish O-GlcNAc glycosylated and phosphorylated peptides in this work using a synthetic peptide platform.

2.2. Nanopore sensing technique

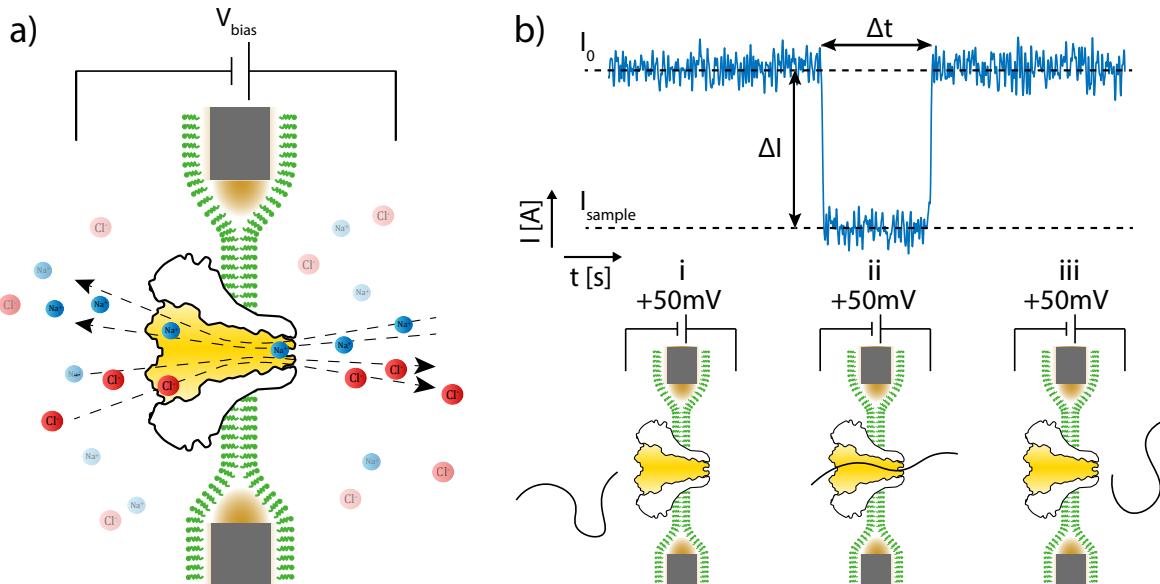


Figure 2.3: Overview of a general nanopore measurement. **a)** A diagram of a biological nanopore setup, with a cut-through view of the Fragaceatoxin C nanopore (white and yellow) in a lipid bilayer (green) spanning a teflon aperture (gray blocks). Negative Cl^- ions (red) and positive Na^+ ions (blue) move in opposite directions under the applied potential V_{bias} . **b)** Example current traces for an analyte translocation event for a positive V_{bias} . *(i)* Initially, the pore will have a baseline current I_0 with some noise. *(ii)* The entry and translocation of an analyte (black line) decreases this current by ΔI , up to the point that *(iii)* the analyte has fully translocated after a dwell time Δt , restoring I_0 .

Starting from the mid-1990's, nano-scale pore geometries were used to detect and characterize biopolymers, with ever-increasing degrees of sensitivity [132]. In a typical experimental setup, a nanopore is positioned in between two ionic solutions with concentrations of $0.1\text{M} - 1.0\text{M}$ KCl, and provides a serves as a pathway for ionic conductance. The nanopore sensing technique relies on measuring nA ionic current through a pore upon applying an electric potential V_{bias} . Figure 2.3a schematically shows ion movement in a biological pore inserted into a lipid bilayer. The basic components of sample analysis can then be understood by examining Figure 2.3b. Initially, *(i)* an open pore current I_0 is present, *(ii)* up until the point that an analyte

is captured by the pore through a combination of diffusive, electrophoretic and fluidic forces. This decreases the current by ΔI , which is highly dependent on the samples structure and composition. After dwell time Δt , the analyte leaves the pore, restoring the baseline current I_0 .

By examining the resulting current traces in detail, information on structure, charge and size can be obtained on a single-molecule scale [11, 119, 132]. Most often one looks at ΔI and Δt values to characterize and distinguish samples. With the release of commercial nanopore arrays, containing over 144000 channels capable of parallel measurements, the nanopore sensing technique shows promise of becoming a high-throughput single-molecule technique, with relatively simple data-interpretation steps [28].

2.3. Solid state and biological nanopores

There are many ways to obtain a useful nano-scale pore, each having its corresponding drawback. Generally speaking, nanopores can be roughly divided into two main types based on their constituent materials; biological and solid-state. Biological pores are multi-meric protein complexes, that entropically form into a nano-scale pore shaped geometry. Most often biological pores are inserted into lipid membranes to mimic their natural working environment [103]. Solid state pores are similar in function overall, but have a constriction built entirely from dielectric materials e.g. silicon nitride, glass or graphene [129].

Advantages of biological pores are their reproducible, well-defined geometries, with constrictions in the nanometer range [86]. Site-directed mutagenesis can be used to further alter specific characteristics of the pore to suit experimental needs, such as hydrophobicity, surface charge, and constriction diameter [11, 47, 129]. However, these perks are paired with some trade-offs. Lipid bi-layers are considerably less mechanically stable, and cannot be stored and reused practically over multiple measurements [129]. Biological pores are also comparatively more sensitive to measurement conditions, such as concentration of denaturants, salt or the solution's pH [129]. Other disadvantages include the protein synthesis and purification steps, which can be quite laborious and time-consuming, aside from requiring high levels of expertise. Lastly, protein pore stock solutions have a limited life-span, as they thermally denature on the timescale of months depending on the storage conditions, rendering them useless for future nanopore experiments [73].

In contrast, solid state pores are mechanically robust, and have no problems operating in denaturing conditions [129]. The most notable drawback here is significantly higher levels of dielectric noise, due to higher capacitance [116]. In addition, manufacturing consistent pore sizes and geometries at $\approx 1\text{nm}$ scale is still a challenging technical task. What is more, high surface charge density of most dielectrics lead to interactions with oppositely charged analytes, leading to semi-permanent pore clogging, temporarily hindering further experiments.

2.3.1. PTM detection with Nanopores

Nanopore-based PTM detection is still a relatively young field, with only a handful of reports published from 2015 onward [102]. Among those is the work by Fahie and Chen [34], showing that N-linked glycosylated avidin can be distinguished from unglycosylated streptavidin by transiently holding the protein in an OmpG pore entrance. Interestingly, no translocation of the glycoprotein was required to make the distinction. The change in surface charge due to the glycan was presented as the key component to enable detection.

Rosen et al. [104] showed that phosphorylated thioredoxin can be distinguished from its unmodified variant, but also that mono- and di-phosphorylated species are dissimilar enough for detection. In their set-up, they made use of the partially folded state of thioredoxin as a means of transiently stalling the translocation. A folded region close to the C-terminus of thioredoxin provided a high free energy barrier for translocation, effectively stalling the protein within the nanopore. A negatively charged polynucleotide label was then attached to the C-terminus end, providing an increased capture rate and enough force to unfold thioredoxin. The system allowed for extremely long read lengths of the C-terminal end, where the phosphorylation sites were strategically placed. It was observed that the setup was significantly less sensitive to phosphorylated sites away from the C-terminus, suggesting that confining the PTM in close to the pore constriction for extended periods of time is vital for detection.

Very recently, a report was released detailing the characterization of Hyaluronic acids (HA) molecules [101]. The authors show that HA molecules purified directly from equine synovial fluid could be characterized in a label-free manner, based on their degree of glycosylation. It should be noted that the HA structures are very large (54kDa - 2000kDa) [101]. For this work, we are however interested in smaller structural differences.

2.3.2. Fragaceatoxin C nanopore

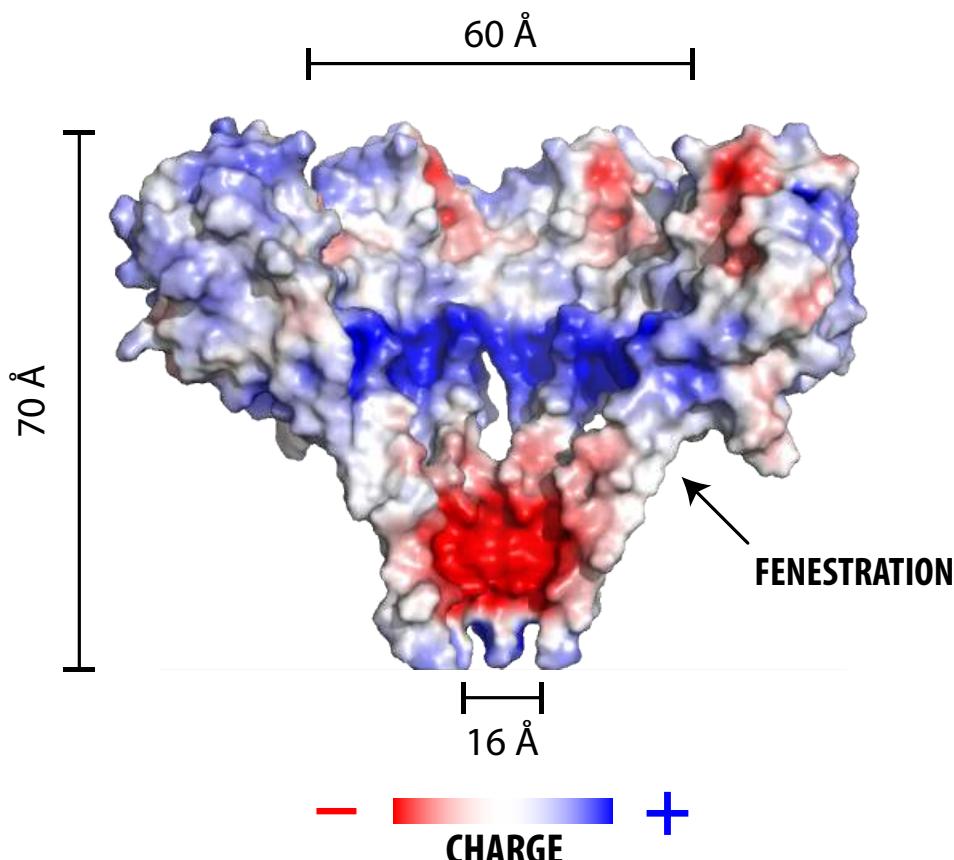


Figure 2.4: Surface plot of the Fragaceatoxin C nanopore with ADB electrostatics simulation [120]. Red and blue surface colour corresponds to negative and positive surface charges. The pore has a roughly conical shape, tapering down from 60 Å-13Å in diameter. The actual pore exit has a diameter of 16 Å. The protein structure contains fenestrations, indicated by the an arrow, allowing lipids to protrude into the pore lumen. Together with aromatic residues, this leads to a hydrophobic region on the height of the fenestration. *Image heavily based on work by Tanaka et al. [120]*

A consistent, very small pore constriction was deemed necessary for the detection of small PTMs. To that end, the nanopore sensing technique in this work is realized by the use of Fragaceatoxin C (FraC), an α -helical pore-forming toxin that naturally occurs in the sea-anemone *Actinia Fragacea* [12, 73, 120]. Biologically, the venom of *Actinia Fragacea* contains water-soluble FraC monomers, which can entropically insert into the lipid bilayer of a target cell upon unfolding its trans-membrane alpha barrel. After insertion, the monomers conglomerate and form octameric complexes in the shape of a nanopore. Massive perforation of a target cell will result in lethal lytic activity [12]. *In-vitro*, FraC has been used to distinguish proteins over a broad range of masses as a nanopore sensing platform, ranging from 1 – 25kDa [47].

A diagram of the pore is shown in Figure 2.4. FraC was chosen for its distinguishing features, including a relatively simple conical geometry, small constriction diameter of 1.3nm and localized charge distributions pictured in red (negative) and blue (positive) in 2.4. This makes it an excellent candidate for nanopore translocation experiments that require high sensitivity [73]. Other structural features of note are fenestrations, indicated by an arrow in Figure 2.4. These allow for lipid molecules to protrude into the lumen of the pore, creating local hydrophobic pockets. What is more, changing the lipid type decreases heptamer complex formation, suggesting that the lipids have a structural role as well [120].

Similar to other nanopores, FraC has shown cationic selectivity, due to the negatively charged residues in its constriction (red in the figure). Mutations of FraC with positively charged constrictions have conversely found to be anion-selective [47], giving further evidence that selectivity arises from surface charges. The solution's pH can be fine-tuned to regulate the protonation states of the acidic and basic residues in the pore. This can in turn alter surface charge and thereby ion selectivity. It has been hypothesized that the charges in the constriction could be structurally important, as mutations with aromatic residues lead to gating effects of

the pore (*F Lucas, unpublished work*).

2.3.3. Noise

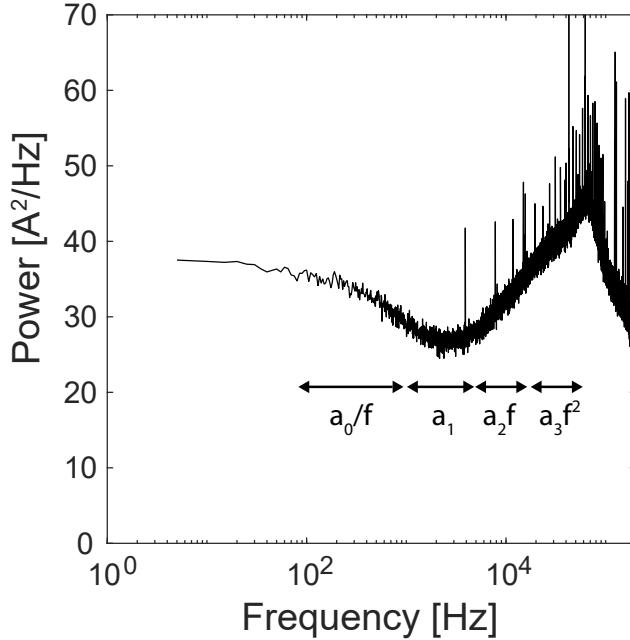


Figure 2.5: Power spectral density plot of FraC obtained by averaging 100 discrete Fourier transform spectra of a segment size of 0.2s. The different noise regions are indicated symbolically.

A limiting factor in nanopore research is high-frequency noise, necessitating the use of active low-pass filtering to obtain any recognizable current modulations from the measurement signal. This naturally limits the measurement's time resolution to the filter's finite rise time, which was found to be $\Delta t_{min} \approx 2 \cdot 0.34/f_{filt}$ for the Gaussian filter used in this study [91]. The factor 2 is included to account for the rise and fall of the signal. Events with Δt close to or below this limit will have significantly distorted shapes due to the filter, affecting both the pulse width and height. Consequently it is important to understand the noise sources, their contributions to the overall noise level, and ways to combat them, in hopes of maximizing frequency bandwidth. Figure 2.5 shows a representative plot of the current power spectral density ($S_I(f)$) for FraC in 1M NaCl solution at $V_{bias} = -90\text{mV}$. The general shape of this graph is commonly explained by four types of noise sources, categorized on their frequency dependence [118]. This is captured by a four-part polynomial model,

$$S_I(f) = \frac{a_0}{f} + a_1 + a_2 f + a_3 f^2 \quad (2.1)$$

in units of A²/Hz, with f being the frequency. The first two terms refer to flicker and white noise respectively, and cannot be easily avoided [118]. The next term is linear in f , and is attributed to dielectric noise. Non-ideal capacitors fundamentally lose energy by the production of heat in the dielectric material and the dissipative leakage of electromagnetic waves, causing dielectric loss [126]. These two contributions can be captured with

$$S_d = 8\pi k_B T D_{loss} C_m f \quad (2.2)$$

, where D_{loss} is the dielectric loss, and C_m is the membrane capacitance. For $f \gg 10\text{kHz}$ noise originates from an interplay of voltage noise at the trans-impedance amplifier input and the capacitances in the system [118]. Voltage fluctuations cause currents in capacitors as $I = CdV/dT$. High frequency voltage noise is essentially a quickly changing voltage, and therefore causes large current noise in capacitors. This can be effectively modelled by [118]

$$S_c = (2\pi f C_{system} e)^2 \quad (2.3)$$

, where C_{system} is the total system capacitance, most often dominated by the membrane, and e is the elementary charge. From the above we can see that capacitance is the leading cause for high frequency noise, making it essential to minimize it. This can be done in biological nanopore set-ups by e.g. minimizing lipid bilayer surface area and using lipids with minimal capacitative properties. It is also useful to minimize any parasitic capacitances from e.g. input capacitors or cable connections.

Other features captured in Figure 2.5 are the built-in 10kHz low-pass filter, causing the decrease in spectral power for $f > 50\text{kHz}$. Furthermore, some noise spikes are visible, which are thought to be originating from high capacitances within the measurement devices. Seeing as most of them are filtered out during data processing with a 10kHz Low-pass filter, these were not explored further.

2.4. Theoretical model of transport in Nanopores

In the past decades, some substantial progress has been made to understand transport phenomena in nanopores from the perspective of theoretical physics. Accurate analytical conductance models have been proposed and verified for simple cylindrical solid-state pores with large length to diameter (L/d) ratios [63]. Steps have also been made to describe deviations from the ideal cylindrical pore, such as conically and hour-glass shaped solid-state pores [19, 62]. Effects of surface charge and fluid-flow phenomena have also been elucidated numerically, showing that these can be the deciding factor in analyte transport directionality [6, 36]. However, significant difficulties arise when attempting to mathematically model biological pores, where complex protein pore geometries, pore flexibility and irregular surface charge distribution hinder analytical understanding and continuum modelling. In this section, we will start by giving an overview of the current theoretical work on transport in nanopores, focussing on models relevant to understanding the FraC nanopore. We then move on to discussing analyte capture.

2.4.1. Charge transport through nanopores

As ionic current is the main measurable quantity in any nanopore experiment, it seems fitting to start this section with charge transport. A host of literature can be found that aims to model charge transport using continuum models, disregarding any effects due to the discreteness of ions [6]. The most common continuum description of ionic fluxes in solution under an applied potential are the Poisson Nernst-Planck (PNP) equations [3, 19, 61, 133]. For these models, it is often assumed that dynamical effects occur at timescales $\ll 2 \cdot 0.34/f_{filt}$. We are therefore interested in steady-state solutions which can be written as

$$\nabla \cdot \mathbf{J}_i(\mathbf{r}) = 0 \quad (2.4)$$

$$\mathbf{J}_i(\mathbf{r}) = -D_i \nabla c_i(\mathbf{r}) - \frac{D_i z_i e}{k_B T} c_i \nabla \phi(\mathbf{r}) + \mathbf{u}(\mathbf{r}) c_i(\mathbf{r}) \quad (2.5)$$

$$\nabla^2 \phi = \frac{F^2}{\epsilon R T} (z_+ c_+(\mathbf{r}) + z_- c_-(\mathbf{r})) \quad (2.6)$$

, with D_i being the diffusion constant of ion species i , $\mathbf{v}(\mathbf{r})$ fluid flow velocity, z_i ionic valence of i , k_B boltzmann constant, T temperature, F Faraday constant, ϵ permittivity of vaccuum and R universal gas constant. The variables of interest are the ionic concentration $c_i(\mathbf{r})$ of species i , electric potential $\phi(\mathbf{r})$, and ionic flux $\mathbf{J}_i(\mathbf{r})$ at point \mathbf{r} . From this point on, explicit coordinate dependence is omitted for brevity.². Eq. 2.4 is simply a continuity statement for the ions. Eq. 2.5 describes the relation of diffusion, electric field $\mathbf{E} = -\nabla \phi$ mediated drift and fluid flow contributions to the ionic flux \mathbf{J}_i , as seen in the first, second and third term on the right hand side respectively. Finally, eq. 2.6 describes the electric potential due to the ions with the Poisson equation.

The next step is to find a description for the fluid flow field \mathbf{u} . Principally, \mathbf{u} arises through momentum transfer from ions to the solvent as a result of viscous drag. In an electro-neutral solution, $\mathbf{u} \approx 0$ should hold everywhere as the contributions of oppositely moving positive and negative ions mostly cancel each other out, given that they haev similar electrophoretic mobilities [6]. However, in most nanopore applications, charge imbalances are created by for instance static negative wall-charges on the pore surface. These attract positive counter-ions, creating the so-called electric double layer (EDL) which we will treat in the Section 2.4.3. Electro-neutrality is now locally broken, leading to $\mathbf{u} \neq 0$, which couples to the ionic current through eq. 2.5. This contribution is referred to in the literature as electro-osmotic flow, and has been shown to play a major role in both ionic conductance and analyte transport through nanopores [3, 36, 47, 66].

²For a derivation starting from the equations of motion, the reader is referred to Albrecht and Edel [7]

To describe \mathbf{u} , Navier-Stokes' equation is often used, with the assumption of incompressible flow and negligibility of inertial terms due to the small pore dimensions [3, 66, 133]. Here too, we are interested in the steady-state equations, which can be written as

$$\nabla \cdot (\rho \mathbf{u}) = 0 \quad (2.7)$$

$$\mu \nabla^2 \mathbf{u} - \nabla p - (z_+ c_+ - z_- c_-) \nabla \phi = 0 \quad (2.8)$$

, with fluid density ρ , pressure p , and viscosity μ . Eq. 2.7 is again a continuity statement for fluid flow. Eq. 2.8 describes the interplay of a pressure gradient, diffusive flow and an externally applied electrostatic force mediated by ion movement, in the first, second and third terms respectively.

Essentially, all charge transport characteristics can now be obtained by solving these equations for c , ϕ , \mathbf{u} and \mathbf{J} . However, no literature could be found on accurate analytical boundary conditions for protein pores. As a result, analytical models remain elusive. Previous studies have applied finite-element methods [3, 66], numerical integration [19, 20] or molecular dynamics [4, 67, 74, 123], obtaining great agreement with experimental current behaviour of both solid-state nanopores and biological pores.

Other authors have derived analytical models by applying appropriate approximations. These not only deepen qualitative understanding of nanopore conductance behaviour in general, but have also been quantitatively accurate for solid-state pores [3, 21, 63, 133].

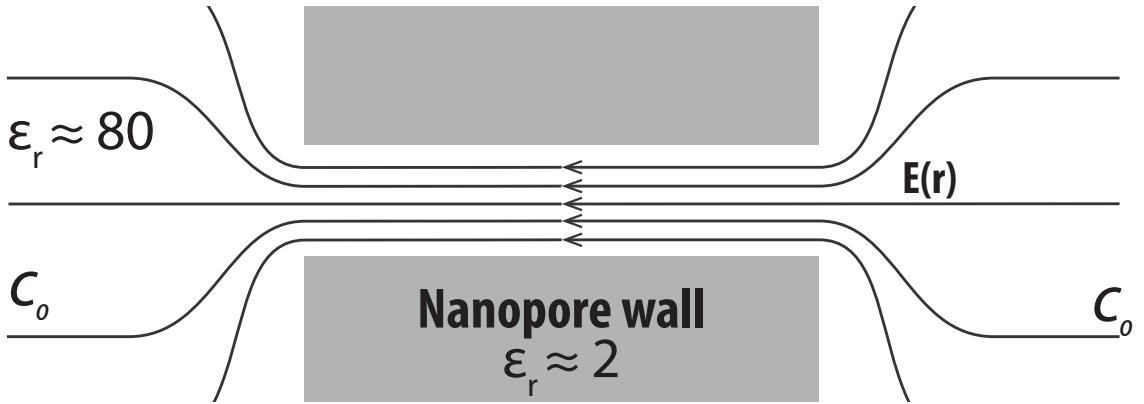


Figure 2.6: Diagram of cylindrical nanopore. Note that the field lines are initially directed orthogonally to the nanopore wall, and converge to the constriction. Within the channel, the are again parallel. *This image is heavily based on work by Albrecht et al. [6]*

The simplest model describes an ideal cylindrical nanopore, sketched in Figure 2.6. Here, the assumption is made that the electric field lines emanating from the electrodes all converge at the pore entrance, as the pore wall is built from electrical insulating material ($\epsilon_{r,water}/\epsilon_{r,wall} \gg 1$) [6]. We can then also conclude that the resulting electric field E inside of the pore must be constant, as the field lines are parallel throughout the length L of the pore, $E \approx V/L$. For large enough aspect ratio $L/d \gg 1$ we can then neglect any edge effects [7]. With these simplifications, we can treat the nanopore straightforwardly as a conductor in an electric circuit

$$I = \sigma \frac{A}{L} V_b = \sigma \frac{\pi d^2}{4L} V_b \quad (2.9)$$

$$G_{pore} = \frac{\pi d^2}{4L} \sigma \quad (2.10)$$

, with A the surface area of the pore entrance, d the pore diameter, G_{pore} the pore conductance and σ the conductivity of the ionic solution. Note is that the conductance scales linearly with nanopore diameter and ionic conductivity, which agrees with intuition.

For smaller L/d , edge effects can no longer be neglected. The convergence of ions toward small pore constriction then represents a significant contribution to the overall conductivity, and is referred to as the access resistance R_a . FraC for example has $L/d \approx 1$. [45] and Hall [41] described it by noting that the resistance between two electrodes in an ionic solution can be related to the their mutual capacitance as $R_a = \epsilon/\sigma C$. Taking the electrodes as a hemi-spheres and the nanopore entrance as a circular disk, they found

$$R_a = \frac{1}{\sigma d} \quad (2.11)$$

$$G_{total} = \sigma \cdot \left[\frac{4L}{\pi d^2} + \frac{1}{d} \right]^{-1} \quad (2.12)$$

, where G_{total} is the total conductance of the system. Kowalczyk et al. [63] show convincingly that this model fits experimental data with great precision for solid state pores with a large range of diameters between 5 – 100nm. In this work, the L was taken as a free parameter, and had to be set unrealistically high to achieve a good fit. To solve this problem, the authors proposed an hourglass shaped model instead [63]. This was validated by scanning-electron microscopy images, which show that nanopores made by electron-beam lithography and wet etching of SiN indeed have an approximately hourglass-shaped passage [9]. The resulting analytical expression shows that the data falls within the expected range of pore diameter variability, without having to change its length. This signifies that pore geometry is extremely important for quantitative accuracy at smaller L/d .

2.4.2. Surface charge effects

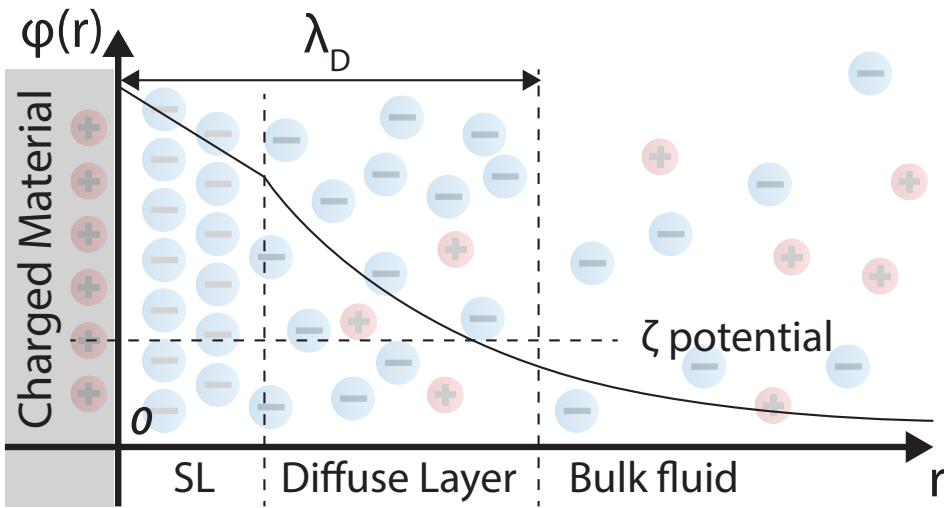


Figure 2.7: Qualitative plot of electric potential field versus distance r close to a positively charged surface in an ionic solution. SL indicates the Stern Layer of tightly bound counterions, adjacent to a diffusive layer of mostly counterions. Distance from the material to the so-called slipping plane between the diffuse layer and bulk fluid is the characteristic Debye Length λ_D . Eventually, the bulk solution ionic concentration is obtained for large r .

To further the discussion, we must first turn to surface charge effects. As mentioned previously, these are the leading cause for electro-osmotic flow, but also for other interesting phenomena such as local build-up and depletion of ions, electrophoretic drag forces and biased current/flow rectification [3, 6]. A qualitative diagram of the electric potential ϕ and ionic concentration over distance r near a charged object is shown in Figure 2.4.2. Here, a positively charged (red) material is placed in an ionic solution, whereby a tightly bound layer of counter-ions (blue) forms the so-called Stern-Layer (SL) [1]. The wall-charges are then shielded for larger r , resulting in a more loosely bound diffuse layer containing a majority of counter-ions. Using equations 2.4-2.6, it can be shown that the electric potential decreases exponentially, with a characteristic length of λ_D [1]. This value is directly related to the concentration of ions as follows,

$$\lambda_D = \sqrt{\frac{\epsilon_0 \epsilon_r k_B T}{2 N_A e^2 I_s}} \quad (2.13)$$

, where $I_s = \sum_i z_i c_i$ is the ionic strength, e is the elementary charge and N_A is Avogadro's number. For a 1M NaCl solution in water at room temperature, one obtains $\lambda_D \approx 0.3\text{nm}$.

2.4.3. Pore assymetry: biased current and flow rectification

The nanopores found in literature all possess surface charges, either due to the solid-state material interface or due to acidic/basic residues protein pores. For pore dimensions close to λ_D , the pore interior has an increased amount of counter-ions due to EDL formation. As an example, lets consider a negatively charged

cylindrical pore, with positively charged counter-ions. Here, the majority of ions in the pore will be positive, and will move with the electric field until it reaches the edge of nanopore. Here, the electric field strength decreases due to access resistance, causing a build-up of incoming cations and eventually reaches a steady-state equilibrium with diffusive flow [66]. At the same time, a depletion of ions occurs at the other end of the pore as the anions are repelled by the surface charges, whilst the cations are migrating away. The redistribution of charges described here is referred to as concentration polarization and causes rectification in the ionic current due to the need for ions to overcome a concentration gradient [3, 66, 133]. In other words, nanopores can deviate from their Ohmic, linear current-voltage (I-V) behaviour due to the effects of surface charges. It was observed that larger λ_D increase the rectification effects [3, 77], allowing one to tune the behaviour by varying the ionic strength, see eq. 2.13.

What is even more interesting is that many biological pores (and also some particular solid state pores) exhibit bias in their rectification behaviour. That is, the magnitude of current and flow rectification depend on the electric field's polarity. From experimental and theoretical studies, it was concluded that an asymmetry in either pore geometry or surface charge distribution is necessary in the direction of flow for this phenomena to occur [3, 19, 22, 56, 77, 114, 115, 133]. Biological pores α -Haemolysin, MspA and also FraC fall within this category [35, 117, 120], making it useful to understand these phenomena. For solid-state pores, conical shapes have been fabricated that have been reported to exhibit similar characteristics, adding further evidence to this claim [22, 56, 114, 115, 133]. The asymmetry leads to a similarly asymmetric concentration polarization effect, with significantly different steady-state ionic concentration profiles depending on the polarity of V_{bias} .

2.4.4. Application of theory on FraC

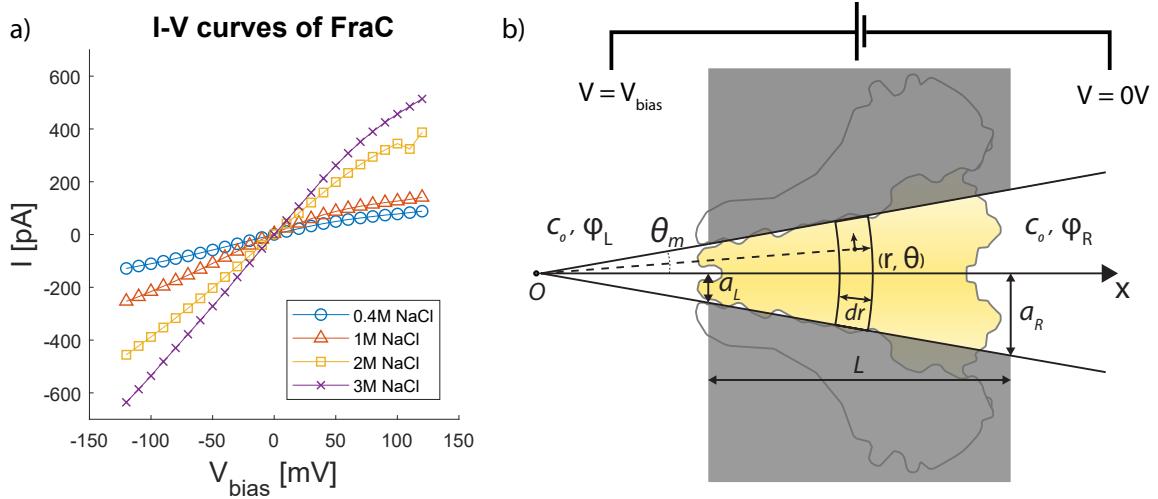


Figure 2.8: Overview of FraC nanopore. **a)** Experimental I-V curves taken of the FraC nanopore at 0.4M, 1M, 2M and 3M NaCl. Points represent averages over a 5s trace. The lines function only as a guide to the eye. **b)** Cut-through diagram of the FraC nanopore, with a spherical coordinate system overlaid on top, based on work from Cervera et al. [20]. Analogy to an ideal conical pore is made by the gray boxes. The origin O is put at the projected cone tip, with the angle θ measured counter-clockwise to the x -axis, and r the distance to O . The positive bias is defined at the narrow end of the pore.

We will now analyse the FraC nanopore in the theoretical context presented above. To that end, we have plotted experimental I-V curves of FraC for 0.4M, 1M and 2M NaCl in Figure 2.8a. Biased rectification can be observed, where the positive V_{bias} at the narrow end of the pore shows rectification. As expected, the degree of rectification decreases with salt concentration, and therefore with λ_D .

We compute the pore conductance G , by linear fitting of the curve between $-50mV < V_{bias} < -150mV$ as to omit the rectified values. The result for 1M NaCl is $G_{FraC} = 2.29 \pm 0.1nS$, with the error indicating a s.d. of 3 measurements. We compare this with the cylindrical model using eq. 2.12, taking into account the access resistance. Taking $L = 7nm$, $\sigma = 9.86S/m$, and the narrowest pore diameter $d = 1.3nm$, we obtain $G_{cyl} = 2.18nS$, which is surprisingly well in agreement with the experimentally obtained conductance. This suggests that the narrowest pore constriction of 1.3nm has the largest influence on the overall conductance of the pore.

However, this model can in no way explain the rectifying behaviour shown in Figure 2.8a. As a result, a more sophisticated model is required. We follow the procedure by Cervera et al. [20], who has obtained excel-

lent agreement of numerical and experimental data, showing quantitatively accurate rectification behaviour for conical solid state nanopores. We include key-steps of their derivation here, and refer the to the original articles Cervera et al. [19, 20] for a full derivation. The aim is to build a conical continuum model using eqs. 2.4-2.6. A major simplification is made by omitting the flow component $\mathbf{u}c_i$. Another key step in the derivation is the use of mean-sphere approximations, in which the average values of \mathbf{J}_i , c_i and ϕ is computed through a spherical shell of thickness $d\mathbf{r}$, instead of at every point \mathbf{r} . To do this, a spherical coordinate system is set-up with the origin corresponding to the cone-tip, see Figure 2.8b. Using this approximation, one can rewrite eq. 2.4 - 2.6 in spherical coordinates as

$$\langle J_i \rangle = -D_i \left[\frac{d\langle c_i \rangle}{dr} + \frac{D_i z_i e}{k_b T} \langle c_i \rangle \frac{d\langle \phi \rangle}{dr} \right] \quad (2.14)$$

$$\frac{d}{dr} \left[-r^2 D_i \left(\frac{d\langle c_i \rangle}{dr} + \frac{D_i z_i e}{k_b T} \langle c_i \rangle \frac{d\langle \phi \rangle}{dr} \right) \right] = 0 \quad (2.15)$$

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\langle \phi \rangle}{dr} \right) = \frac{F^2}{\epsilon R T} \left(\langle c_- \rangle - \langle c_+ \rangle - X(r) \right) \quad (2.16)$$

, where r is the radial distance from the origin, θ is the angle to the x -axis, θ_m is the opening angle of the cone, and averaging operator $\langle \rangle$ and $X(r)$ are defined as

$$\langle \rangle = \frac{1}{1 - \cos(\theta_m)} \int_0^{\theta_m} \sin(\theta) d\theta \quad (2.17)$$

$$X(r) = \frac{\sin(\theta_m) \sigma}{Fr(1 - \cos(\theta_m))} \quad (2.18)$$

. $X(r)$ is obtained by applying Gauss' law for fluxes, and can be interpreted as the charge contribution due to counter-ions. σ is the surface charge density of the pore in units of C/m^2 . To further the derivation, non-radial fluxes are assumed to be negligible for high L/d pores. This corresponds mathematically to $\mathbf{J}_i = J_i \hat{\mathbf{r}}$. Lastly, to obtain a simple set of equation, we need to approximate $\langle c_i \frac{d\phi}{dr} \rangle \approx \langle c_i \rangle \frac{d\langle \phi \rangle}{dr}$. With this, Eq. 2.14 - 2.16 is a system of coupled, one-dimensional, non-linear ordinary differential equations, and can be integrated numerically with relative ease.

Boundary conditions for ϕ and c_i are found by assuming a Donnan-equilibrium between the cis and trans side of the pore, together with local electro-neutrality at the pore interfaces. This can be formulated mathematically as

$$\langle (c_i(r_j)) \rangle = \frac{1}{2} \left[-z_i X(r_j) + \sqrt{X(r_j)^2 + 4c_0^2} \right], i = +, -; j = L, R \quad (2.19)$$

$$\langle \phi(r_j) \rangle = \phi_j - \sum_i \frac{1}{z_i} \ln \frac{\langle c_j(r_j) \rangle}{c_0}, i = +, -; j = L, R \quad (2.20)$$

. Finally, we can find the ionic current at given V_{bias} as

$$V_{bias} \equiv \frac{RT}{F} (\phi_L - \phi_R) \quad (2.21)$$

$$I = 2\pi(1 - \cos(\theta_m)) Fr^2 \sum_i z_i \langle J_i \rangle \quad (2.22)$$

. To validate our method, we have simulated the parameters given by Cervera et al. [20] and obtained practically identical results.

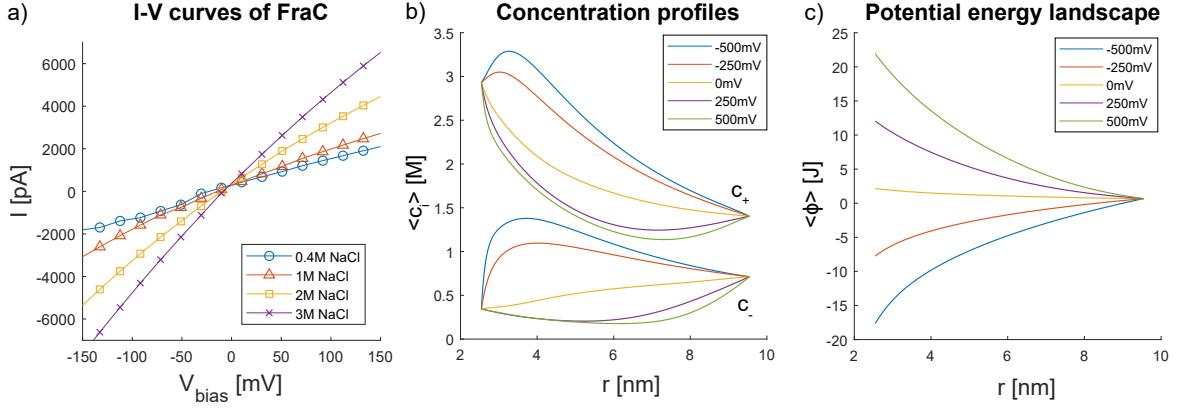


Figure 2.9: Numerical integration results for $a_L = 1.3\text{nm}$, $a_L = 6\text{nm}$, $L = 7\text{nm}$, $T = 293K$ and $\sigma_c = -0.976C/m^2$. **a)** I-V curves for 0.4M, 1M, 2M and 3M NaCl, indicated by the blue circles, orange triangles and yellow squares respectively. The lines are purely meant as a guide to the eye. Current rectification is observed for $V_{bias} > 0$. **b)** Steady-state ion concentration profiles along the pore axis for 1M NaCl. The colors represent voltages from -500mV to $+500\text{mV}$. The top lines refer to Na^+ and the bottom lines to Cl^- concentrations. The concentration does not reach bulk concentrations of 1M, due to boundary conditions. **c)** Electric potential ϕ along the pore axis for 1M NaCl. Note that the right side of the pore is set to $\phi \approx 0$ due to boundary conditions. The gradient in potential results in charge redistribution.

Reasonably confident in our methods, we now continue by performing the integration for FraC, with the parameter values $a_L = 1.3\text{nm}$, $a_L = 6\text{nm}$, $L = 7\text{nm}$, $T = 293K$ at concentrations of 0.4M, 1M, 2M and 3M NaCl. We approximate the surface charge σ by only considering the negative charges near the narrowest constriction, as these are thought to be most influential to the rectification behaviour [77]. We further assume that all 16 acidic residues are de-protonated, and divide by the truncated conical surface area. This gave $\sigma_c = -0.976C/m^2$.

The numerical (I-V) curve was plotted in Figure 2.9a, together with the concentration and electric potential profiles in b and d. Current rectification is observed at positive V_{bias} , which indicates a qualitative agreement with the experiment. We furthermore see the same general decreasing trend in current with salt concentration, as expected. However, the numerical current magnitudes are an order of magnitude too large, note the difference in x- and y-axis. What is more, the current rectification strength is decreased compared to the experimental results.

It is at this point interesting to note that both these numerical results and experimental results in Figure 2.8a show rectification at positive V_{bias} . In contrast, other works in literature show rectification at negative V_{bias} for conical pores [9, 19, 20, 66, 114, 115, 133]. Experimental and theoretical works have shown that rectification bias inversion occurs when the pore radius $r_{pore} \approx \lambda_D$ [3, 53, 77]. At this regime, the dominating local electric potential field is no longer due to pore geometry, but due to the high density of surface charges. This creates an extreme enrichment near the pore tip, which behaves in an exactly reversed manner compared to 'classical'-rectification, see Figure 2.9b-c. We can therefore conclude that FraC exhibits inverted rectification bias, which was qualitatively captured by the model.

We now conclude that the simple, uniformly charged, conical nanopore model with radial flows does not quantitatively describe current and rectification magnitude, but does captures qualitative properties of the system. From the theory above, we identify potential root causes for the experimental and theoretical mismatch that for biological nanopores: 1) Electro-osmotic flow must be incorporated to be quantitatively accurate. 2) Non-radial current and fluid flows are not negligible, as $\theta_m \gg 1^\circ$. 3) Overestimation of ionic fluxes due to the approximation for $\langle c_i \rangle \frac{d\langle \phi \rangle}{dr}$ [20].

As a result, we pose that increased accuracy in I-V behaviour must be obtained by using more sophisticated models. Finite-element models have proved to be very effective in modelling solid state pores, enabling a unified description of charge-polarization and biased rectification effects [3, 66, 77, 133]. Seeing as the main focus of this study was experimental work, we have opted to not pursue this any further due to time constraints.

2.4.5. Capture rate of particles

With the background of charge transport in mind, we are now ready to discuss protein/peptide capture in nanopores. Theoretical work by Grosberg and Rabin [39] and Muthukumar [80] offers convincing arguments for the existence of two main regimes, denoted as the diffusion- and barrier-limited regime. Within this definition the rate-limiting process is either the protein's diffusive approach to the pore vicinity, or the free energy barrier for translocation. The latter could be due to entropic effects of confinement in the pore, or enthalpic effects due to protein-pore interactions and the protein structure [74, 80]. It must be noted that the molecule of interest in these works were long DNA strands, but general principles still apply to protein capture relevant to this study.

To start, [39] derives a simple expression for the electric potential close to the nanopore mouth, using the ideal cylinder approximation and conservation of ionic flux. The result is

$$V(r) = \frac{d^2}{8Lr} V_{bias} = \frac{q_{eff}}{\epsilon r} \quad (2.23)$$

, where r is the distance to the pore, ΔV is the applied bias voltage. The point of focus here is that the nanopore is likened to a point charge q_{eff} , capturing oppositely charged analytes in a distance dependent manner. They then look for the equilibrium point between energy gain from ΔV and random thermal motion $k_B T$. This distance defines a capture-radius r^* , as for $r < r^*$ the analyte will hypothetically no longer be able to escape its arrival at the pore due to thermal fluctuations. At point r^* , an effective thermal stall force can be defined, working against electrophoretic drift. By equating the work done by this stall force to the thermal energy $k_B T$, one obtains [80]

$$r^* = \frac{q_{eff}\mu}{\epsilon D} \quad (2.24)$$

, where μ is the analytes electrophoretic mobility, and D is its diffusion constant. It is taken then that all particles within r^* will inevitably diffuse toward the nanopore. The flux of analytes to the pore can then be written as

$$J_{analyte} = cR_c = 2c\pi Dr^* = \frac{\pi d^2 \mu}{4L} V_{bias} \quad (2.25)$$

. It must be noted at this point that a cylindrical approximation for the nanopore was used for this derivation, and is therefore not quantitatively correct for FraC and other biological nanopores, where geometric and edge effects are significant. Nonetheless, qualitative linear scaling with concentration and V_{bias} still seems reasonable.

Next, we discuss barrier-limited diffusion. This process was intensively discussed by Grosberg and Rabin [39] for long DNA coils, which have an (approximately) uniform surface charge. Unlike DNA, protein can have positively, negatively or neutrally charged regions, invalidating many of the detailed approximations made by the authors, such as dielectric polarization, enhancement effects, and more. However, a qualitative statement still holds by considering that the protein must overcome some entropic energy barrier U . This barrier is effectively lowered by aqV_{bias} , where a accounts for decrease in effective applied voltage over the pore due to e.g. R_{access} . This allows us to write

$$J_{analyte}^{barr} = cR_C^{barr} \propto ce^{U/k_B T - aqV_{bias}/k_B T} \propto ce^{aqV_{bias}/k_B T} \quad (2.26)$$

, where the main point is that the capture rate should be again linear in concentration but now exponential in applied voltage. It can thus be concluded that the translocation type can be recognized based on the voltage dependency of the capture rate.

The theory is, however, further complicated by introducing the concept of protein accumulation. Freedman et al. [37] have shown that analytes in the barrier limited regime accumulate near the pore due to constant electrophoretic influx of ions, and slow barrier limited outflux. Practical examples include large, folded protein with a high entropic translocation barrier. Of experimental importance is the fact that this accumulation causes pore clogging as a result of multiple protein entering the pore at once. It was therefore suggested that in the barrier-limited regime, the polarity of V_{bias} should be regularly switched, together with intervals of $V_{bias} = 0$ to allow for the samples to equilibrate.

Experimental studies have also shown that electro-osmotic flow also plays a significant role in analyte capture [36, 47]. Theoretical work was done by Wong and Muthukumar [136], showing that analogously to

electrophoretic capture, a capture radius can be defined. The quantity of interest here is the Zimm relaxation time of a molecule compared to the time required to reach the pore using electro-osmotic flow alone. We can then write the general relation

$$r_{EOF}^* \propto R_g \quad (2.27)$$

, where R_g is the particle's radius of gyration. This suggests that EOF particle capture is size-dependent. EOF therefore provides a mechanism by which uncharged particles can be captured.

2.4.6. Dwell-time distributions

Aside from capture rates, the distribution in dwell-times Δt of the particle is also often used to identify and/or distinguish a set of analytes [43, 47, 73]. It is usually assumed that a translocation event can be seen as a two-state reaction, from confinement to translocation. Depending on the physical parameters such as size, polarity, and hydrophobicity, the pore-analyte interactions can vary. To quantify this, [64] rate theory is applied, and the dwell-time population is fit to a single exponential distribution $P(\Delta t) = a \cdot e^{-kt}$ with k being the 'reaction rate'. The mean dwell-time is then taken as the inverse of the reaction rate $\bar{\Delta t} = 1/k$.

This model is very effective in describing approximately linear polymers, such as DNA strands, short peptides and even bumping of large globular protein. More complex interactions such as unfolding of secondary and tertiary protein structures no longer fit the model [119]. Instead, some studies have opted for multiple exponential fits, indicating equally many pore-protein 'interaction-reactions' [75, 87]. Statistical mechanics models have shown that gating effects due to pore fluctuations usually fit to a power-law distribution $P(t) = at^b$ instead, with $b = 3/2$ [76, 85]. The applicability of this theory on FraC however is still not clear.

For this study we mostly expect single-exponential distributions, as we study short peptide sequence of roughly 30 amino acids. The exception to this is our study of Ribonuclease A and B, which are globular protein of 126 amino acids. For these we will apply somewhat more involved fitting when necessary in an attempt to further characterize our experimental results.

3

Material and Methods

In this chapter, we described the methods and materials used in this work. Firstly, the specifics regarding nanopore translocation are discussed. Secondly, data acquisition and analysis methods are treated in detail. We then describe the sample preparation protocols. Finally, we gave a detailed list of all chemicals used in this work.

3.1. Nanopore translocation

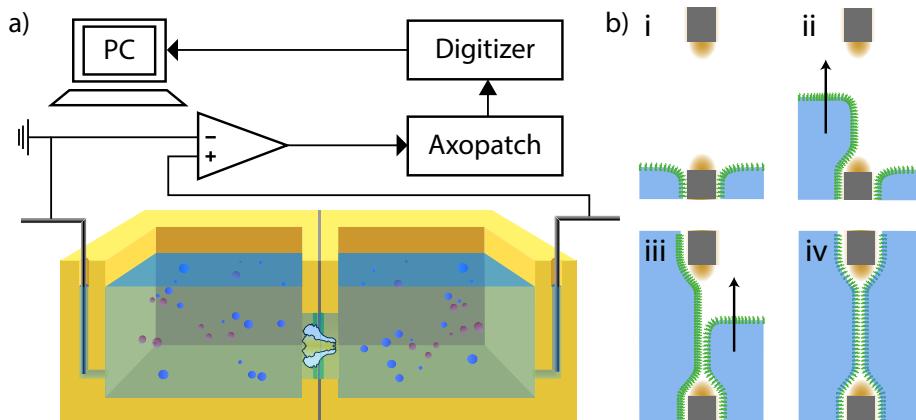


Figure 3.1: Diagram of measurement setup. **a)** A 2-part polyether ether ketone flowcell was created in house (yellow). These are tightly fixed together with screws with a teflon film in between (gray line). The lipid membrane and nanopore are built on top of the teflon film. The signal propagation from electrode to measurement PC is visualised schematically. **b)** Schematic overview of the Montal Mueller method for creating lipid bilayers. Water levels are sequentially changed to reorient the lipids on the water interfaces.

This study largely follows the measurement protocol described by Maglia et al. [73]. A cut-through diagram of the experimental set-up is shown in Figure 3.1a. Measurements were performed using a polyether ether ketone flowcell made in house, comprised out of 2 identical chambers, visualized in yellow in the diagram. Each chamber is filled with $400\mu L$ of ionic buffer solution. A $3\mu m$ diameter circular pathway is drilled on the side of each chamber to allow a secure contact between electrode and ionic solution. The two flowchambers are screwed together, trapping a teflon film in between them (gray in the figure). To connect the chambers, an aperture was made in the teflon film by a high-voltage shock. The aperture, forms the framework for installing the lipid bilayer, and eventually the nanopore. Aside from the aperture, the film must form a fluid-tight seal to prevent leakage currents.

The aperture diameter decides lipid-bilayer surface area, and with it the largest capacitative contribution in the system [73]. Deciding on a diameter is therefore a careful balancing act between signal-to-noise ratio (S/N) versus membrane stability. The lipid bilayer must envelop the bilayer thickness, see Figure 3.1b. Small apertures are therefore affected comparatively by edge curvatures, bringing the lipid molecules toward a less energetically favourable conformation. We follow Maglia et al. [73]'s protocol and aim for aperture diameters of $\approx 50\mu m$. This was found to give adequate S/N for the measurements in this work.

3.1.1. Lipid bilayer formation

The next step is to build a lipid bilayer, spanning the teflon aperture. We employ the so-called Montal-Mueller method [78] for this purpose, chosen for its practical simplicity and reproducibility [73]. A schematic overview of the procedure is given in Figure 3.1b. To prepare the teflon film, a droplet of 5% hexadecane oil dissolved in pentane is first applied on both sides of the teflon film using a glass pipette. The pentane quickly evaporates, leaving a layer of hexadecane (indicated in brown in Figure 3.1b.). The oil layer cushions the lipid bi-layer, effectively decreasing the lipid bilayer curvature required to span the teflon thickness gap, see Figure 3.1b. (i) Ionic buffer solution is added to the chamber, together with a droplet of 10mg/mL 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC) lipids. DPhPC has a remarkable thermal and chemical stability, allowing them to stay stable in a range of -120 deg C to 120 deg C without gel-matrix transitioning and forming lipid bilayers that stay stable for days [54]. It has been hypothesized that this is brought about by the dense packing of its carbon tails due to their structure [128]. (ii) The lipid molecules collect at the air-water interface and arrange in their energetically favoured orientation, with their hydrophobic carbon tails and charged phosphate group pointing toward the air and water respectively. (iii) The water level in one flow chamber is slowly raised by pipetting in ionic solution. This forms one half of the lipid bilayer, spanning the teflon aperture. (iv) By repeating the process in the opposite chamber, a second lipid layer is formed in a similar manner, (iv) giving the required lipid bilayer.

A successful formation can be verified by measuring an increased system capacitance due to the lipid bilayer. In practice, a voltage is applied across the bilayer in a triangular wave fashion with slope $dV/dt = 1$. From basic circuit theory we know that $I = CdV/dT$. In this case, the triangular wave voltage signal should result in a square wave current signal upon bilayer completion, where the current I gives a direct reading of the membrane capacitance C . An estimate of the surface area can then be made by approximating the bilayer as a parallel plate capacitor: $C = \epsilon A/d$, where $\epsilon_{lipid} \approx 2$ [128]. The bilayer is then tested for stability by alternating $+300mV, -300mV$ pulses. If the formation was initially unsuccessful, one should repeatedly perform the water raising and lowering steps, as the lipid molecules might need to be reoriented in the solution [73].

3.1.2. Nanopore insertion

After obtaining a stable lipid bilayer, FraC octamer complexes are added to the cis-side of the chamber. The other chamber is then by definition the trans-side. FraC heptamers diffuse and stochastically arrive at the lipid bilayer. It then thought to preferentially insert its transmembrane alpha-barrels into the lipid bilayer [120], giving it a consistent orientation. An insertion event can be recognized by a stepwise increase in for a constant V_{bias} . The pore's orientation can subsequently be verified by measuring its I-V behaviour. Positive V_{bias} has been empirically found to give strong current rectification, as discussed in Section 2.4.4. With this, the cis- and trans-side can then be confidently related to the wide and narrow constriction of FraC respectively, as shown in Figure 3.1a.

Naturally, there is nothing preventing more than one pore from inserting entropically. Although measuring with multiple pores in parallel inherently gives higher capture-rates, significant drawbacks are found in subsequent data-interpretation as a result of e.g. pore-to-pore conductance variability and overlap of signals due to simultaneous translocations. For this reason, measurements were only performed with solitary nanopores in this work. It was therefore ill-advised to add too many FraC octamers, as the probability of a second stochastic insertion event increases.

3.1.3. Silver/Silver-Chloride electrodes

Silver/Silverchloride (Ag/AgCl) reference electrodes were used to connect the ionic solution to the headstage amplifier. These electrodes convert ionic currents to electron currents through the electrochemical reaction of solid silver (Ag) and chloride ions (Cl^-) to solid silverchloride ($\text{AgCl}(s)$)



. Ag/AgCl electrode is affordable, easily produced and, most importantly non-polarizing. The latter means that no capacitative current arises due to build-up of counter-ions, enabling the electrodes to mediate current accurately even for rapidly changing V_{bias} . It was found experimentally that properly prepared electrodes can stay stable for hours in a nanopore setup with minimal drift. As a result, Ag/AgCl electrodes are the most common electrodes used in nanopore experiments in Cl^- containing solutions.

For this study Ag/AgCl were fashioned in house by submerging Ag wire in commercial bleach solution. Prior to bleaching, the electrodes were cleaned by abrasion with sandpaper to remove any impurities and

rinsed thoroughly with 100% ethanol and water. Hypochlorite in bleach solution creates an AgCl layer on the electrode surface. Submersion is maintained for at least one hour and preferably overnight to ensure a deep plating. Successful preparation is verifiable by a black-purple colour. Improperly plated electrodes are quickly exhausted of solid AgCl, causing significant drift in measured current and/or unexpected surface reactions. Care must also be taken to store the electrodes in a dark place, as AgCl is sensitive to light [73]. Electrodes can be reused by repeating the cleaning and plating steps. Replacement should be considered when electrode diameter decrease too much due to abrasion. The decreases in electrode surface area can increase its resistance, adding to systematic errors [73].

3.1.4. Translocation measurements

Translocation measurements are performed in ionic solution of 1*M* sodium chloride (NaCl), 10*mM* tris(hydroxymethyl)aminomethane (TRIS) and 1*mM* ethylene-di-aminetetra acetic acid (EDTA) at pH = 7.2, unless stated otherwise. *KCl* is the most common salt used in the nanopore field, as the mobilities of K^+ and Cl^- ions are nearly identical. *NaCl* is used to retain consistency with measurements done in this group prior to this study. It must therefore be noted that biological nanopores show somewhat stronger cation selectivity in *NaCl* solutions compared to *KCl* [6]. This must be kept in mind when comparing results from this work with literature.

3.1.5. Data acquisition and analysis

A schematic overview of the signal acquisition pathway is presented in Figure 3.1a. V_{bias} is applied over the electrodes by an Axon Axopatch 200B Amplifier. Output current signals are then pre-amplified in the headstage, before being fed back to the Axopatch further amplification and measurement. Both the flowcell and the Axon Instruments CV203B4 headstage were located in a aluminium box with roughly 0.5*cm* thick walls for electromagnetic shielding. To further eliminate external sources of noise, the entire box is placed on a vibration free table (Newport). The Axopatch allows for a maximum sampling rate of 500*kHz*, with a maximum low-pass filter bandwidth of 100*kHz* of its analog bessel-filter. In most nanopore studies found in literature, Axopatch amplification values are $\alpha = 1$ and $\beta = 10$. This, however, causes current overload for high $|V_{bias}|$ used in this study. For consistency, current amplification is therefore set to $\alpha = 0.1$ and $\beta = 1$. The signal is then propagated to a Axon Digidata 1550B digitizer and further processed in a PC.

The resulting digital current-time traces for this study were generally analysed using automated MatLAB software package Transalyzer developed by Plesa and Dekker [93]. Data processing can be split into two main parts: event detection and event analysis. The former refers to distinguishing nanopore blockade events from noise. The latter entails processing and interpreting the population characteristics of the events, commonly done by translating event collections to dwell-time and current-blockade distributions. Transalyzer presents solutions for both of these tasks in an automated and user-friendly manner. Similar to many other event analysis programs in the field of nanopore research, a threshold detection algorithm is employed to distinguish events from noise. The threshold value is determined by multiplying a moving average of the signal's standard deviation with a user-defined factor. For most applications, the signal is first low-pass filtered to remove high-frequency capacitative noise which enables thresholding to be more effective. For this study, signals are filtered with a Gaussian Low-pass filter with bandwidths of either 10*kHz* or 2*kHz*, whichever bandwidth the S/N of the signal in question permits. Dwell-time is then determined as the Full-Width at Half Maximum (FWHM), and the average blockade amplitude is computed as the area under the current signal divided by the FWHM time. This integrating method can be further used to detect multiple levels inside of a single events in e.g. DNA-protein interactions or protein-unfolding in the pore.

Distinguishing features of Transalyzer are an adaptive baseline detection functionality. A naive baseline current detection algorithm using a moving average will inherently biased toward zero, as the events decreased current values contribute to the moving average. Instead, Transalyzer iteratively removes blockade values after detection to obtain the true baseline value. As an extra check to ensure proper data processing, all traces are initially manually trimmed, to remove any V_{bias} switching during an experiment and/or any irregularities such as long-term pore clogging and gating.

Modifications to the original Transalyzer code were implemented which decreased overall analysis time by a factor of two. Furthermore, bug-fixes were applied for current blockades at small dwell-times $\approx 1/f_{filt}$, effectively increasing the time resolution of the data analysis. Lastly, Translayzer's event detection software was expanded to also enable data processing of an episodically stimulated waveform from the Axon Axopath 200B used in this work. This enables researchers to use a fixed waveform with multiple defined V_{bias} values and process the resulting current trace. More in-depth report of these changes are included in the Appendix A.

Studies have reported that the standard deviation of the blockade signal [104], or even the spectral density at certain frequency ranges [139] can also be used to distinguish different protein species. These methods have been implemented but gave no added clarity in distinguishing populations. Alternative data processing techniques based on machine learning could also be useful for complex multi-state processes for which simple thresholding is not enough. Examples include Hidden-Markov modelling, or neural network algorithms. However, these were deemed to be outside of the scope for this study and were therefore not pursued any further.

3.2. Spectrophotometer unfolding studies

RNAse molecules have inherent fluorescent properties due to their aromatic residues. These can be used to monitor their folded state, as will be explained in Section 6. The unfolding of RNAse A in Urea and GdnHCl was studied using spectrophotometry (Agilent Cary Eclipse Spectrophotometer). To that end, $20\mu M$ of RNAse A was allowed to unfold at different concentrations of either Urea or Gdn-HCl overnight, containing 1M NaCl 10mM TRIS and 1mM EDTA for pH buffering and consistency with nanopore measurements. Sample concentrations were verified using DeNovix DS-11+ nanodrop, using an extinction coefficient of $\epsilon = 8640 cm^{-1} M^{-1}$ and molecular weight of 13.7 kDa [137].

The samples were then transferred to a quartz cuvette, which was cleaned thoroughly both inside and outside prior to the experiment with 1M KOH, Helmanex detergent, 70% ethanol and MiliQ water ($18.2 M\Omega$). Excitation wavelength was set to $280 nm$, and the emission spectrum was measured from $290 - 370 nm$. Background emission spectra obtained for solutions without RNAse molecules, and subtracted from the signal to normalize.

3.3. Sample preparation

Target sample details are listed in their respective sections below. All other materials can be found in section 3.4 Peptide and protein sequences are shown from N- to C-terminus.

3.3.1. Peptide D

A synthetic Peptide model with sequence *EEEEEEES^{*} GSGSGSGRRRRRRRRR* was obtained from Biomatik ($\geq 95\%$ purity from HPLC), including Phosphorylated and Glycosylated variants at the S^* residue. Acetylated variants were likewise obtained from Biomatik ($\leq 95\%$ Purity from HPLC) with the sequence *EEEEEEEK^{*} GSGSGSGRRRRRRRRR*, with K^* being the modified residue. All peptides were obtained in solid form, and resuspended in water. For long-term storage, they were kept at -20 deg C . Peptide D with a concentration on the order of $100 nM$ were pre-diluted in buffer solution $1M NaCl$, $10mM$ TRIS, $1mM$ EDTA and $4mM$ TCEP to a total volume of $100\mu L$ to promote homogeneous mixing in the flowcell. Further mixing was achieved by gently pipetting $100\mu L$ up and down three times. TCEP was functionally unnecessary, but was used for consistency with previous measurement containing disulfide bonds.

3.3.2. Arg-adaptor Ligation

A target peptide MUC5AC with sequence *GTT^{*} PSPVPTTST^{*} TSAP* was obtained from Anaspec ($\geq 95\%$, AS-61329), as well as its two O-GlcNAc glycosylated variants, located on residues marked T^* at positions 3 ($\geq 95\%$, AS-61334) and 13 ($\geq 95\%$, AS-61333) respectively. N-Hydroxysuccinimide (NHS) click-chemistry, was attempted with a synthetic Arg-Adaptor region, with sequence *RRRRRRRRRRHHHHH-(CH₂)₆-NHS* obtained from Thermo-Fisher ($\geq 95\%$) in solid form. It was resuspended in dimethyl sulfoxide (DMSO) to prevent spontaneous hydrolysis of the NHS-group in aqueous environments [44], and stored in -20 deg C .

The NHS-click chemistry is described in Hermanson [44], where it can be seen that the C-terminal end containing an NHS-group can be attached to the N-terminal α amine end of a target peptide, upon forming an amine bond. In this way, the Arg-Adaptor was attached to MUC5AC, forming a longer peptide. To pre-

vent cross-linking of Arg-Adaptors, its N-terminus was acetylated beforehand. The reaction was performed in a 10mM phosphate buffer at $pH = 7$, with a 10:1 molar ratio of Arg-adaptor:MUC5AC. The target peptide concentration was 1mg/mL . The sample was incubated for 4h, as NHS groups have been shown to nearly completely hydrolyse in aqueous conditions in under 4h, depending on the solution's pH value [44]. For control experiments, 800nM Arg-Adaptor was dissolved in water overnight to hydrolyse the NHS group, leaving only the peptide region of the adaptor.

The sample was purified using Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) on a Agilent 1260 Infinity machine, with a C-18 column, and an acetonitrile gradient ramping from 1-15% in 15min. The results were verified using MALDI-TOF mass spectrometry (Autoflex Speed) in negative ion mode. The matrix consisted of 10 mg/mL α -Cyano-4-hydroxycinnamic acid and 0.2% TFA.

3.3.3. RNase A, B and glycans

RNase A (≥ 70 Kunitz units/mg, R6513) and B (unknown purity, P7817S) from bovine pancreas was obtained from Sigma-Aldrich and New-England Biolabs respectively. Similar pre-dilution steps were performed for RNase A and B, where 400nM RNase was added to 1M NaCl, 10mM TRIS, 1mM EDTA and 4mM TCEP in a total volume of $100\mu\text{L}$. Measurements at acidic conditions were performed with 10mM Citric acid instead, at $pH = 4.2$. For measurements in Urea, 400nM RNase was denatured in 8M Urea and 4mM TCEP for atleast 1h and preferably overnight. The sample was mixed repeatedly by pipetting during the unfolding process and before injection into the flowcell. Experiments in denaturing conditions have been performed in 1M , 2M , and 4M Urea, 1M NaCl, 10mM TRIS, and 1mM EDTA at $pH = 7.2$, or 10mM citric acid at $pH = 4.2$.

Seeing as the RNase molecule diameter is larger than the narrowest pore constriction, we opt to following the protocol of Muthukumar [80] to prevent potential pore clogging due to simultaneous insertions of RNase molecules. To that end, we apply $+50\text{mV}$ for 0.5s, followed by 1.0s of 0mV , and finally the desired measurement voltage for 2.0s. For data analysis, the first 1.5s are omitted from the measurement window.

3-mannose(86%, CN-M3N2-10U), 8-mannose (90%, CN-MAN8-10U), as well as sialic acid (96.26%,CN-A2-10U) glycans were purchased from Ludger Bioquant. These were prediluted in the same way as Peptide D, and measured at multiple concentrations on the order of $1\mu\text{M}$.

3.4. Materials

HPLC-grade plus water was purchased from Merck (115333). 5M NaCl, 1M TRIS, and 1M EDTA stock solutions filtered at $0.2\mu\text{m}$ were obtained from Invitrogen. Sodium citrate ($\geq 99.5\%$, 71497), citric acid ($\geq 99.5\%$, 251275), sodium-phosphate monobasic ($\geq 99.0\%$, S8282), sodium-phosphate dibasic ($\geq 99.0\%$, S7907), urea ($\geq 99.0\%$, U5128), guanidine hydrochloride ($\geq 99.0\%$, 50940), DMSO ($\geq 99.9\%$, 472301-M), pentane ($\geq 99\%$, 236,705), hexadecane ($\geq 99\%$, H6703) and TCEP (C4706) were obtained from Sigma Aldrich. Absolute ethanol was obtained from Honeywell ($\geq 99.98\%$, 40347H) 0.50mm diameter silver wire ($\geq 99.99\%$, AG5486) was obtained from Advent. 1,2-diphytanoyl-sn-glycero-3-phosphocholine were purchased from Avanti Polar Lipids ($\geq 95\%$) in powdered form and resuspended in pentane. Consumer-grade bleach with a concentration of < 5% sodiumhypochlorite was used. Wild-Type FraC octamers in solution were generously provided by the Maglia Lab in the University of Groningen, as well $25\mu\text{m}$ thick polytetrafluoroethylene (PTFE, teflon) films.

4

Model Peptide: Peptide D

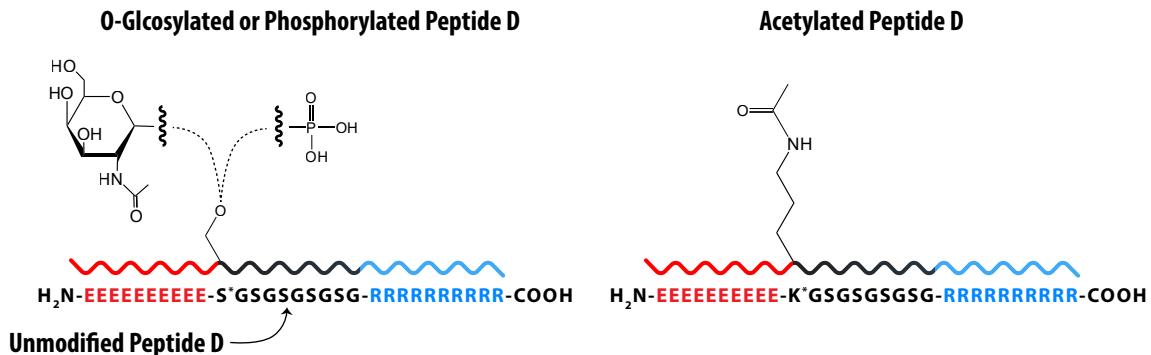


Figure 4.1: Diagram of Peptide D structure and PTM variants. Peptide D principally has two oppositely charged stretches, consisting of negatively charged glutamic acid (E) and positively charged arginine (R) residues. The regions are connected by alternating glycine (G), and serine (S) residues. Acetylation was performed on the lysine closest to E-stretch (left), whereas for O-linked Glycosylation and phosphorylation a serine is placed at that position instead. Note that unmodified peptide D refers to the construct on the right without PTMs.

In this section we detail the results of measuring the PTMs of lysine acetylation (AcetylK), O-GlcNAc glycosylation (O-Glyc) and phosphorylation (PO_3^{2-}) using the model peptide platform Peptide D (PepD)¹. It has been schematically represented in Figure 4.1, with on the left the AcetylK variant and on the right the O-Glyc and PO_3^{2-} variants. The design of this peptide was meant to tackle some common hurdles in nanopore-based analysis of peptides. Specifically, analytes can have very short translocation times $\ll 1\text{ms}$ [36], or can contrarily hardly be captured at all by the nanopore. What is more peptides also do not have a defined orientation whilst translocating and could be folded [119], giving rise to different interactions with the nanopore and thus complicating data interpretation. To solve these problems, ten positively charged arginine residues (R, pKa = 12.48) and negatively charged glutamic acid (E) (pKa = 4.07) residues were placed at each end. In the $V_{bias} < 0$ case primarily used in FraC, the positive R-stretch will serve as a pointer, electro-phoretically guiding the peptide towards the nanopore whilst also increasing capture rate. When the negative E-stretch enters the electric field, the peptide will experience an opposition force, effectively stalling the construct. Combined efforts of both stretches will then hypothetically result in a more defined orientation and linearisation of the molecule. A linker region consisting of 5 pairs of flexible glycine (G) and small, hydrophilic serine (S) residues connect the two charged regions. It must be noted that a functionally similar construct was proposed by Asandei et al. [10].

Relevant to PTM detection, it was hypothesized that a (transient) force equilibrium is achieved within the nanopore, where the forces on E- and R-stretch cancel. This should then give rise to extremely long read-lengths, enabling one to extensively probe the equilibrium position of the peptide until it is thermally disturbed and translocated. Substantial evidence was found for the existence of such an equilibrium position by tagging PepD with large hydrophobic molecules of different size, at various locations along PepD

¹The letter D was used simply due to it being the fourth peptide model to be tested.

(Restrepo-Perez, unpublished work). Tags resulted in current modulation only for specific voltages and position combinations, suggesting that there is a well defined probing region. What is more, it was found that the position was consistent over multiple measurements, and tunable by varying V_{bias} or the salt concentration. For $V_{bias} = -90\text{ mV}$ in 1M NaCl solution, it was found to be situated near the 11th residue, adjacent to the E-stretch, see Figure 4.1. Consequently, PTM'ed variants of PepD have been synthesized with PTMs strategically placed at this position, maximizing the use of the PepD construct, see Figure 4.1. Note that the references to unmodified PepD refer to the construct on the right-hand side, without PTMs. PTM'ed variants would then hypothetically give characteristic increases in ionic current blockades due to the addition of a chemical group, enabling their identification through current measurements. In the next section we show the experimental results thereof, and discuss the data with a focus on data interpretation.

4.1. Peptide D characterization

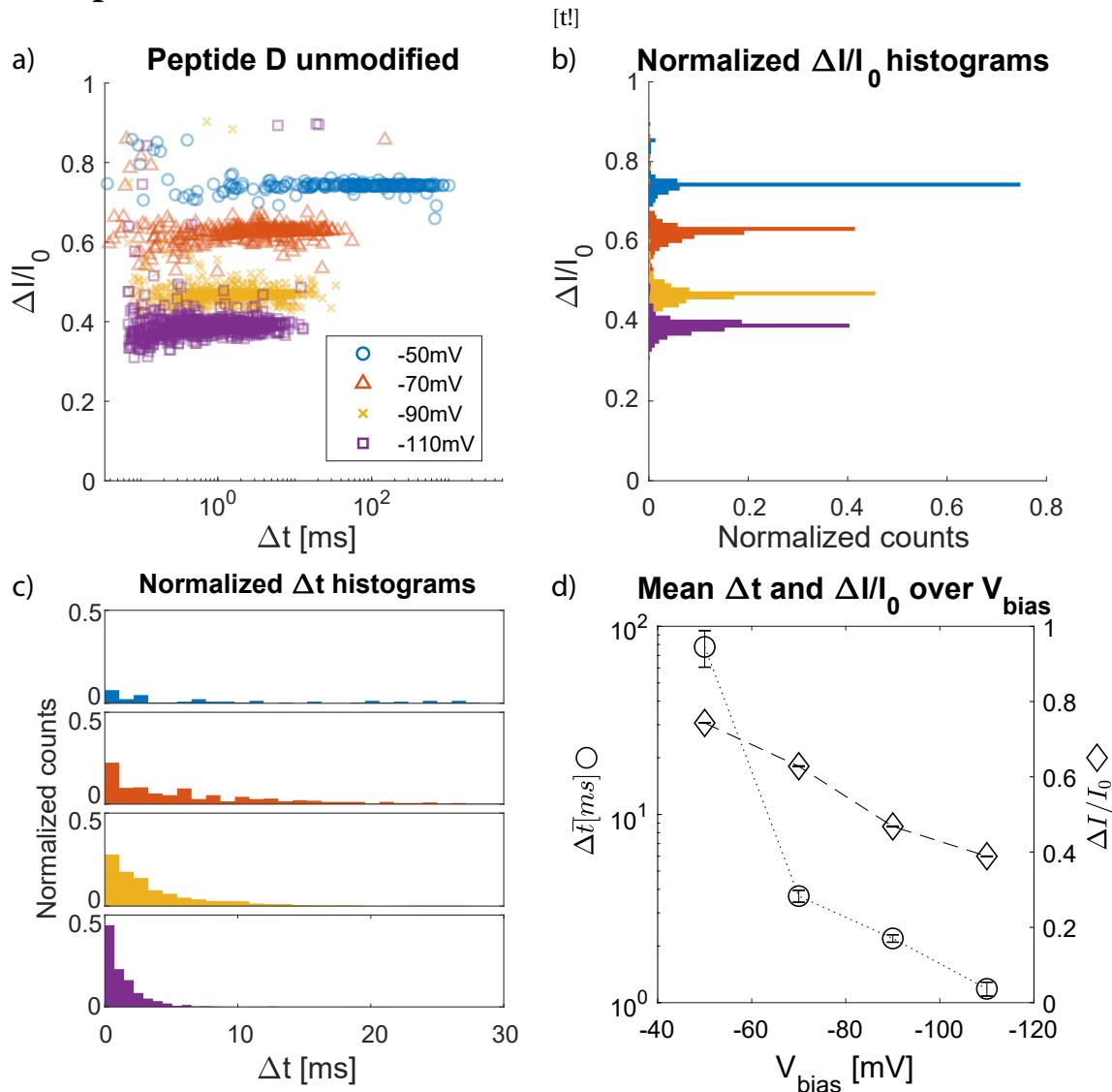


Figure 4.2: Measurement of unmodified PepD. Each color and shape combination corresponds to a given V_{bias} value. **a)** $\Delta I/I_0$ vs. Δt graph, where each point represents a single event in the current-time domain. Consistent blockade levels are found at each voltage. **b)** Normalized histogram of $\Delta I/I_0$, showing that the populations are indeed very consistent in $\Delta I/I_0$, with little spread. **c)** Normalized Δt histograms, with each window representing a given voltage. Clear exponential distributions can be recognized for $V_{bias} \leq -70$. **d)** Statistical averages obtained from b) and c), summarized in one plot. The left and right y-axis correspond to Δt and $\Delta I/I_0$ respectively.

To introduce data representation methods of nanopore experiments in general, we start by characterizing unmodified PepD's behaviour under different V_{bias} values. Current-time measurements were performed in 1M NaCl, 10mM TRIS and 1mM EDTA solution at $pH = 7.2$, followed by event identification and extraction

methods outlined in Section 3. To visualise a population of events, we choose to plot their relative current blockade $\Delta I/I_0$ against the dwell-time Δt in Figure 4.2a. We have opted to normalize the blockade current ΔI with the baseline current I_0 to account for slow drifts due to e.g. evaporation of water. As a result, 0 and 1 on the y -axis refers to open and fully blocked pore respectively. The x -axis indicates dwell-times Δt in ms , shown in logarithmic scale.

A general trend is visible in the event populations of Figure 4.2a, suggesting that both $\Delta I/I_0$ and Δt decrease with magnitude² $|V_{bias}|$. To further quantify this, we analyse the $\Delta I/I_0$ and Δt distributions separately in Figure 4.2b-c. Very sharp distributions $\Delta I/I_0$ are visible in Figure 4.2b, indicating that the blockade level is consistent and well defined. The exponential distributions in Figure 4.2c get progressively steeper, indeed indicating a decrease in mean Δt . Note that the x -axis in Figure 4.2 is converted back to a linear scale in this plot. Also note that the distribution of $V_{bias} = -50mV$ does not fit an exponential very accurately. Most likely, more events are required for a better fit. By subsequently fitting Gaussian and single-exponential functions on $\Delta I/I_0$ and Δt distributions respectively, statistical averages can be obtained.

These have been summarized in Figure 4.2d over the measured V_{bias} range. The error-bars indicate the 95% fitting confidence interval for Δt and one s.d. for the $\Delta I/I_0$. Indeed, both $\Delta I/I_0$ and $\Delta \bar{t}$ are observed to decrease monotonically with $|V_{bias}|$. The decrease in $\Delta I/I_0$ has been hypothesized to be a result of stretching the PepD molecule under influence of the opposing electrical forces (*Restrepo-Perez, unpublished work*), with molecular dynamics simulations giving supporting evidence for this claim. Next, the decrease in Δt with $|V_{bias}|$ is typical for translocation events: higher V_{bias} leads to stronger electro-phoretic forces, pulling the peptide through the nanopore faster. Analytes that are not able to translocate will instead be more strongly pulled into the pore region, increasing their Δt . Δt versus $|V_{bias}|$ trends are thus generally used to distinguish translocation from blocking, although it is still an active area of discussion whether this method is reliable [47]. We accept this line of thinking, and therefore conclude that PepD is translocating at $|V_{bias}| \geq -50mV$.

²We take the magnitude of V_{bias} , to prevent confusion regarding the polarity of V_{bias} . Negative applied potentials are used due to ionic current rectification at $V_{bias} > 0$

4.2. Distinguishing Peptide D from PTM'ed variants

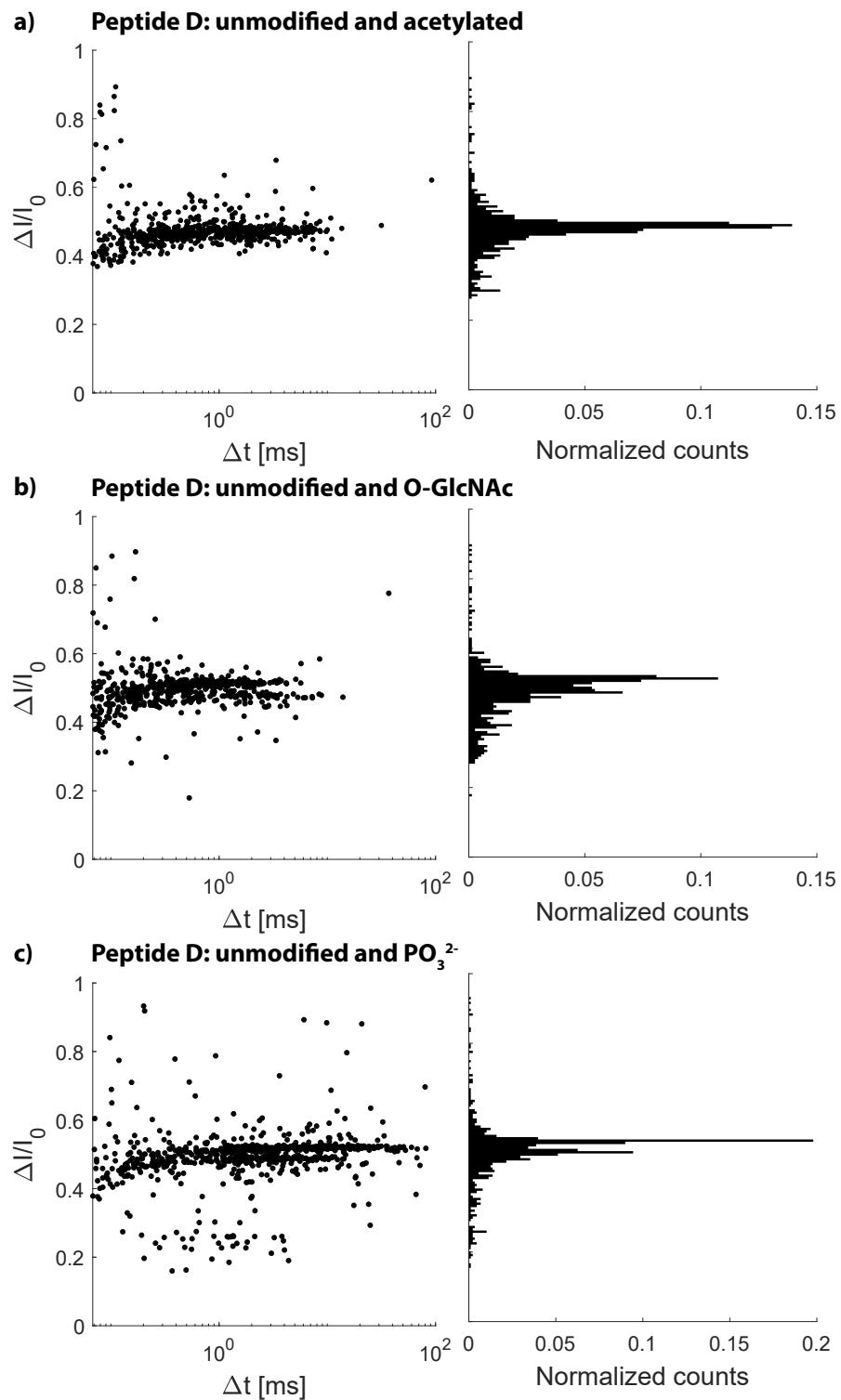


Figure 4.3: Nanopore translocation data for PepD and a PTM'ed variant. Measurements were performed at $V_{bias} = -90\text{mV}$, in 1M NaCl, 10mM TRIS, 1mM EDTA. All current traces were Gaussian filtered with a bandwidth of 10kHz for data analysis. Measurements with two species were performed in equimolar amounts. **a)** PepD and acetyl PTM. No distinct second blockade level could be found, even when further zooming in on the population. **b)** PepD with O-GlcNAc PTM. A distinct second $\Delta I/I_0$ level can be found above the PepD level, with an overall decrease in Δt . **c)** PepD with Phosphorylated variant, where a second level can be seen in $\Delta I/I_0$ once again, but with an increased Δt .

We now move on to the focus of this study, namely the detection of PTMs, where we will use the sensitive probing region of PepD to its full extent. In a series of experiments, a single PTM variant was injected with equimolar amounts of unmodified PepD, with the aim of distinguishing them from each other. Measurements were performed at $V_{bias} = -90mV$ at total peptide concentrations of $200nM$. The measurement results are summarized in Figure 4.3. From the previous section we obtain for unmodified PepD $\Delta I/I_0 = 0.467 \pm 0.009$ and $\Delta t = 1.5611 \pm 1.1ms$ was found, where the uncertainty represents s.d. over 3+ measurements. This will be used as a reference.

Figure 4.3a shows the result of measuring equimolar amounts of AcetylK and unmodified PepD. No discernible second population arises, suggesting that $\Delta I/I_0$ of both species are very similar. Two repeat measurements were performed, over a range of V_{bias} values between $-30mV$ to $-110mV$, all with similar results. It was concluded that the addition of a small 43 Da, uncharged moiety simply falls outside of the detection limit of this technique, and serves as a lower limit to the sensitivity of this platform. More advanced data processing was briefly pursued in the form of spectral density analysis. However, no distinct signals could be recognized by eye, and further attempts at employing e.g. machine learning tools were not attempted in the interest of time.

The next measurement contains O-GlcNAc glycosylation together with unmodified PepD. In Figure 4.3b, two distinct populations are discernible in their $\Delta I/I_0$ values, each with comparable densities. This is in line with the expected result of an equimolar measurement, leading us to conclude that the nanopore is able to distinguish O-GlcNAc PTMs from unmodified PepD. Upon comparing $\Delta I/I_0$ values from previous control measurements, the lower population can be identified as unmodified PepD, meaning that the upper one must correspond to the O-GlcNAc variant. Further analysis of two repeat measurements show that the O-GlcNAc variant has an average $\Delta I/I_0 = 0.504 \pm 0.02$, a 5.8% increase compared to unmodified PepD. For the dwell time we find $\Delta t = 0.405 \pm 0.24$, 74.1% lower compared to unmodified PepD.

We therefore observe that adding a O-GlcNAc moiety to PepD increase the blockade, but somehow also decreases the dwell time. The $\Delta I/I_0$ increase is attributed to the O-GlcNAc volume, blocking more pore than the unmodified serine residue. For the decrease in Δt , we propose two possible mechanisms. The first is based on entropic considerations of confining O-GlcNAc. From the $\Delta I/I_0$ modulation we can be reasonably confident that the PepD construct brings the O-GlcNAc moiety into the pore constriction, limiting its degrees of freedom. As a result, the equilibrium state created by the charged regions of PepD is less energetically favourable for the Glycosylated variant than for the unmodified variant. This increases the probability to thermally disturb the force equilibrium, leading to shorter dwell-times. The other is of enthalpic nature, and considers the interactions of O-GlcNAc with the pore. A general trend was observed experimentally and in Molecular Dynamics simulations, where Δt decreased with increasing hydrophilicity (*L. Restrepo-Perez, Unpublished Work*). The O-GlcNAc acts through its hydrophilic qualities. The latter mechanism must however be further explored and quantified.

Finally, we move on to the PO_3^{-2} variant, the results of which is shown in Figure 4.3c. Again, we can see two clear populations of similar density, from which we can conclude that PO_3^{-2} PTMs can be distinguished using nanopores as was reported previously in literature [104]. The $\Delta I/I_0$ of the bottom population once again corresponds to unmodified PepD. The PO_3^{-2} PTM was found to have $\Delta I/I_0 = 0.506 \pm 0.005$, a 6.2% increase compared to unmodified PepD. Note that this is almost identical to O-GlcNAc blockade levels, despite the molecular weight of PO_3^{-2} being less than half that of O-GlcNAc. Average $\Delta t = 3.211 \pm 2.1ms$ was found, amounting to a 141.6% increase compared to Peptide D.

For PO_3^{-2} we find an increased average $\Delta I/I_0$ and Δt . To explain the relatively strong current blocking prowess of PO_3^{-2} despite its small size (78Da), it was hypothesized that EDL of counter-ions leads to an effective increase of size. In line with this, we note that $\lambda_D = 0.3nm$ is of the same order of magnitude as O-GlcNAc's atomic cross section. If this were conceptually true, the $\Delta I/I_0$ value tunable with ionic strength by eq. 2.13. We pursue further this idea in the next section. Moving on to the increase of Δt , we attribute the increase in Δt to a charge imbalance in the construct. The addition of PO_3^{-2} shifts the force equilibrium brought about by the positive and negative tail regions of peptide. As a result, Δt increases on average. From our results it can be observed that the FraC is very sensitive to negatively charged moieties, both in $\Delta I/I_0$ and Δt .

4.3. Distinguishing O-GlcNAc and Phosphorylated variants

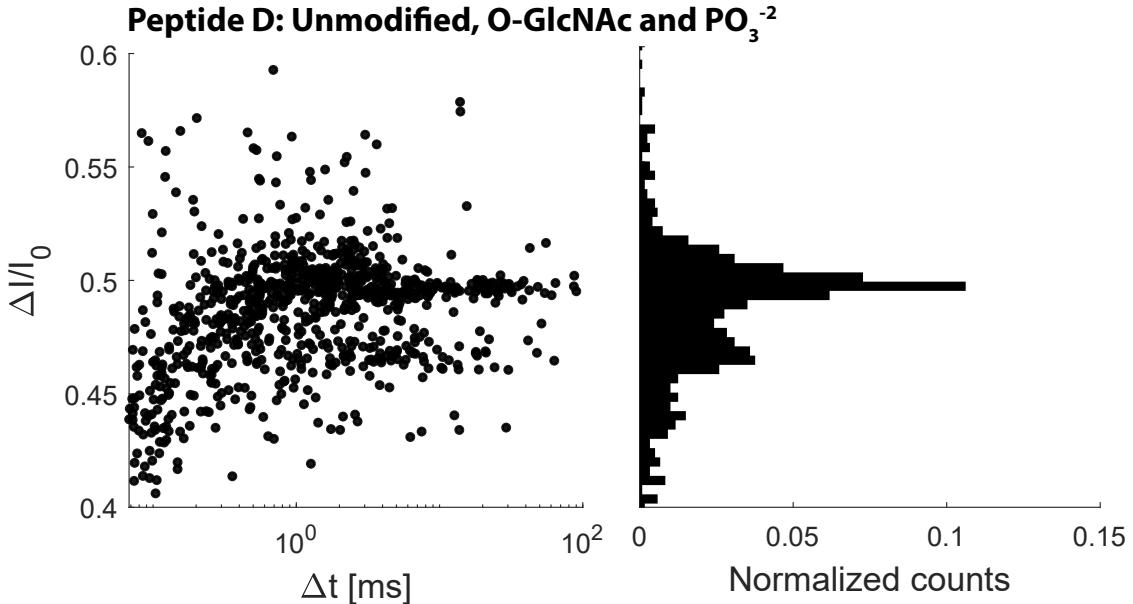


Figure 4.4: $\Delta I / I_0$ vs. Δt for unmodified, O-GlcNAc and PO_3^{-2} PepD variants at 1M NaCl. Note that the y-axis is zoomed in. Measurement was performed in equimolar conditions of the three species. Two levels are visible, with the top one approximately twice as dense as the bottom population.

With the successful distinction of O-GlcNAc and phosphorylated peptides from unmodified PepD, we confidently march on to the biologically relevant distinction of O-GlcNAc and PO_3^{-2} *amongst themselves* as described in Section 2.1.3. The previous datasets show that the $\Delta I / I_0$ levels of O-GlcNAc and PO_3^{-2} are very similar. Indeed, when measuring equimolar amounts of unmodified, O-GlcNAc and PO_3^{-2} PepD, only two populations are revealed, as visible in Figure 4.4. Counting the events, we find a 2.27:1 ratio between the upper and lower populations respectively, further suggesting that the upper one contains two species.

To separate the O-GlcNAc and PO_3^{-2} populations one needs to somehow adapt one species' $\Delta I / I_0$ selectively. Further to the discussion on the effect of PO_3^{-2} current blocking behaviour, we explore the idea of varying the ionic strength, and with it λ_D . For the sake of exploring this conceptually, we make the simple assumption that EDL surrounding the PO_3^{-2} moiety leads to a spherically shaped volume increase, with radius λ_D . We further naively assume that $\Delta I / I_0 \propto V_{phos}$. Using eq. 2.13, we can then calculate the effective increase in $\Delta I / I_0$ when lowering the salt concentration from 1M to 0.8M as follows³:

$$\Delta I / I_0 \propto V_{phos} \propto \lambda_D^3 \propto I_s^{-2/3}$$

, giving an increase of 16%, whilst leaving the O-GlcNAc blockade strength unchanged. From the above we expect $\Delta I / I_0 = 0.511$ for Phosphorylated PepD in 0.8M NaCl. Naturally, the ionic strength also affects EDLs at the pore surface and the charged regions of PepD, which in turn also affects hydrodynamic contributions to the current. It is therefore essential to normalize the $\Delta I / I_0$ values of the PTM'ed variants with that of unmodified Peptide D.

³Other salt concentrations have been tested, but have proven to be unfruitful. A summary was added in Appendix B.2.

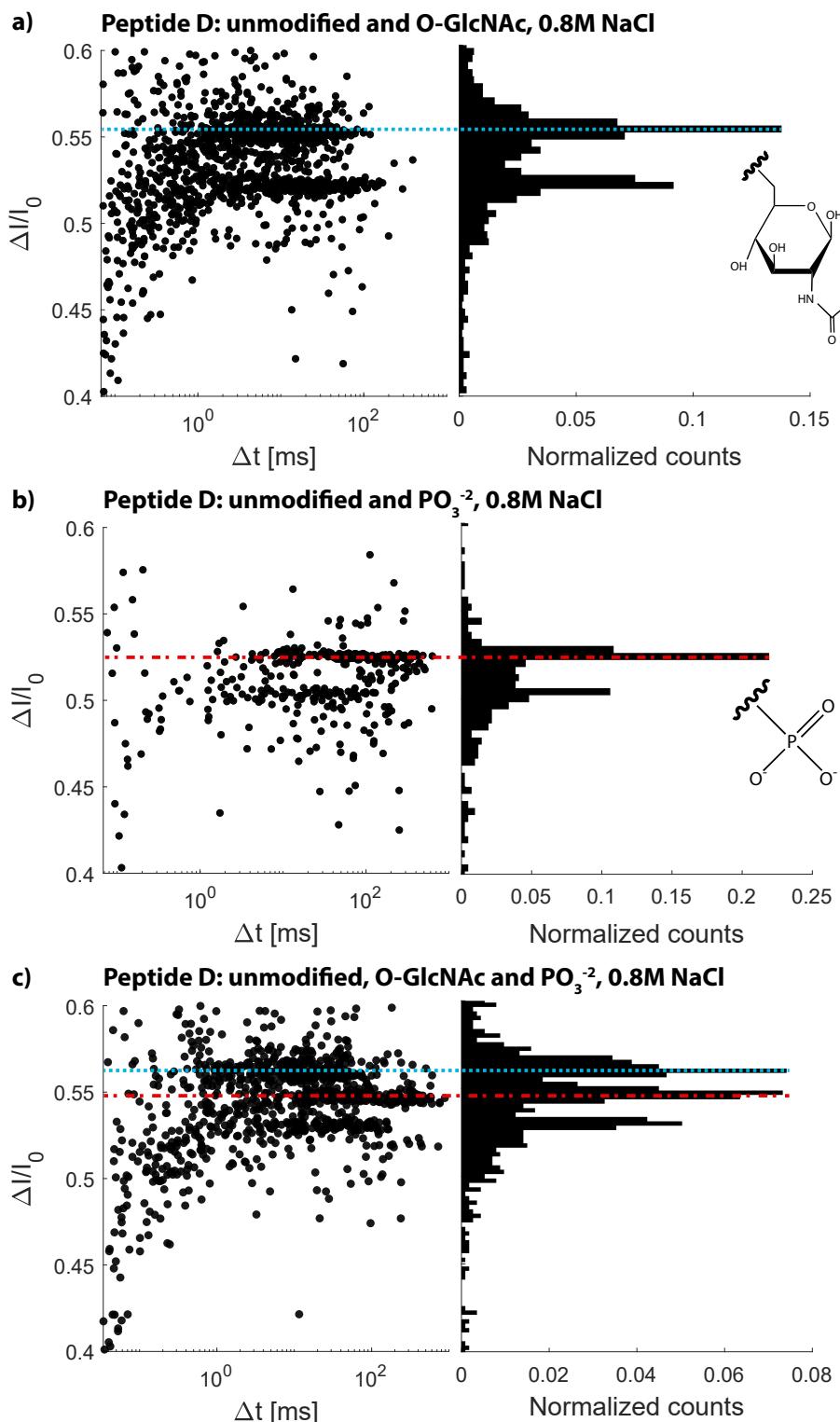


Figure 4.5: $\Delta I / I_0$ vs. Δt for unmodified, O-GlcNAc and PO_3^{2-} PepD variants at 0.8M NaCl. Note the zoomed in y-axis for all graphs. **a)** Equimolar mix of 75nM unmodified PepD and PO_3^{2-} at 0.8M NaCl. The bottom population can be identified as unmodified PepD, and the top population as PO_3^{2-} PepD (red dashed dotted line). Note that the average Δt is increased compared to unmodified PepD, similar to the 1M NaCl case. **b)** An equimolar mix of 75nM unmodified PepD and O-GlcNAc variant. The top population can now be identified as O-GlcNAc (dashed blue line). Note the decrease in average Δt , similar to 1M NaCl. **c)** Same experiment as in a), but at 0.8M NaCl. The appearance of a third level allows us to identify O-GlcNAc, PO_3^{2-} and unmodified PepD from top to bottom using results in b) and c).

Table 4.1: Comparison of blockade at 1M NaCl and 0.8M NaCl. Average $\Delta I/I_0$ of PTM'ed variants are normalized by unmodified PepD's $\Delta I/I_0$ at 1M NaCl and 0.8M NaCl. Standard deviation is computed from at least 3 measurements. It is seen that O-GlcNAc blockade stays constant, whilst the PO_3^{-2} decreases with salt concentration.

	1M NaCl	0.8M NaCl
unmodified	1	1
O-GlcNAc	1.06 ± 0.02	1.06 ± 0.02
PO_3^{-2}	1.06 ± 0.01	1.04 ± 0.01

To that end, the measurements were performed in a similar manner to Section 4.2, using 0.8M NaCl concentration instead. The results are shown in 4.5. Figure a) and b) show the results of mixing equimolar quantities of unmodified PepD with O-GlcNAc and PO_3^{-2} variants respectively. We again recognize two distinct $\Delta I/I_0$ levels, corresponding to the PTM'ed and unmodified variant. It was also observed that the characteristic dwell time increase and decrease of PO_3^{-2} and O-GlcNAc respectively is maintained at 0.8M NaCl.

The next measurement was naturally of an equimolar mixture of all three. The experimental results of O-GlcNAc, PO_3^{-2} and unmodified PepD mixture at 0.8M NaCl with total peptide concentration of 150nM are shown in Figure 4.5c. These conditions enabled us to confidently recognize three separate populations, as visible in both the scatter plot and histogram. Using the results in Figure B.1a-b, we can also confidently assign each species to their respective event populations, based on their shift in average Δt . The populations indicated with a blue dotted and red dashed-dotted line then correspond to O-GlcNAc and PO_3^{-2} respectively, with the bottom population corresponding to PepD. Two independent repeat measurements were performed using different stock solutions, giving similar results. These are added in Appendix B.1.

With this, we see that nanopores have the capability of detecting O-GlcNAc and PO_3^{-2} , as well as distinguishing them amongst each other on the PepD platform. A potential next step is to adapt peptides found *in-vivo* with positive and negative stretches, with the aim of creating similar sensitive probing regions. We explore this option in the next Chapter.

Although we are now able to distinguish these PTMs, we must note that our initial hypothesis of PO_3^{-2} 's salt-dependent current blocking behaviour was incorrect. To see this, we compute the $\Delta I/I_0$ of O-GlcNAc and PO_3^{-2} over at least 3 measurements, for 1M NaCl and 0.8M NaCl. We normalize this value by the $\Delta I/I_0$ values found for unmodified PepD, to account for overall changes in behaviour of PepD itself. The results are shown in Table 4.1. As seen in the previous section, the blockades of O-GlcNAc and PO_3^{-2} are identical at 1M NaCl. At 0.8M NaCl, we see no change in O-GlcNAc normalized $\Delta I/I_0$, but we do observe a *decrease* in normalized $\Delta I/I_0$ for PO_3^{-2} . This is in contrast with our expected 16% increase calculated above.

We speculated that the decrease in blockade of the PO_3^{-2} moiety is due to a similar effect as the inverted current rectification bias. We have seen in Section 2.4.4 that when pore diameter $d_{pore} \approx \lambda_D$, the polarity of rectification is reversed, leading instead to an *increase* in current for $V_{bias} < 0$. In line with this reasoning, it is speculated that the confinement of PO_3^{-2} on the time scale of ms essentially creates an annulus in the pore constriction, with decreased cross-sectional area and increased surface charges. This would then hypothetically lead to stronger overlap of the λ_D regions, an increase in current and finally a decrease in $\Delta I/I_0$ compared to the 1M NaCl case.

5

Arginine-Adaptor: Proof-of-concept

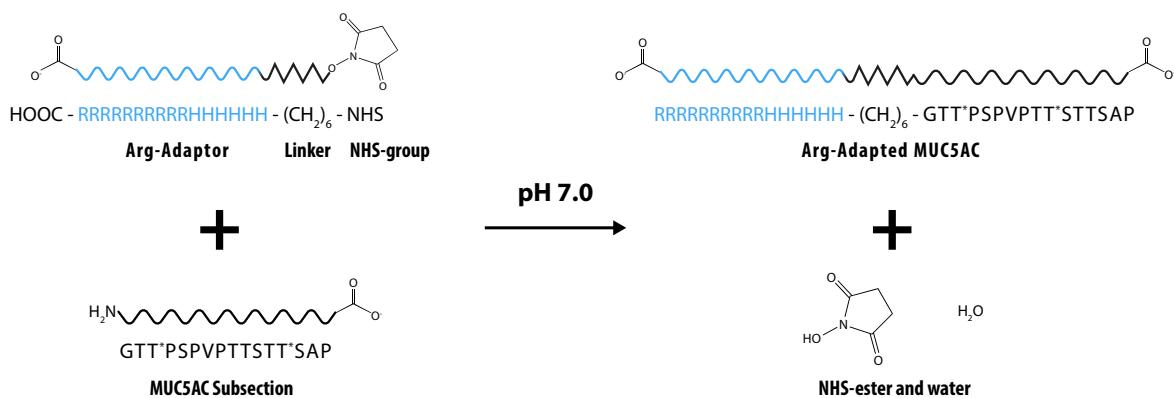


Figure 5.1: NHS reaction scheme for the Arginine Adaptor attachment. The adaptor region consists of six His and 10 Arg residues, abbreviated as 6H10R and coloured blue to indicate positive charge. The target is MUC5AC, where two Thr indicated as T* are potential glycosylation sites. Reaction was performed at pH = 7.0.

In this chapter we present our efforts in building a proof-of-concept Arginine-adaptor region, with the ultimate goal of attaching a peptide-based stretch of positive charges to any target protein or peptide. As discussed in the previous section, nanopore analysis of protein is complicated due to heterogeneity in folded structures, charge distribution and orientation with respect to the pore [47, 102]. Furthermore, uncharged molecules can only be captured electro-osmotically, which was experimentally found in this Chapter to severely limit capture rate.

Analogous to the PepD platform, the Arg-Adaptor is designed to tackle these problems. A stretch of positive charges at a terminal end hypothetically provides enough electrophoretic forces to unfold secondary structures [11, 104], adds to the directionality of the target molecules, and facilitates the electrophoretic transport of uncharged molecules. Conceptually, this idea was already worked out by Biswas et al. [15] and [104], with the key difference being their use of polynucleotides instead of amino acids. What is more, their proposed ligation methods that were not generally applicable.

In an attempt to improve generality, we propose the reaction schematically represented in Figure 5.1. On the top left hand side, the Arg-Adaptor is shown, consisting first and foremost of a peptide with 16 positive residues *HHHHHHRRRRRRRR*¹. NHS-functionality was to the N-terminus via a $(CH_2)_6$ linker.

The NHS-ester can mediate ligation to the amine of any target peptide/protein. At $pH \leq 7$, it was found that N-terminal α -amine specificity can be achieved over ϵ -amine on lysine side-chains [44]. This gives the current construct the potential of adapting a random target protein, given that their N-terminal amine is solvent exposed and unmodified².

¹A histidine tag was for Ni-NTA purification. It was later found that HPLC-purification was more effective, rendering the His-tag obsolete.

²The author is aware that this is not generally the case *in-vivo*, but stresses that NHS chemistry was chosen for its practical simplicity to verify the proof-of-concept. Other potential ligation strategies are discussed in Section 7.

To test this ligation scheme, a synthetic 16 amino acid subsection of the human MUC5AC O-linked mucin-type glycoprotein was selected as a target peptide. For simplicity, the sequence contains no lysine residues and is uncharged to test the Arg-Adaptor's capability of mediating electrophoretic transport. *In-vivo*, MUC5AC is secreted in the nasal pathways to create mucus. For that reason, two O-linked glycosylation sites can be found position Thr3 and Thr15, marked with a * sign. In the following sections we report the results of the ligation reaction and the experimental effect of Arg-Adaptor attachment to MUC5AC, including preliminary results on detecting the Thr3 O-linked glycosylated variant.

5.1. Ligation and purification

Arg-Adaptor ligation to MUC5AC

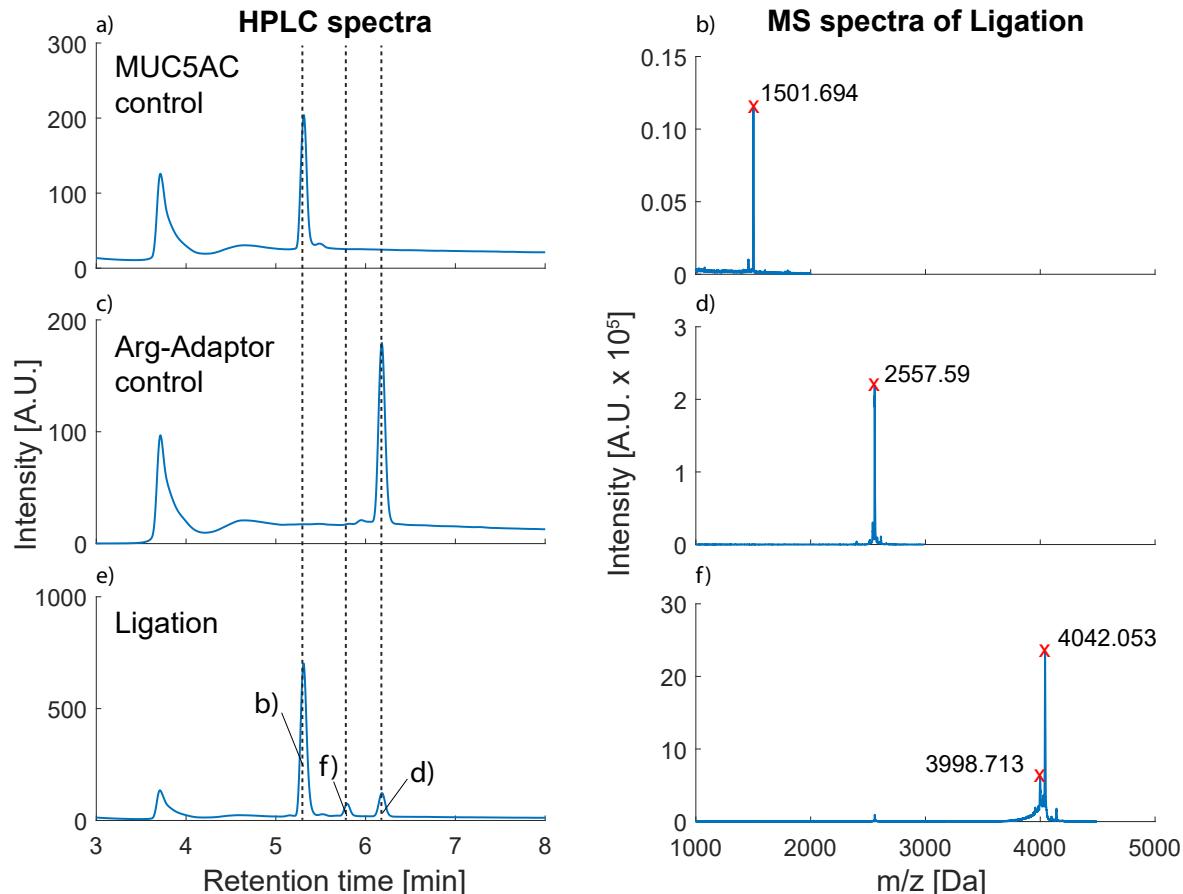


Figure 5.2: Purification of NHS-ligation reaction. The left column shows HPLC spectra measured at a wavelength of 214nm, whereas the right column shows MALDI-TOF mass spectra in negative ion mode. HPLC spectra all show a void peak around 3.7 min inherent to the HPLC method. **a)** and **b)** show control measurements of pure MUC5AC and Arg-Adaptor samples. A single well-defined peak is observed at 5.3min and 6.2min respectively. **c)** HPLC measurement of NHS-reaction volume. Great overlap was found with the control experiments, allowing the identification of the MUC5AC and Arg-Adaptor within the spectrum. The third peak is then recognized as the Arg-Adapted MUC5AC. The peaks are subsequently extracted and subjected to MS. **b), d) and f)** then show the mass spectra of the indicated peaks. The m/z values correspond to the theoretical values within 1 Da.

The ligation was performed at pH = 7.0 for 4h at a 10x molar excess of Arg-Adaptor, with a concentration of 1mg/mL of MUC5AC. The entire reaction volume was analyzed and purified by RP-HPLC. The resulting fractions were then individually analyzed by MALDI-TOF MS. The results are shown in Figures 5.2 a-f. The left hand column shows HPLC spectra, with the y-axis showing the fluorescent emission intensity at 214nm wavelength, and the x-axis the retention time in the column. Figure 5.2a and c are control measurements of pure MUC5AC and Arg-Adaptor samples respectively. For Figure 5.2c, the Arg-Adaptor was incubated in water for 4h, to hydrolyze the NHS-ester. In both cases, two distinct peaks can be seen. The first peak around 4minutes is the so-called void-peak, inherent to HPLC. The second peak is then identified as control sample.

The NHS-ligation reaction volume was then injected into the HPLC, resulting in Figure 5.2e. Here, excellent overlap is found with the control experiments, as indicated with the dashed line. However, we also see a smaller peak at 5.8min, which supposedly is the Arg-Adapted MUC5AC. This results suggests that the ligation was not very efficient, as much of the MUC5AC is still unmodified. No exact quantification was done.

We then extract the fractions shown in Figure 5.2 and subjected them to MALDI-TOF MS analysis in negative ion mode. These resulting mass spectra are shown in Figure 5.2b, d, f. The theoretical mono-isotopic masses of the reagents and products have been found to be MUC5AC 1501.6Da, Arg-Adaptor 2558.9Da and Arg-Adapted MUC5AC 4042.5Da. This corresponds to within 1 Da of the peak values found in Figures 5.2b, d and f respectively. Slight deviation are ascribed to Carbon-13 isotopes and measurement errors. A second peak was found in Figure 5.2 at 3998.713 Da, which could potentially be accounted for by arginine oxidation [13]. This was however not further explored. From the above, we conclude that substantial evidence was found for the successful ligation of MUC5AC and Arg-Adaptor visualized in Figure 5.2f.

5.2. Experimental effect of adaptor region

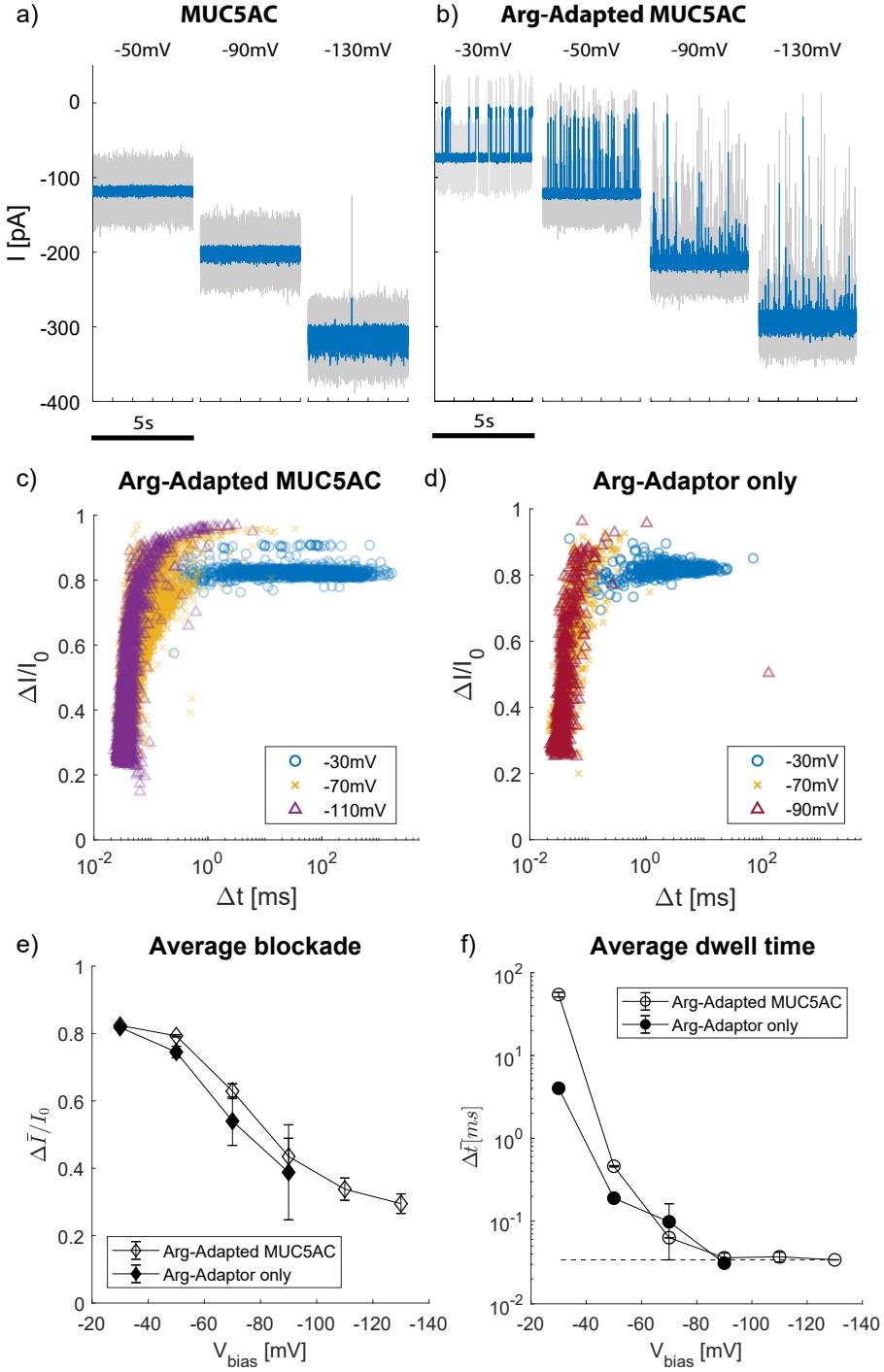


Figure 5.3: Measurement results for Arg-Adaptor ligation to MUC5AC. **a)** Example traces for $5\mu M$ MUC5AC for increasing V_{bias} values, with the gray and blue signal filtered at 10kHz and 2kHz respectively. Practically no events were detected. **b)** Example traces after adding the Arg-Adapted sample purified using HPLC. A striking increase in capture rate can be seen at all voltages. It can also be seen that the dwell-time decreases with $|V_{bias}|$. **c)** Relative blockade versus dwell time scatterplot of the Arg-Adapted MUC5AC. Colors and markers indicate different V_{bias} values. Long dwell times were found for $-30 mV$, which quickly decreased with V_{bias} . A transition in $\Delta I/I_0$ value is observed at $-70 mV$. **d)** Control experiment with only the hydrolyzed Arg-Adaptor. Dwell-times are significantly lower, although blockades are comparable to the Arg-Adapted MUC5AC. **e)** Gaussian fit of blockade distributions, with errorbars indicating one s.d. of the gaussian. A strong decrease in blockade is found, mainly attributed to filter distortions. Of note is the large s.d. at $-70 mV$ due to the transition. **f)** Single exponential fit results of dwell time distributions, with errorbars indicating 95% confidence intervals. Note the strong decrease with V_{bias} . The dashed line indicates the filter limit.

We now attempt to quantify the effect of the Arg-Adaptor on the nanopore translocation behaviour of MUC5AC. To that end, $5\mu M$ of MUC5AC injected into the flowcell, mixed by pipetting and allowed to equilibrate for 15 minutes. A series of measurements were subsequently performed at a range of $-30 < V_{bias} < -130$, and representative traces were plotted in Figure 5.3a. The gray and blue trace were filtered at a bandwidth of 10kHz and 2kHz respectively. Hardly any capture was recorded, attributed primarily to MUC5AC's charge neutrality. Event rates of less than $0.5Hz$ were found for all voltages, making the acquisition of a statistically relevant population sizes ($N > 100$) practically challenging. What is more, the events that are visible could potentially be attributed to low-concentration contaminants in the solution, based on reference measurements. Repeat experiments at $8\mu M$ revealed similar low capture rates. The concentration could be increased by an order of magnitude to further increase capture rate, but was not further attempted in this work.

Within the same measurement, the HPLC purified Arg-Adapted MUC5AC at a maximum concentration of $5\mu M$, assuming 100% ligation efficiency and full sample retention in the purification process. The actual concentration is most likely an order of magnitude lower. Nevertheless, a striking increase in event rate was found for all attempted V_{bias} values within seconds of addition. Representative traces are given in Figure 5.3b. It can also already be observed that average Δt decreases with $|V_{bias}|$.

To further characterize Arg-Adapted MUC5AC's behaviour, $\Delta I/I_0$ versus Δt scatter plots are shown in Figure 5.1c. For $V_{bias} \geq -30mV$, a population with $\Delta I/I_0 = 0.824 \pm 0.001$ can be found, with the error indicating one s.d. of a gaussian fit. Overall a decrease in Δt is seen with increasing $|V_{bias}|$. Interestingly, a transition occurs in $\Delta I/I_0$ at $-70mV$, where a second level can be observed at $\Delta I/I_0 \approx 1$. This suggests that at stronger $|V_{bias}|$, Arg-Adapted MUC5AC is in a conformation that blocks more current. What is more, for $V_{bias} = 110mV$, the original blockade level is seemingly completely absent. We also see that at this high V_{bias} , most events are close to the filter bandwidth limit, giving rise to the characteristic downward shaped curve due to filter distortions.

Control experiments are performed by measuring only the Arg-Adaptor, to observe how much of the behaviour can be attributed to only the Arg-Adaptor itself. To that end, 200nM Arg-Adaptor was first incubated in water for 4h to incubate the NHS-ester, and subsequently measured in a nanopore. The result is shown in Figure 5.3d. $\Delta I/I_0 = 0.818 \pm 0.001$ was found, very similar to the ligated construct, suggesting that most of the blocking can be attributed to the Arg-Adaptor itself. Similar to the Arg-Adapted MUC5AC construct, the average dwell time quickly decreased with $|V_{bias}|$, leading to filter distortions. However, there was no strong evidence for a transition to an increased $\Delta I/I_0$ value for higher V_{bias} .

We further quantify these results by performing gaussian and single-exponential fits on the $\Delta I/I_0$ and Δt distributions. These have been shown in Figures 5.3e and f respectively. The errorbars indicate one gaussian s.d. and 95% fitting confidence intervals. With this, we can verify that both the average $\Delta I/I_0$ and Δt decrease with $|V_{bias}|$. The former is attributed mainly to the increasing effect of filter distortions at high $|V_{bias}|$, biasing the the $\Delta I/I_0$ value toward zero. The latter indicates that both the Arg-Adapted MUC5AC and the Arg-Adaptor on its own translocate. We also note a striking decreased average Δt for the Arg-Adaptor control, indicating that the MUC5AC slows down translocation.

From the above we conclude that the addition of the Arg-Adapted construct has a significantly increased capture rate compared to bare MUC5AC. This showcases the Arg-Adaptor's capabilities of electrophoretically facilitating transport of uncharged peptides. Furthermore, it was shown that the Arg-Adapted MUC5AC and Arg-Adaptor control experiment were significantly different in Δt , indicating that the MUC5AC region has some interaction with the pore. To explain the Arg-Adapted MUC5AC's transition in $\Delta I/I_0$ at high V_{bias} , we propose that the constructed could be forced into a coiled state due strong electrophoretic forces, transiently blocking more current.

6

N-linked Glycosylation: Ribonuclease A and B

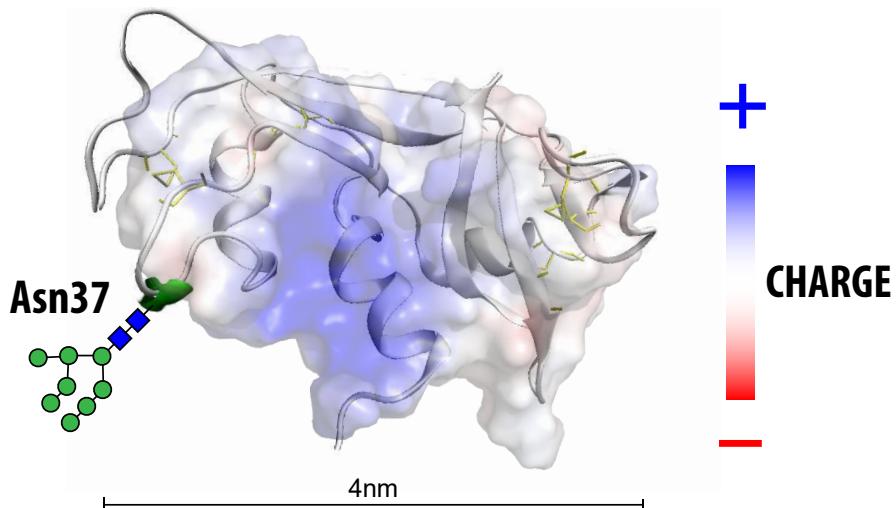


Figure 6.1: Structural overview of RNase A and B. RNase contains a positively charged pocket for the binding of RNA molecules, to perform its degradation function. Four disulfide bonds indicated in yellow give added structural stability to the protein. For RNase B, a mannose rich glycan group is attached at the Asn37 position, indicated in green. The glycan can contain, 5-9 mannose subunits, with the 8 mannose variant shown on the left as an example [108, 134]

In this chapter, we present the results of measuring a full-length protein: Ribonuclease A (RNase A). These proteins bind and cleave RNA molecules in the cell, either degrading or adapting them wherever necessary. The substrate was chosen due to the availability of a N-linked glycosylated variant *in-vivo*, referred to simply as Ribonuclease B (RNase B). RNase B was shown to be practically identical in structure, with the noteworthy exception of one single N-linked glycosylated site at Asn37 [108, 134]. On this site, we find mannose-rich glycans with 6-8 mannose subunits [134]. A diagram of an 8 mannose glycan, together with a cartoon representation of Ribonuclease A is shown in Figure 6.1. NMR studies have shown that the movement of residues within 50 amino acids of 37-Asn were significantly hampered. This suggests that the glycan enhances protein stability by steric hindrance [48, 52, 94]. It was also shown that the surface charge was affected by decreasing the degree of solvent-exposure of charged residues [52].

With the increased structural complexity of RNase A/B compared to the samples reported in the previous chapters, additional challenges are presented with regards to nanopore analysis. First and foremost are size considerations. It was found that the longest diameter of RNase is found to roughly 4nm, more than 3 times the size of the narrowest constriction in FraC [108, 134]. As a result, RNase A and B are both not expected to translocate in their native state. Previous reports confirm that RNase A indeed does not translocate within

the context of an α -Haemolysin nanopore [65]. Therefore it is not deemed not very likely that the glycan itself will be probed by the narrowest constriction of FraC in a consistent manner. As a matter of fact, the glycan itself is already quite formidable, with molecular weights ranging from 1.4 - 1.8 kDa being comparable to the MUC5AC subunit analyzed in Chapter 5. To combat this problem it was deemed necessary to unfold the protein, such that it does translocate, hypothetically allowing the glycan to transiently pass the pore constriction. In an ideal case, RNase B is fully linearized and translocates from terminus to terminus. The current traces of RNase B would then reveal an additional blockade increase on top of the protein backbone signal, indicating the glycan's translocation. Fundamental to PTM detection is that it would then also hypothetically give information regarding its location, based on the timing of the glycan's passage through the nanopore. The second challenge is the stability of RNase. RNases are often used as protein models due to their tremendous stability in denaturing conditions and high melting point [81, 88, 98]. Four disulfide bonds created by pairwise binding of 8 cysteine residues have been found to be critical for this stability [26, 98]. Cross-linking of the linear chain provides significant resistance to denaturing and high temperature conditions, as these mainly affect non-covalent interactions [81, 88]. These challenges make RNase an excellent case study for PTM detection within the context of full protein, as many of the problematic properties described above are not unique to RNase [86, 102].

In the rest of this chapter, we present our attempts at differentiating RNase A from B, with the ultimate goal of directly probing the PTM on RNase B. As a first step, we measure RNase A and B in their native state. We then show our efforts of translocating purified mannose-rich glycans. Finally, we measure in a series of different denaturant concentrations and pH = 4.2, with the goal of destabilizing the RNase structure [68, 89] to promote translocation.

6.1. RNase A and B: Native and acidic state

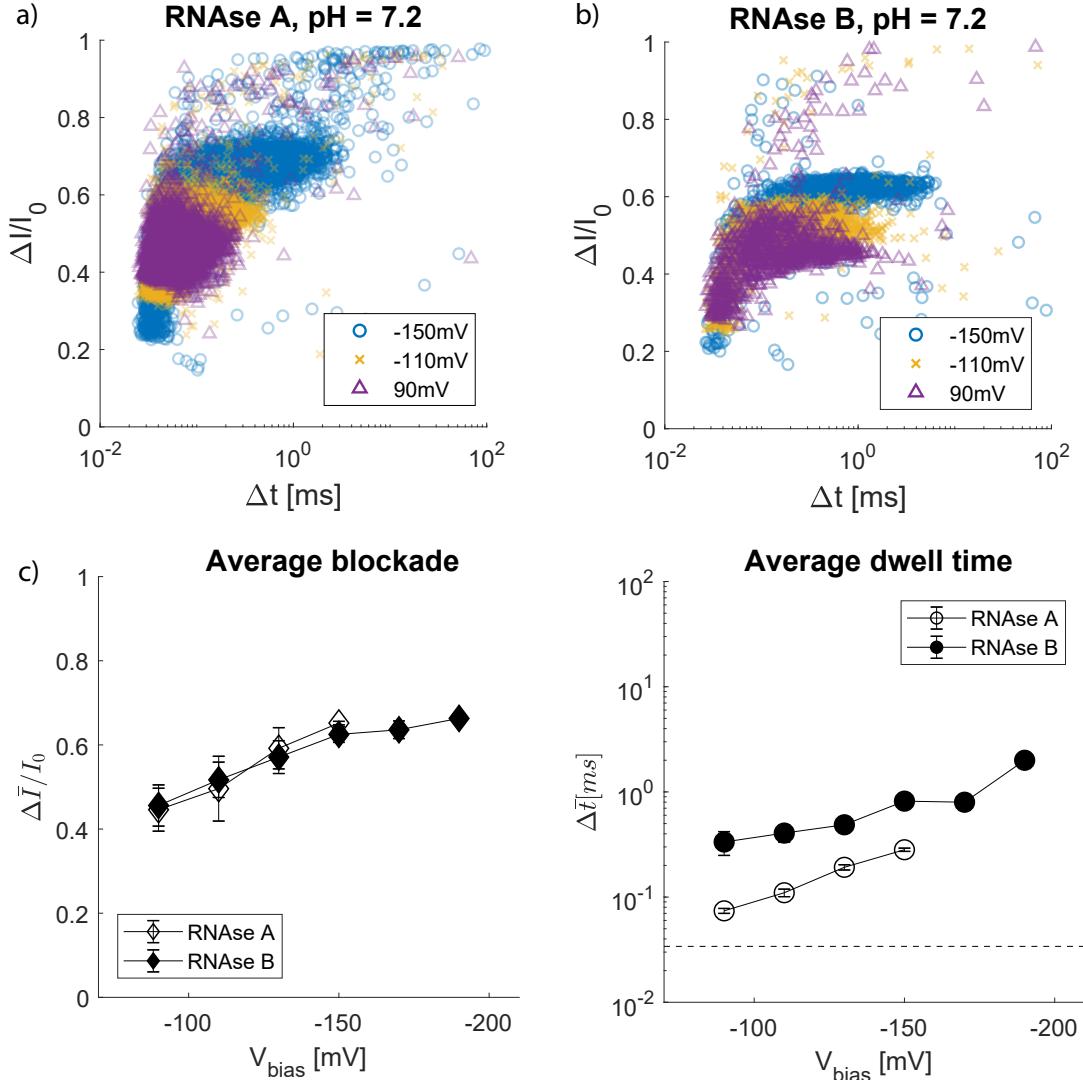


Figure 6.2: Blockades and dwelltime results for RNase A and B at pH = 7.2. The measurements were performed at a concentration of 200nM, in a solution of 1M NaCl, 10mM TRIS, 1mM EDTA and 1mM TCEP. **a**) Scatterplot of RNase A at different V_{bias} values. Both the average blockade and dwell-time increase with $|V_{bias}|$. **b**) Scatterplot of RNase B at the same V_{bias} range. A smaller spread in blockade is seen for every voltage. **c**) Gaussian fit results on $\Delta I/I_0$ distributions, with errorbars indicating one s.d. within the Gaussian. Hollow and filled markers refer to RNase A and B respectively. A monotonic increase is observed over the measured V_{bias} range. **d**) Single exponential fit result on Δt distributions, with errorbars indicating 95% fitting confidence intervals. Hollow and filled markers indicate RNase A and B respectively. Dwell times are found to increase over the entire V_{bias} range. Note the logarithmic y-axis.

To start exploring RNase A and B, we measured them in their native state at pH 7.2 in a solution of 1M NaCl, 10mM TRIS and 1mM EDTA. The resulting relative blockade versus dwell-time scatter plots are shown in Figure 6.2a and b respectively for RNase A and B. The colours and markers indicate different V_{bias} values. For both RNase A and B, an increasing trend can be seen in average $\Delta I/I_0$ and Δt . It can also be seen that the spread in blockade is larger for RNase A in comparison to B, particularly noticeable at $V_{bias} = -150\text{mV}$. We further quantify the behaviour of both molecules by fitting $\Delta I/I_0$ and Δt distributions with a Gaussian and single exponential fit respectively and plotting their average values. The results are visible in Figure 6.2c and d, where the hollow and filled markers refer to RNase A and B respectively. Errorbars refer to a single Gaussian s.d. and 95% fitting confidence intervals respectively. From Figure 6.2c, we can conclude that for both species, the average $\Delta I/I_0$ increases monotonically over the measured V_{bias} range. This indicates that for stronger applied voltages, a larger portion of the current is blocked by the molecule. Interestingly, average $\Delta I/I_0$ value of both species is very similar, despite RNase B having a $\approx 10\%$ larger molecular weight due to

the glycan attachment. Figure 6.2d then shows that the average Δt also increases exponentially, noting the logarithmic scale on the y-axis. This leads us to conclude that RNase A and B or both not translocating within the measured V_{bias} range. What is more, it can be observed that RNase B has consistently higher average Δt over the measured V_{bias} range.

The decreased spread in $\Delta I/I_0$ for RNase B is attributed to structurally stabilizing effect of its N-linked glycan structure [48, 52, 112]. Previously reported simulation results show that the unfolding energy barrier of RNase A was increased by 20% due to the presence of the glycan [112]. This is then thought to counteracts the formation of destabilized intermediates with different $\Delta I/I_0$ values. With regards to the monotonic increase of $\Delta I/I_0$ and Δt values with $|V_{bias}|$ for both RNase A and B, we speculate that the RNase molecules are pulled into the pore with increasing force, leading to higher blockades and longer dwell-times. Previous studies have shown this general trend for proteins significantly larger than the constriction of FraC [47]. It is however still surprising that RNase B has practically the same average $\Delta I/I_0$, especially considering that the glycan structure points radially outward [134]. RNase B could potentially be consistently oriented with its glycan away from the pore, explaining why the $\Delta I/I_0$ remain unchanged compared to RNase A, but no direct evidence for this was found. With regards to the increased average Δt of RNase B compared to A, we speculate that the glycan reduces the effective surface charge of the molecule. In line with this thinking, previous studies were found showing that RNase B's glycan structure reduces the degree of solvent exposure of residues near Asn37 [48, 52]. The reduced surface charge would then allow the RNase B molecules to stay within the charged regions of the pore in a more energetically favourable manner.

Seeing as the iso-electric point of RNase A is reportedly $pI = 9.83$ [125], it was believed that RNase A would behave predominantly as a positively charged molecule at $pH = 7.2$. It is therefore surprising that relatively strong $V_{bias} \leq -90mV$ must be applied to obtain dwell times well above the filter bandwidth limit. A previous report has indicated that RNase A has a negative surface charge [65], whereas other authors have shown that this is highly dependent on the valency of the salt solutions [92]. As a result, the surface charge of RNase A in the present conditions must still be further explored by performing e.g. ζ -potential measurements.

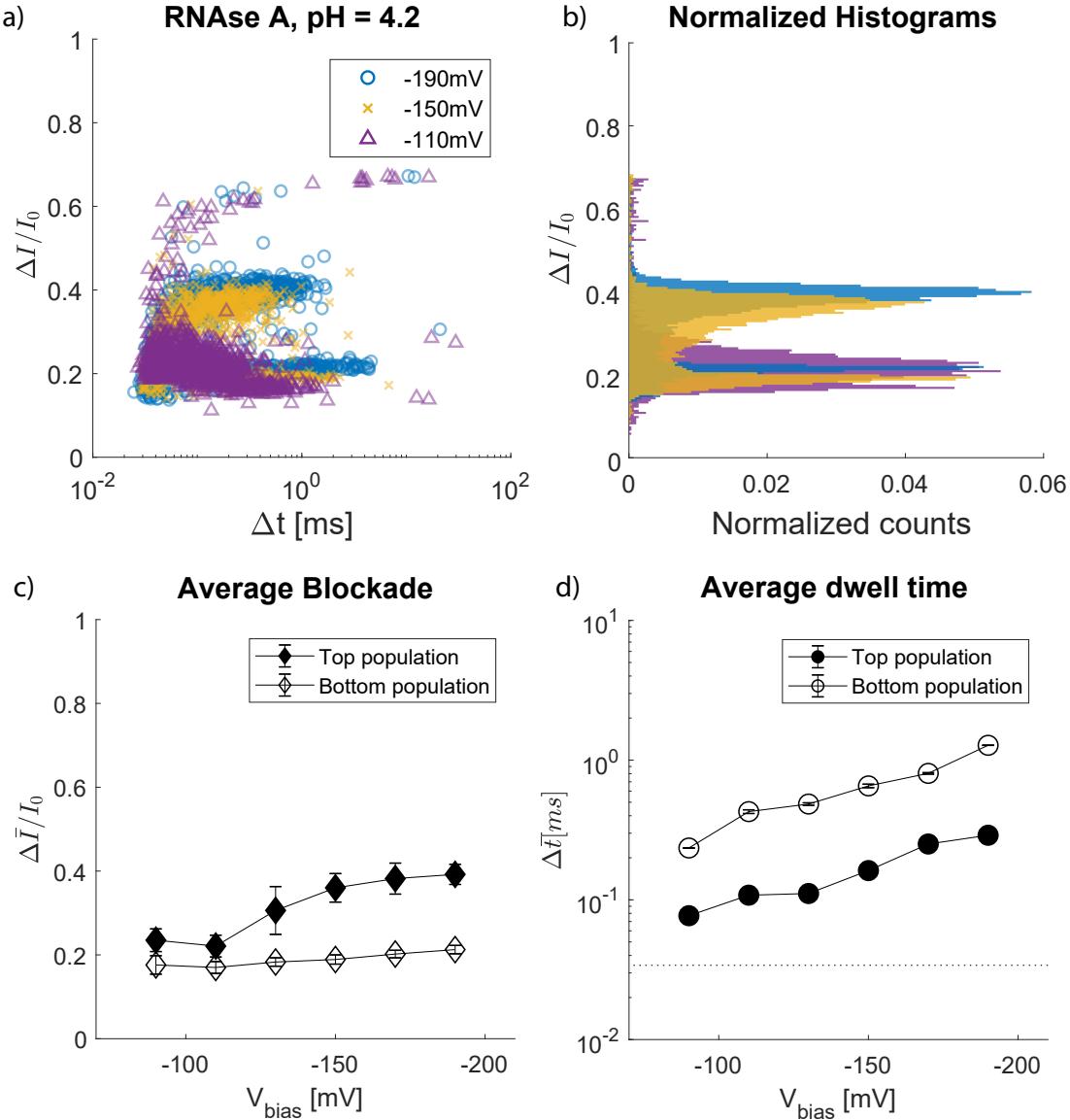


Figure 6.3: Measurement of RNase A at pH = 4.2. Measurement were performed with 200nM sample, at 1M NaCl, 10mM citric acid, 1mM EDTA and 1mM TCEP. **a)** Blockade versus dwell time scatterplot. Colors and markers indicate different V_{bias} values. Two populations are visible at blockades of 0.4 and 0.2. **b)** Histogram of $\Delta I/I_0$, normalized with total event count. The splitting in populations is visualized more clearly. **c)** Average blockades obtained by Gaussian fitting, with errors indicating one s.d. of the Gaussian fit. Filled and hollow markers refer to the top and bottom population respectively. **c)** Average dwell time obtained by single-exponential fitting, with errorbars indicating 95% fitting confidence intervals. The dashed line represents the filter bandwidth limit. Filled and hollow markers refer to the top and bottom population respectively. Both populations increase in average Δt with $|V_{bias}|$.

In an attempt to circumvent problems with RNase's surface charge, we instead measure at acidic conditions of $pH = 4.2$ to increase the degree of protonation for the solvent-exposed residues. Previous studies have also shown that RNase A is more prone to unfold at acidic conditions [89], which contribute to our overall goal of unfolding RNase. The measurement results for RNase A at $pH = 4.2$ are visualized in Figure 6.3. The notation is kept consistent with Figure 6.2, with the exception of V_{bias} values. Drastically different behaviour was found in acidic conditions for RNase A, both in blockade and dwell time. Most notably in Figure 6.3a, two distinct populations can be recognized for strong $|V_{bias}|$ around $\Delta I/I_0 = 0.2$ and 0.4 . Both these populations seem to increase in $\Delta I/I_0$ as well as Δt with $|V_{bias}|$. Populations are more distinct in the $\Delta I/I_0$ histogram shown in Figure 6.3b. From this plot, we can clearly see two populations even at $V_{bias} = -110\text{ mV}$. As was done multiple times in this work, we further quantify the results by fitting. The events were split into two populations based on their $\Delta I/I_0$ value. The exact border between the two population was chosen some-

what arbitrarily, but was seen not to influence fitting results significantly. With that said, Figure 6.3c allows us to verify the increase in $\Delta I/I_0$ of both populations with $|V_{bias}|$. Filled and hollow markers refer to the top and bottom populations respectively. As could be seen in Figure 6.3a, the top population has a stronger increase in blockade with V_{bias} , but seems to plateau at a value of $\Delta I/I_0 = 0.4$. When comparing with the blockades at $pH = 7.2$ of Figure 6.2, we observe that RNase A blockades have in fact decreased. For example, the $\Delta I/I_0$ value at V_{bias} have decreased from 0.652 ± 0.004 at $pH = 7.2$ to 0.360 ± 0.034 $pH = 4.2$. Conversely, average Δt have increased at acidic conditions, as concluded from Figure 6.3d. We furthermore recognize an increase in average dwell time with $|V_{bias}|$ for both populations, indicating that RNase A does not translocate at these conditions.

The two populations arising at acidic conditions are thought to be two different partially destabilized conformations of RNase A, as a result of the acidic conditions [89]. It is thought that the increased positive charge of RNase A at $pH = 4.2$ then facilitates longer dwell-times, due to stronger electrophoretic forces. This, however, seems to contradict the decrease in $\Delta I/I_0$ at acidic conditions, and must therefore be explored further. We speculate that only a partly linearized section of RNase is trapped in the constriction, instead of the entire protein.

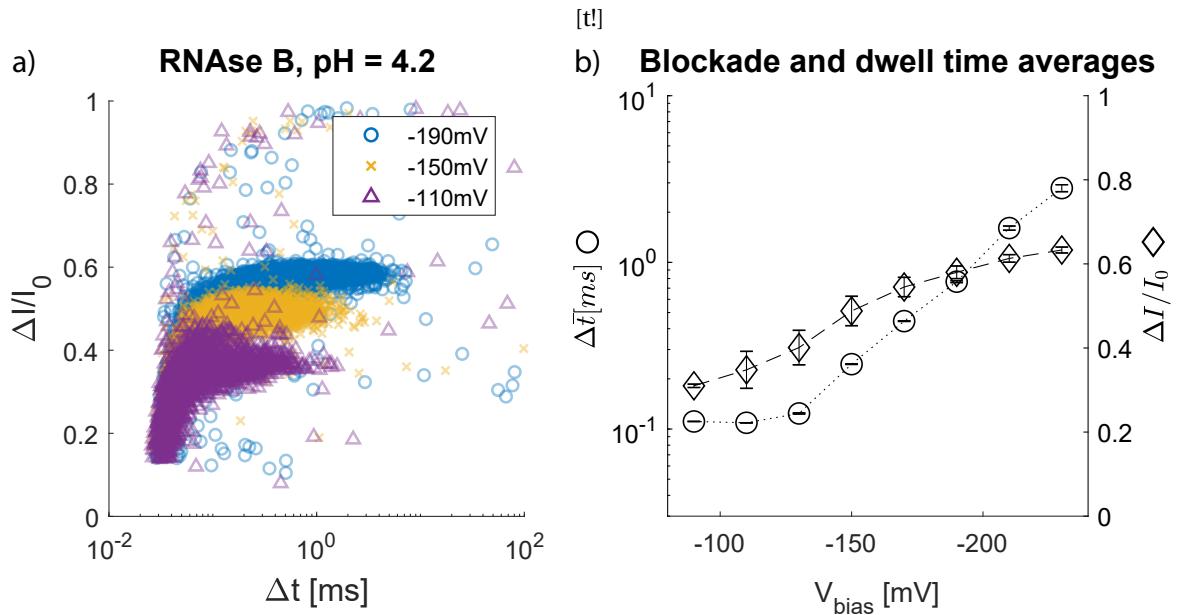


Figure 6.4: Measurement results of RNase A and B at $pH = 7.2$, measured at a concentration of 200nM, in a solution of 1M NaCl, 10mM TRIS, 1mM EDTA and 1mM TCEP. a) Scatterplot of RNase A at different V_{bias} values. b) Scatterplot of RNase B at the same V_{bias} range. c) Fitting results of the Δt (circles) and $\Delta I/I_0$ (diamonds) distributions. Hollow and filled markers indicate RNase A and B respectively.

Measurements in the same conditions are performed for RNase B, resulting in Figure 6.4. Notably, no second population was found, attributed to the superior stability of RNase B [48, 52]. On first glance, the relative blockade versus dwell time scatter plot in 6.4a looks almost identical to the result obtained at $pH = 7$ (Figure 6.2b). From our fitting results in Figure 6.4b, we can identify a key difference: although in both the $pH = 7.2$ and $pH = 4.2$ conditions blockade levels plateau $\Delta I/I_0 \approx 0.6$, the plateau is reached faster at weaker $|V_{bias}|$ at neutral $pH = 7.2$. This is in contradiction with our expectation of increased electrophoretic force at acidic conditions, due to the increase in positive charge. Repeat experiments must be performed to verify this behaviour.

It must at this point be noted that RNase B molecules can be clearly distinguished from RNase A molecules at $pH = 4.2$, based on their relative blockades. Interestingly, this was achieved without any evidence for RNase translocation. Similar results were reported by Fahie and Chen [34], who showed that localizing Streptavidin near the opening of an OmpG nanopore gave enough resolving power to distinguish it from the N-linked glycosylated variant avidin.

6.2. Translocating Purified Glycans

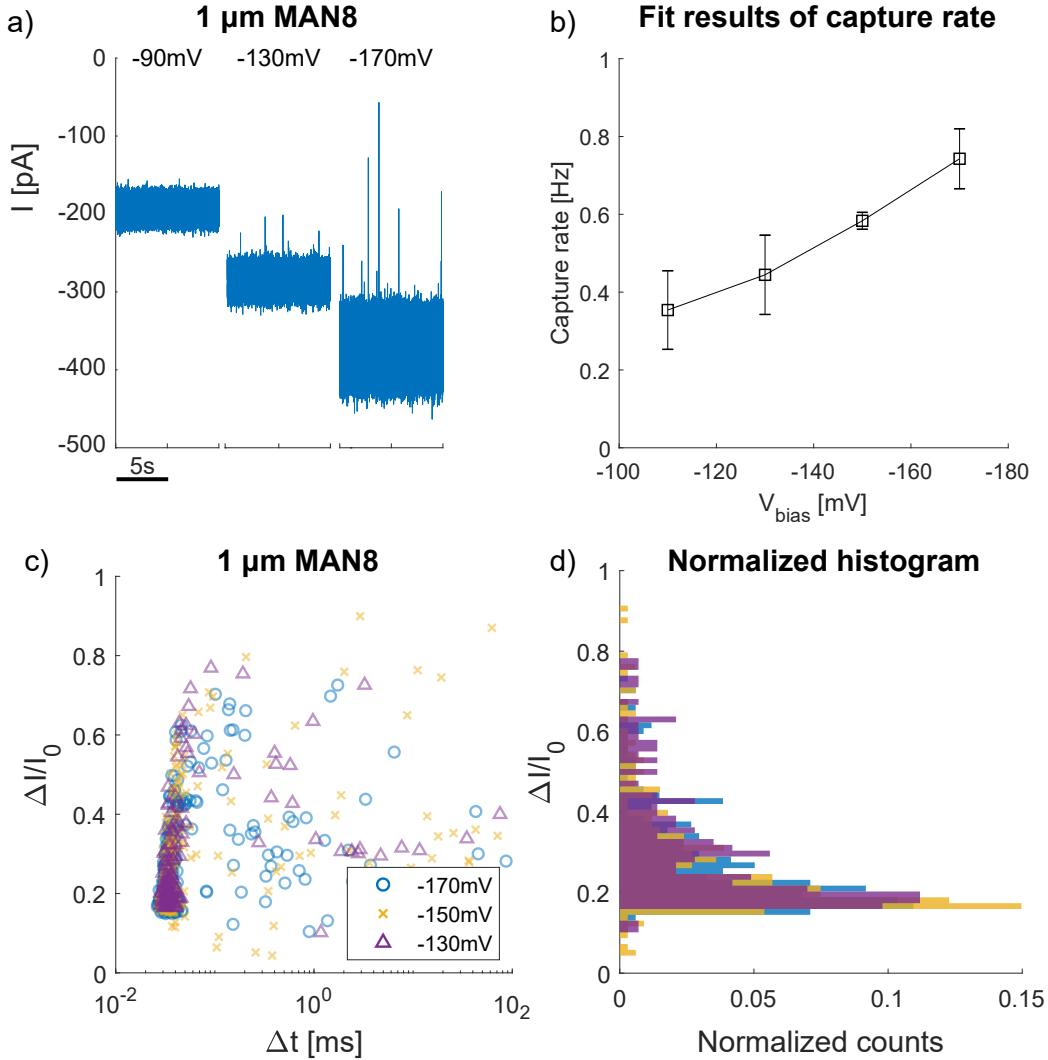


Figure 6.5: Summary of measurement on 8mannose glycan. **a)** Example current traces for increasing $|V_{bias}|$. **b)** Fit results of time-between-events, using a single-exponential. Average time-between-events was plotted, with errorbars indicating 95% confidence intervals. **c)** Scatter plot of $1\mu M$ MAN8. **d)** Corresponding blockade histogram.

As an intermediate step towards probing the N-glycan on RNase B directly, we first present our efforts of translocating purified 8-Mannose glycans. At the same time, we explore the possibilities of performing nanopore-based glycan analysis, also known as glycomics. Similar to PTM detection, large scale glycan characterization faces similar problems with regards to structural complexity and sample heterogeneity [105]. Nanopore-based glycomics has only recently very recently been explored [55, 131]. At the time of writing, no literature was found on characterizing glycan structures using biological nanopores.

To that end, the present preliminary results of measuring $1\mu M$ 8-mannose glycans (MAN8) in $1M$ NaCl, $10mM$ TRIS and $1mM$ EDTA at $pH = 7.2$. Figure 6.5a shows example traces for 3 V_{bias} voltages, demonstrating the low capture rate of the molecule. This was further quantified by fitting a single-exponential to the distribution of time-between-events. The fitted exponent is then the average capture rate. The fitting results are summarized in Figure 6.5b, where the errorbars signify the 95% confidence intervals. A near-linear increase in capture rate is found with V_{bias} , which is indicative of a diffusion-limited capture process. It must be noted that the capture rate is relatively low, even at $V_{bias} = -170mV$, with a value of 0.74 ± 0.08 events/s. This is attributed to the fact that MAN8 is uncharged, and therefore subjected to electro-osmotic flow mediated capture only.

Blockade versus dwell time scatter plots are shown in Figure 6.5c, showing that most events have Δt very close to the filter bandwidth limit. This is quite a fundamental problem, as high V_{bias} is required for the cap-

ture rate, but at the same time strong V_{bias} lead to fast translocations. A histogram of blockade distribution is finally shown in Figure 6.5d, indeed indicating that most events have low $\Delta I/I_0$ attributed to filter distortions.

6.3. Unfolding of Ribonuclease

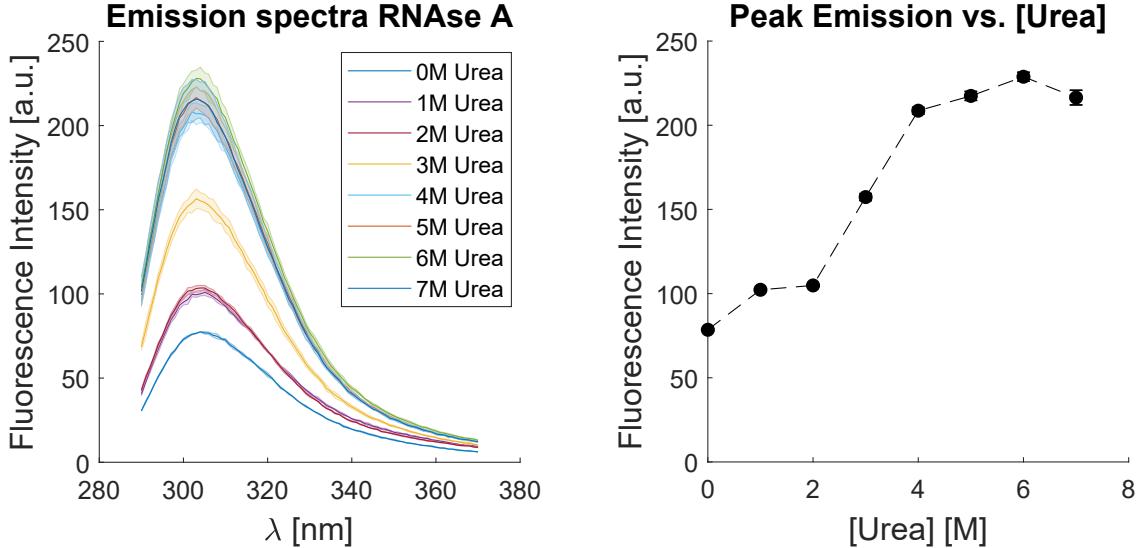


Figure 6.6: Fluorescence spectroscopy measurements of RNase A in urea. **a)** Emission spectra of RNase A in 0-8M Urea, with excitation wavelength of 280nm. **b)** Peak emission over urea concentration. An S-shaped curve can be found, with peak emissions plateauing for $[Urea] > 4M$.

After this brief detour, we return to RNase A and B. We strive to translocate RNase molecules, with the ultimate goal of directly probing the glycan's structure and location within a nanopores. To promote translocation, denaturing conditions were explored. One of the most common compounds used for protein denaturation is urea. It was found to disrupt hydrophobic/hydrophilic interactions between side chains by forming hydrogen bonds with the peptide backbone structure [138]. Ultimately, this leads to protein unfolding at high urea concentrations [68]. As mentioned earlier, RNase A is very stable protein, mainly contributed by its four disulfide bonds [26, 83, 88]. It is therefore essential to break these bonds using reducing agents to obtain an unfolded structure [97]. Tris(2-carboxyethyl)phosphine (TCEP) was employed for this purpose.

The degree of unfolding of RNase, and protein in general, can be monitored using fluorescence spectroscopy. The aromatic residues Tyrosine, Tryptophan and Phenylalanine have inherent fluorescent properties, with characteristic emission spectra upon excitation with a laser. Their fluorescent properties are highly dependent on the direct environment of the aromatic residues, which varies during an unfolding process. Previous studies have shown that RNase A's six tyrosine residues have a peak in emission intensity $\lambda_{em} = 305\text{nm}$ at an excitation of $\lambda_{ex} = 280\text{nm}$ [26, 83, 95, 110]. It was furthermore shown that the fluorescence of Tyr residues at positions 25, 73, 76, 92 97, and 115, are strongly quenched by neighbouring disulphide bounds [83, 89]. Unfolding of the protein then increases the distance to the disulphide bonds, leading to an increase in fluorescence intensity [83, 89].

We therefore measure the emission spectra of RNase A between 290nm-370nm in concentrations of 0 – 8M Urea and 1mM TCEP. The result is summarized in Figure 6.6a. Background emission of a clear Urea buffer was used to correct the signal by subtraction. Peak intensity values at 305nm were then extracted and plotted in Figure 6.6b. We see that the a transition in fluorescence occurs between 2-4M Urea, plateauing for $[Urea] > 4M$, in line with the literature [83, 89].. This suggests that an unfolded intermediate occurs with increasing fluorescence intensity starting from 2M urea. We recognize a 50% fluorescence intensity point at 3M Urea, in line with the literature [88]. Control measurements suggest that adding only TCEP has no significant effect on the fluorescence intensity. It is speculated that the TCEP molecules are unable to react with disulphide bonds, as they are located within the folded structure of the protein [83, 98].

Other common denaturants include Guanidinium Hydrochloride (GdnHCl) [89] and sodium-dodecyl sulphate (SDS) [99, 122]. Similar spectrophotometer measurements were performed for 0-5M GdnHCl and 1mM TCEP, but were inconclusive as to the exact concentration at which a fully unfolded RNase molecule

was achieved. The results have been summarized in Appendix C.0.2. Preliminary measurements in sodium-dodecyl sulphate (SDS) showed that both the lipid bilayer and nanopore were significantly destabilized at a concentration of $\approx 10^{-4}M$. This is an order of magnitude lower than the amount required to unfold RNase A [122], and SDS was therefore deemed inappropriate. Destabilization was found to be due to insertion of SDS into the bilayer, creating membrane instabilities [128].

We therefore propose to incubate RNase in 8M urea and 4mM TCEP, followed by measurement in 4M urea and 1mM TCEP to preserve the unfolded state. Higher concentrations are more likely to also affect FraC's stability and are therefore avoided. It must however be noted that α Haemolysin pores have reportedly been stable at urea concentrations up to 7M [50] giving a rough estimate for the workable range for FraC. The requirement of high $V_{bias} \approx -150mV$ for RNase translocation in tandem with denaturing conditions is expected to greatly limit this range.

6.4. Translocating Ribonuclease in denaturing conditions

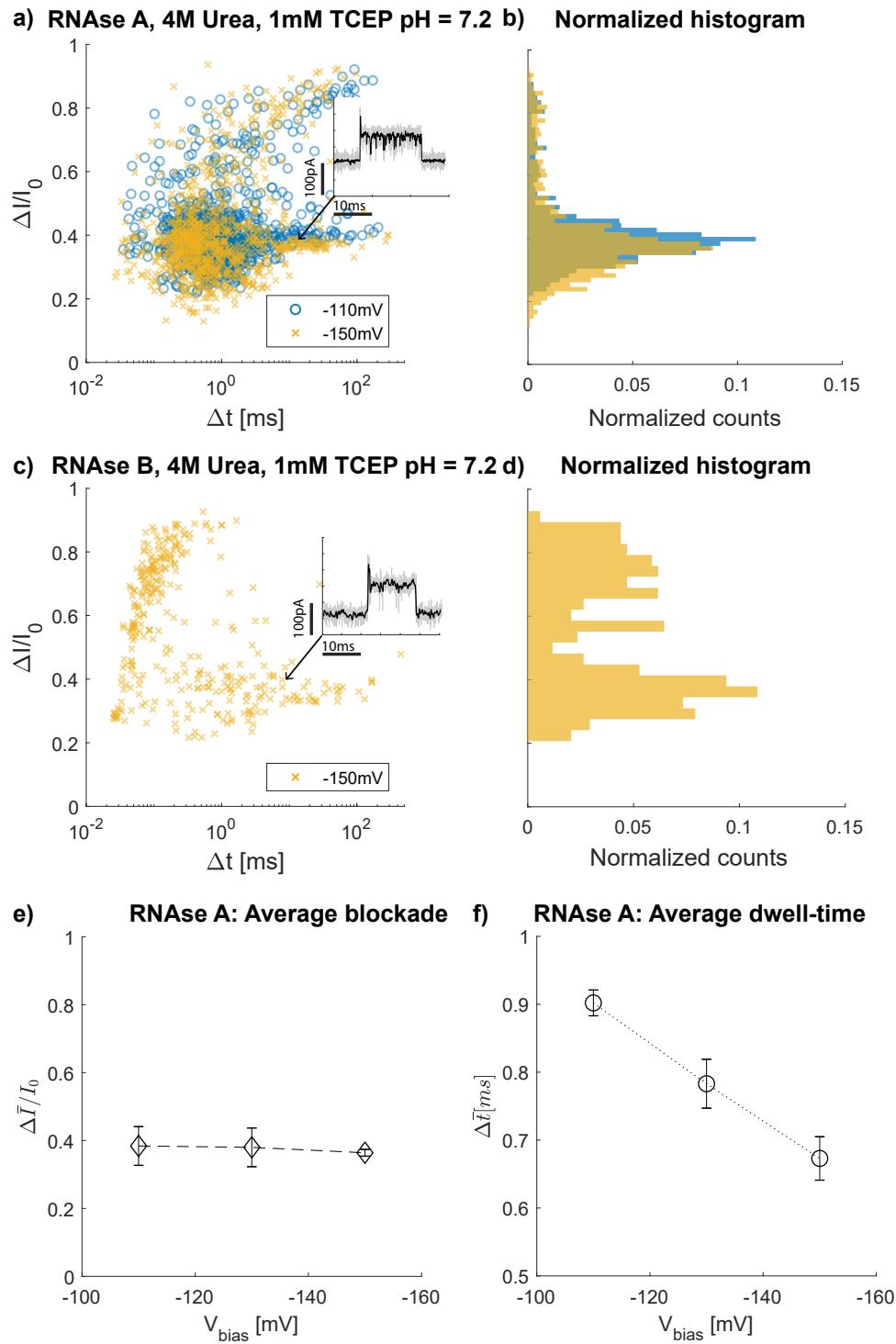


Figure 6.7: Measurement results of RNase A and B in 4M Urea and $pH = 7.2$. **a)** Relative blockade versus dwell time plot for RNase A. Colors and markers indicate different V_{bias} values. A population can be recognized at $\Delta I/I_0 \approx 0.4$, with a large variation in blockade and dwell time. The inset shows an example trace taken from this population. **b)** Normalized histogram of $\Delta I/I_0$ distribution of a). The population at $\Delta I/I_0 \approx 0.4$ can be more readily recognized in this graph. **c)** Measurement results for RNase B for $V_{bias} = -150\text{mV}$. A similarly spread population can be found around $\Delta I/I_0 \approx 0.4$. A second population is now also visible with higher blockade and shorter dwell time. **d)** Normalized histogram of relative blockade distribution from c). The second population is easily resolved in this figure. **e)** Average relative blockade for RNase A, obtained by Gaussian fitting. Errorbars indicate one s.d. of the gaussian. **f)** Average dwell-time for RNase A, obtained by fitting with a single-exponential. Errorbars indicate 95% fitting confidence interval.

As outlined in the previous section, RNase A and B molecules were incubated in 8M urea and 4mM TCEP for atleast one hour, with a buffer solution of 1M NaCl, 10mM TRIS, and 1mM EDTA. Measurements in 4M urea and 1mM TCEP subsequently were performed using the alternating V_{bias} polarity method described in Section 3.3.3. Figure 6.7a shows the blockade versus dwell time scatter plot for RNase A. Colors and markers indicate different V_{bias} values. A large spread in $\Delta I/I_0$ is found, which can be verified the relative blockade histogram in Figure 6.7b. The inset in the Figure 6.7a shows a representative trace for the population around $\Delta I/I_0 \approx 0.4$. A strongly fluctuating current is visible throughout the event, stabilizing when the open pore current is restored. The spread in blockade is therefore attributed to these fluctuating currents.

Next, the experimental results of RNase B measured in the same denaturing conditions is presented in Figure 6.7c. Similar to RNase A, these results also show a spread around $\Delta I/I_0 \approx 0.4$. However, another population with higher blockade and shorter dwell times is recognized. Possibly, these fast translocations of successfully unfolded protein.

We yet again fit the relative blockade and dwell-time distributions with a gaussian and single-exponential respectively. The fitting results of RNase A are shown in Figure 6.7e and f. From the first Figure 6.7e, it is observed that the average blockade level remains practically unchanged within the measured voltage range. Figure 6.7f shows a linear decrease in average dwell-time, giving the first indication that RNase A is translocating the FraC nanopore. Similar analysis must be done for RNase B. However, due time constraints, these measurements have yet to be performed.

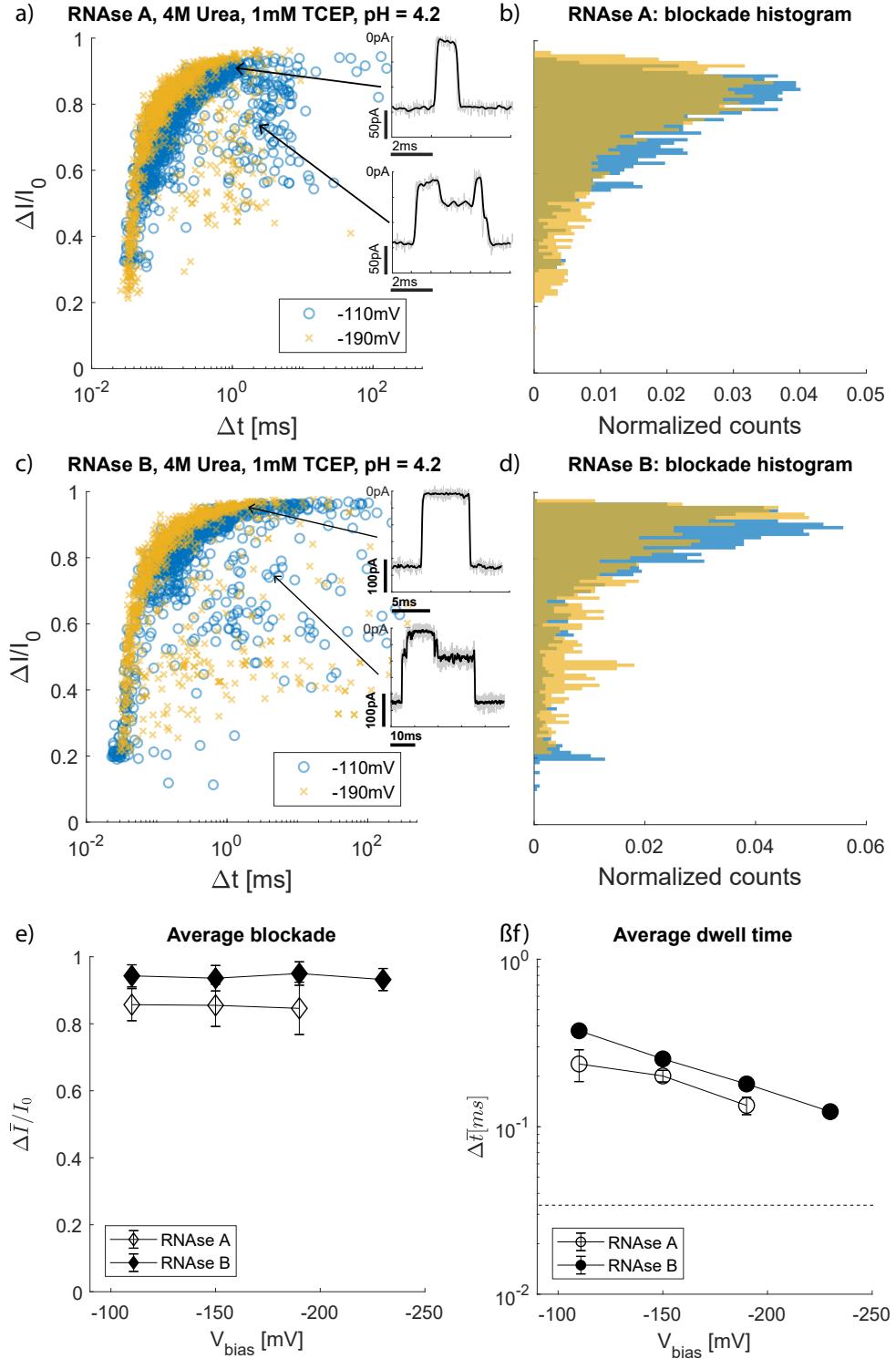


Figure 6.8: Measurements results for Ribonuclease A and B in 4M Urea and $pH = 4.2$. **a)** Relative blockade versus dwell time scatter plot for Rnase A. Colours and markers indicate different V_{bias} values. The inset shows representative traces for the indicated regions. **b)** Normalized blockade distribution corresponding to a) **c)** Relative blockade versus dwell time scatter plot for Rnase B. The inset shows representative traces for the indicated regions. Representative traces for events in the cloud below the measurement. The large spread in $\Delta I / I_0$ is then subsequently due to the irregularity of the fluctuations. **d)** Normalized blockade distribution corresponding to c) **e)** Gaussian fitting results for RNase A (hollow) and B (filled) to obtain average blockade values. Errorbars indicate one s.d. of the gaussian. **f)** Single-exponential fitting results for RNase A (hollow) and B (filled). Errorbars indicate 95% fitting confidence interval.

Instead, it was attempted to further destabilize the protein with a combination of acidic conditions and 4M urea [89], which at the same time facilitates a more uniform positive charge on the RNase molecules. For these measurements, RNase A and B molecules were also incubated in 8M urea, 4mM TCEP and $pH = 4.2$ for atleast one hour in a buffer solution of 1M NaCl, 10mM sodium citrate, and 1mM EDTA. Measurements were subsequently performed at 4M urea, 1mM TCEP, in the same buffer solution. The results are collected in Figure 6.8.

Figure 6.8a shows the relative blockade versus dwell time scatter plot for RNase A. Colours and markers indicate V_{bias} values. The population is distinctively different from its analogous measurement at neutral $pH = 7.2$. Most notably, there seems to be only one defined blockade level near $\Delta I/I_0 \approx 0.9$. The top inset shows a representative trace of these events. A consistent current blockade is observed for the entire the event. We also note that most events are affected by filter distortions, as concluded from the characteristic downward curved slope in the population. From this, we can also recognize that average Δt values are lower compared to the neutral $pH = 7.2$ measurement. A 'cloud' of datapoints is observed, which is deviating from the $\Delta I/I_0 \approx 0.9$ population. A representative trace is shown in the bottom inset. Events within this cloud show current fluctuations, which decrease their reported $\Delta I/I_0$ value in the scatterplot, account for the deviation.

Figure 6.8b shows the relative blockade histogram of RNase A. It can be seen that at higher voltages, a larger portion of events is situated at low $\Delta I/I_0$ values. This is attributed to the increased effect of filter distortions on shorter events.

Experimental results for RNase B in acidic denaturing conditions are subsequently shown in Figures 6.8c and d. The population in Figure 6.8c is on first glance very similar to one found for RNase A. The current traces also look qualitatively similar, as visible in the inset of Figure 6.8c. Figure 6.8d shows the blockade histogram of RNase B. We can once again observe that for higher V_{bias} values, a larger portion of events has low $\Delta I/I_0$ values.

We further analyse this behaviour by comparing fitting results of RNase A and B. To that end, Figure 6.8e shows the average blockade obtained from a Gaussian fit to the blockade distribution. The errorbars indicate one s.d. of the Gaussian. It can be seen that on average, the blockade level stays very approximately constant for both RNase A (hollow) and B (filled) over the measured V_{bias} range. Note that the blockade of RNase A is consistently lower than RNase B. Figure 6.8f shows the average dwell time, obtained by fitting a single exponential to the dwell time distribution of RNase A (hollow) and B (filled) respectively. Errorbars indicate 95% fitting confidence intervals. Noting the logarithmic y-axis, we observe first and foremost that both species have an exponential decrease in average Δt with V_{bias} , indicating translocation of both RNase A and B at 4M urea and $pH = 4.2$. We also see that Rnase A translocates faster on on average compared to RNase B over the measured V_{bias} range. With this, we partly attribute RNase A's lower average $\Delta I/I_0$ value to increased effect of filter distortions.

With substantial evidence for the translocation of RNase A and B, we revisit our initial goal of probing the glycan directly. From the insets of Figure 6.8a and c, we see that current modulation within a single event occurs for both RNase A and B. This suggests that these modulations are the result of some common factor between RNase A and B. Previous studies have suggested denatured RNase A retains more than 10% of its secondary structures [111], which could hypothetically account for the fluctuations. We also speculate that at 4M urea and $pH = 4.2$, the FraC nanopore is comparatively less stable. As a result, measurements at -190mV can potentially cause the pore structure to close transiently, also known as gating. Theoretical models based on statistical mechanics suggest that these gating events can be recognized by a power-law $P(\Delta t) = a * x^{-b}$ with $b = 3/2$, as opposed to a single exponential distribution [76, 85]. By naively fitting a power-law on the Δt distribution of RNase A, we find $b = 1.215 \pm 0.02$ and $b = 1.521 \pm 0.01$ for $V_{bias} = -110\text{mV}, -190\text{mV}$ respectively. No literature was found on the change in exponent b with V_{bias} . Repeat experiments are necessary to confirm the observed pattern and to explore whether the theoretical approaches are applicable for this set-up.

7

Discussion and outlook

In this chapter we critically discuss the validity and value of our measurement results in the broader context of available background literature, and give future outlook prospectives on this work.

Firstly, we have characterized a synthetic PepD platform, and have shown that it facilitates highly sensitive measurements in FrAC. To that end, we have shown that PO_3^{-2} variants and O-GlcNAc variants of PepD could be distinguished from the unmodified variant. This is, to the authors knowledge, the first reported successful detection of a sugar subunit modification using nanopores. Furthermore, the PepD also facilitated the first reported nanopore-based differentiation between phosphorylated and O-GlcNAc glycosylated peptide variants amongst themselves. We proposed possible interaction mechanisms to explain the effect of O-GlcNAc and PO_3^{-2} with the pore. Shorter dwell times were observed for the O-GlcNAc variant. This was thought to be due to the entropic penalty associated with confining the sugar moiety in the pore constriction. This could be tested by attaching incrementally larger glycans to the peptide. Larger glycans would hypothetically have a larger entropic penalty, leading to even shorter translocation times. Most likely, a transition to barrier limited capture will occur, as the glycan will no longer easily fit in the pore constriction. In contrast, phosphorylated variants were observed to have longer dwell times. This was hypothesized to be due to the Coulombic interaction of PO_3^{-2} moiety with negatively charged residues in the pore constriction. Preliminary measurements can be performed at acidic conditions to decrease the protonation of both PO_3^{-2} moiety and the pore constriction.

Next, peculiar behaviour was observed for the PO_3^{-2} PepD variant at 0.8M NaCl. In contrast with the hypothesis, current blockades attributed to PO_3^{-2} decreased. The underlying mechanism was speculated to be in increased effect of concentration polarization, due to stronger overlap of λ_D regions [3, 77]. More salt concentrations must be explored to further characterize this phenomena. Computational models could also be considered, where a the charge on the phosphate group can be artificially varied.

At this point we must note that the concept of using an electrophoretic guide in combination with a stalling mechanism was reported previously. Rosen et al. [104] crafted a construct with a polynucleotide guide, and used folded regions of the target protein to transiently stall it in the nanopore. With this construct, they achieved great sensitivity to phosphorylation, as well as the degree of phosphorylation. They also show that their construct too is highly location dependent, with the sensitivity quickly decreasing away from the sensing region. Asandei et al. [10] have used a conceptually identical peptide-based platform to PepD, with stretches positive and negative amino acids mediating capture and stalling. In their work, they are able to harness the high degree of sensitivity to distinguish between Tyrosine and Alanine residues. Together with the present work, this suggests that confining the site of interest within the pore constriction for extended times is necessary to achieve the reported sensitivity reported sensitivity.

Despite the advantages that the PepD platform offers, it remains a highly artificial construct. As a first step expand PepD's generality, the Arg-Adaptor was designed with the ultimate goal of transferring PepD-like qualities to any target peptide/protein. In this work, it was shown that the Arg-Adaptor was successfully ligated to the target MUC5AC peptide using NHS-ester chemistry, resulting in significant enhancement of target capture rate in nanopores. In contrast with the polynucleotide stretch reported by [15], the Arg-Adaptor has not shown any orientation-dependent blocking behaviour, which simplifies data interpretation. The next step is to attempt PTM detection using the Arg-Adaptor region. Glycosylated variants of MUC5AC with identical sequences can be analysed to verify the Arg-Adaptors capability of facilitating its target region.

Even though NHS-chemistry in theory allows for N-terminal amine specific modification [44], the Arg-Adaptors specificity over lysine sidechain amines was not tested within this work. The reaction was also observed to have low yields, which is problematic for low-copy number targets of *in-vivo* samples. Literature suggest that 30% yield with near 100% N-terminal specificity can be obtained at $pH = 5$ [79]. This suggests that the problem can be partly alleviated by optimizing reaction conditions. Another, more fundamental problem is N-terminal acetylation *in-vivo*. Roughly 80% of the Eukaryotic protein have acetylated N-terminal amines, preventing the NHS-ligation. This poses more limitations on the use of the Arg-Adaptor. With the use of deacetylases, this problem can be partly alleviated.

From the above, it is concluded that different ligation strategies must be explored, focussing on increasing either C- or N-terminal specificity and reaction yields. Recently, 2-pyridinecarboxyaldehydes (2PCA) have shown great promise for N-terminal specificity [72]. However, it still remains challenging to add 2PCA functionality to an peptide such as the Arg-Adaptor. Coincidentally, a highly C-terminus specific labelling reaction was also reported recently, using the difference in oxidation potentials of side-chain aklyl carboxyls versus C-terminal carboxyls [16]. Likewise, it will remain challenging to create a functional adduct on the Arg-Adaptor. Development of these methods highlights the need for termini specific labelling in biophysics, which should be carefully followed to enable widely applicable nanopore translocation techniques.

The last experimental part of this work was performed on RNase A and B. A combination of different pH -values and denaturant conditions were tested, with the aim of translocating RNase A and B. Ultimately, this was thought to enable the direct probing of the glycan on RNase B, and with it conclusive, single-molecule differentiation of RNase A from RNase B. It was found that RNase A and B could be confidently differentiated at $pH = 4.2$, 1mM NaCl and 1mM TCEP based on their current blocking characteristics, even though no evidence was found for translocation. This is analogous to previous studies by Fahie and Chen [34], who reported the distinction of streptavidin and N-linked glycosylated avidin, purely by transiently keeping the molecules in the vicinity of the pore. Next, substantial evidence was found for the translocation of RNase A and B at 4M urea and $pH = 4.2$, based on exponentially decreasing dwell times with increasing $|V_{bias}|$. Repeat measurements must be done at different voltages to conclude the same at neutral $pH = 7.2$. Other potential control experiments include carboxymethylation of the cysteine residues, to permanently prevent the formation of disulfide bonds. It still remains inconclusive whether the glycan was directly probed, due to current fluctuations occurring for both RNase A and B. Fluctuations due to pore gating could potentially be alleviated by optimizing pH and urea concentration. Control experiments of only the FraC pore in denaturing conditions can serve as a reference for the pores behaviour.

Finally we note that the PepD platform was unable to distinguish the acetylated variant of PepD. This suggests that there are PTMs outside of the limits of what can be detected in a label-free manner. Considering the fact that acetylaton and other structurally similar PTMs play pivotal roles in cellular process [33, 57, 71], it is pivotal to devise a method to detect them. What is more, location specific PTM detection was not explored in this work, despite its relevance *in-vivo* [33, 71].

Inspired by the comment of Rosen et al. [104], we propose to use transiently bound, PTM specific labels to tackle these problems. These labels will then hypothetically halt and stall the protein at every PTM site. Electric forces then facilitate the unbinding of the label, which is detected by a current increase. By measuring the time in between labels, an estimation can be made as to the location of the PTM. By designing the labels in such a way that they give characteristic current blockades for each PTM type, allowing it to be identified. The added benefit of the non-covalent nature of the label is that no biochemical sample preparation is required. The transient unbinding of double-stranded DNA for instance has been shown to be detectable in previous studies [135]. Possible labels include PTM-specific antibodies, or enzymes such as deacetylases and phosphatases.

8

Conclusion

In conclusion, this work has contributed experimental work to the field of PTM detection with biological nanopores. The effectiveness of the artificial peptide model PepD within the context of PTM detection was tested. It has been found that the PepD platform provided enough sensitivity to distinguish O-GlcNAc glycosylated and phosphorylated variants of a peptide in a nanopore, which has not been reported previously. We have underlined the decrease in Δt for the glycosylated variant, and the peculiar decrease of $\Delta I/I_0$ with λ_D for the phosphorylated variant. It was also found that the system was unable to differentiate acetylated lysines from unmodified ones. We have proposed conceptually transient binding schemes to overcome these limitations for future research, to further develop nanopore-based PTM detection.

Secondly, we have reported on the Arg-Adaptor region, positively charged amino acids designed to facilitate transport of (uncharged) target peptides conceptually similar to PepD. We have presented evidence for significantly improved capture rates for adapted MUC5AC, compared to unmodified MUC5AC. We have discussed potential termini-specific chemistries to increase the generality of these approaches.

Finally, we have presented our efforts of translocating and distinguishing between RNase A and B. Suggestive evidence was found for the successful translocation of RNase in 4M Urea, 1mM TCEP at $pH = 4.2$. It was also found that RNase A and B behave very differently in both Δt and $\Delta I/I_0$ at $pH = 4.2$ and 0M urea, enabling differentiation between them. Nevertheless, it remains inconclusive whether the glycan itself was actually probe in this work, due to as of yet unclarified current fluctuations. We have presented possible control experiments to ensure the denaturing, unfolding and translocation of RNase in the form of cysteine labelling.

From this work, we conclude that nanopore-based PTM detection still needs formidable improvements in its general applicability and in its capability to detect the location of the PTM'ed residue. However the field is quickly evolving and shows promise in providing an alternative to MS in the form of a facile and cost-efficient single-molecule detection method in the future.

9

Acknowledgements

Firstly, I would like to thank my supervisors Prof. Dekker, Prof. Joo and Laura for giving me the opportunity to work on this project, and creating a positive and open learning environment. I recognize and greatly appreciate your willingness to support and have gained much from our insightful discussions.

Next I would like to thank Prof. Eelkema for accepting the committee member position and taking the time to review my results.

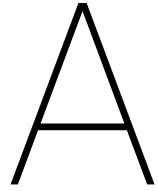
I would like to express an extra thank you to my daily supervisor Laura, for always "being around". Whether you were actually nearby, or all the way in Colombia, I never felt alone in my battle with protein, nanopores, and PTMs. As far as I am concerned, you were the best supervisor a MSc. student could hope for.

Furthermore, I would like to thank the JooC labmembers for showing me the social aspect of academic life in the best way possible. I have thoroughly enjoyed interacting with every one of you, from the steamy sauna in Malaga to the day I graduate and beyond. Specifically, I would like to thank Isa for sharing the happy and sad days, Mike for his violence, speed and momentum, Luuk for always losing at Fussball (and his friendly banter), Thijs for saving me from Luuk during Fussball and his native gel, Ivo for teaching me how to swim like a dolphin, Viktorija for being a great story-teller, Margreet for sharing her passion for music, Sung Chul for teaching me what Korean style means and Sabina for the puns.

To my officemates, I thank you your company throughout the day, the many shared laughs, KitKats and sporadic Bueno's. I especially thank them for putting up with my impromptus singing, and I sincerely hope that the hearing loss is limited. If not the case, you can now enjoy basking in the silence of my absence.

I also want to thank Richard and Sumit for their company during the long weekend, and our in-depth discussions on life. I would also like to thank Sonja for the helpful discussions.

I am sure that I forgot to mention some people, so as a closing word I would like to thank everyone for making my MSc. Thesis a project fun.



Translayzer Changes

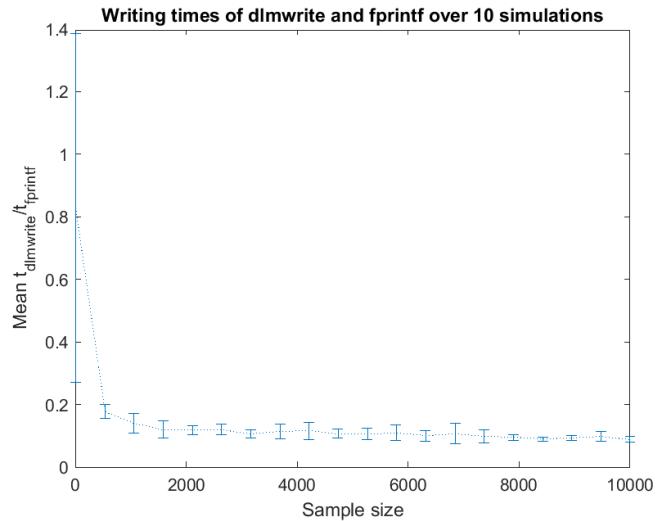


Figure A.1: Ratio of mean writing time using `dlmwrite()` and `fprintf()` for increasing filesizes. A dataset was randomly generated containing increasing amounts of numbers, and written to the hard drive using both functions. The average over 10 simulation was then taken and plotted here. It can be observed that for all but the smallest filesizes, `fprintf()` is almost 10x faster than `dlmwrite()`.

In this section, we detail the changes made to improve the Transalyzer code. Firstly, we decreased the time spent on writing files to disk by a factor of ten, decreasing the overall analysis time by a factor of two. This was achieved by substituting the `dlmwrite()` by the `fprintf()`. To verify this, a random set of numbers was generated and written to disk using both functions. This was done for a multiple set sizes, and repeated 10x per size. The ratio of average writing time using `dlmwrite()` and `fprintf()` is shown in Figure A.1. It can be seen that aside from very small filesizes, `fprintf()` is almost 10x faster.

Next, we implemented changes to the integration method used for determining the relative blockade levels $\Delta I / I_0$. It was found that the integration bounds were determined by obtaining local minima in the event trace. However, local minima below the threshold were also detected, which often arose due to noise spikes. As a result, the integration often overestimated the area under the current. This in turn led to faulty determination of the Full-Width at Half Maximum for dwell-time computation. The effect on the resulting scatterplot of $\Delta I / I_0$ versus Δt is shown in A.2a. A second curved 'tail' is visible, due to the incorrect Δt computation. The problem was fixed by implementing another if-statement, preventing local minima below the threshold to be included in the integration.

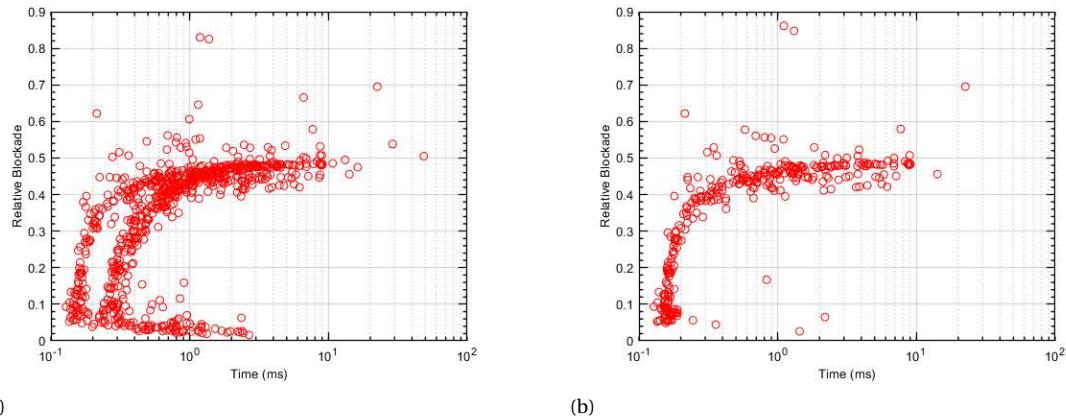


Figure A.2: 2 Figures side by side

```

correct_ind = find(lmval < detectionlevel);
if isempty(correct_ind)
    disp('No_local_minima_found_in_event_outside_of_detection_level')
else
    lmval = lmval(correct_ind);
    minindex = minindex(correct_ind);
end

```

This resulted in Figure A.2b, which shows that the second 'tail' is now gone, effectively decreasing the minimum Δt one can confidently obtain using the software.

B

Supplementary experiments Peptide D

In this chapter experiments performed with relevance to this work.

B.1. Repeat of Glycosylation, Phosphorylation and unmodified mixture

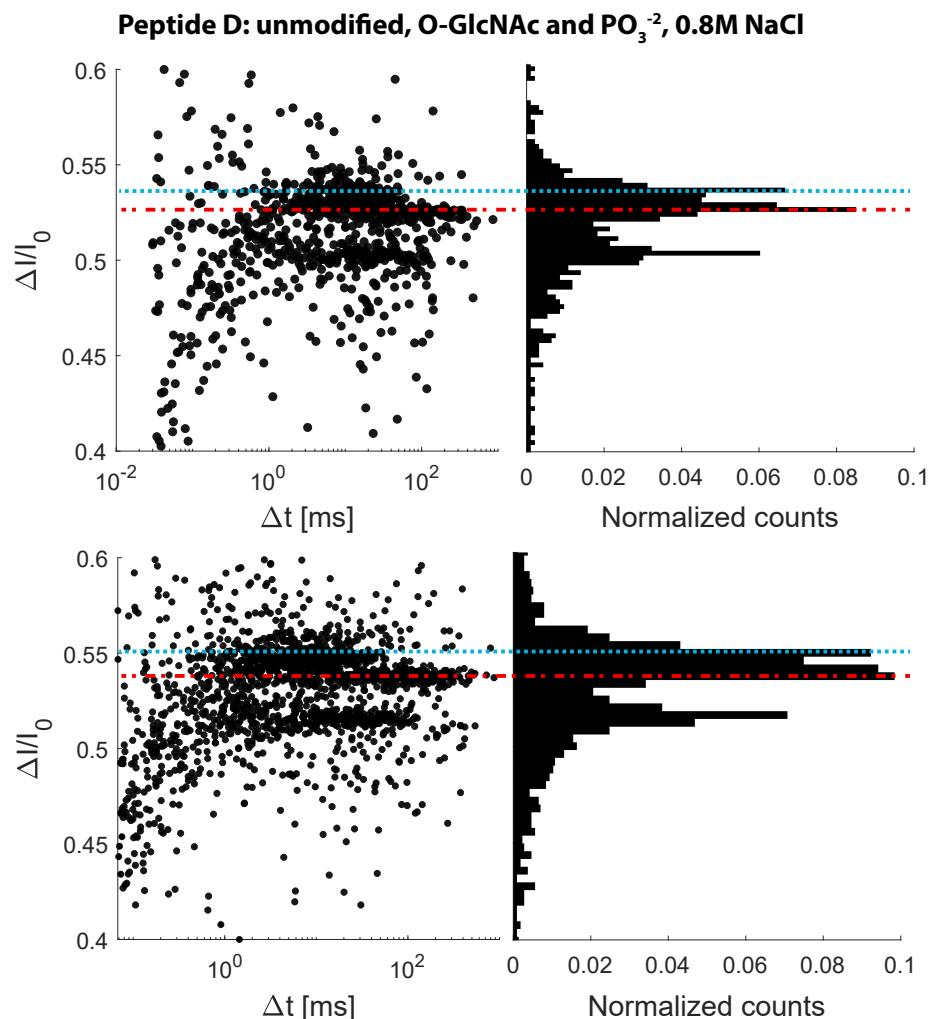


Figure B.1: Repeat experiments for a mixture of unmodified, O-GlcNAc and PO_3^{2-} PepD variants measured at 0.8M NaCl. In both experiments, 3 clear levels are visible, with the blue dashed and red dashed-dotted lines indicating O-GlcNAc and PO_3^{2-} blockades respectively.

2 Repeat measurements were performed for the mixture of unmodified, O-GlcNAc and PO_3^{-2} PepD variants at 0.8M NaCl. Similar to the results shown in Figure 4.4, three $\Delta I/I_0$ levels are visible. Again, the blue dashed and red dashed-dotted line indicate O-GlcNAc and PO_3^{-2} PepD variants respectively.

B.2. High-low salt measurements PepD

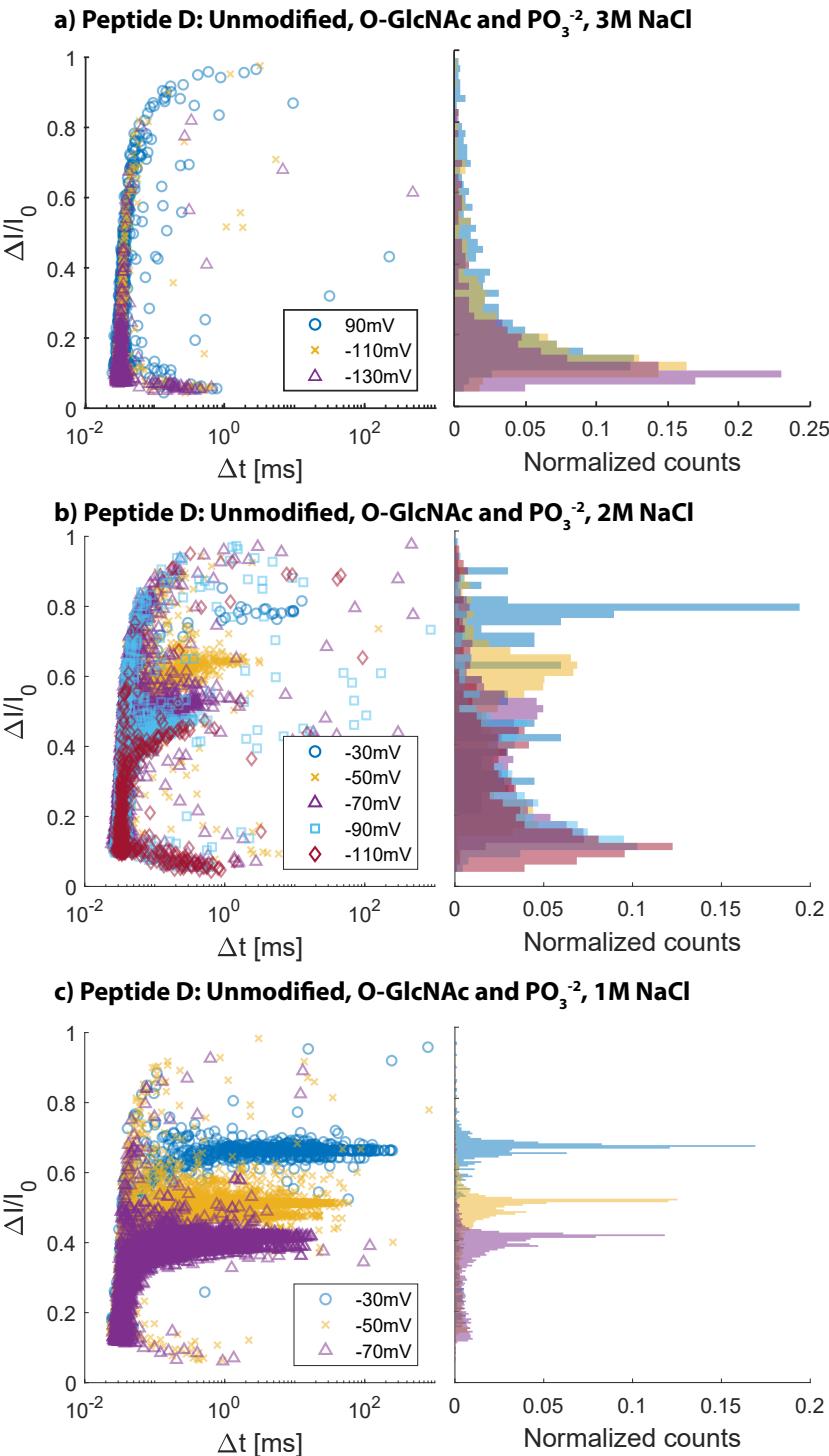


Figure B.2: Mixture of unmodified, O-GlcNAc and PO_3^{-2} PepD at high NaCl concentrations. **a)** Measurement at 3M NaCl. Colors and markers represent different V_{bias} values. Note that for the measured voltages, the events are all close to the filter limit. Higher voltages lead to shorter dwell times **b)** Measurement at 2M NaCl. Measurements show larger Δt values compared to the 3M NaCl case, and begin to have a recognizable $\Delta I/I_0$ level. Higher voltages again lead to shorter Δt and lower $\Delta I/I_0$. Measurement at 1M NaCl, clearly defined $\Delta I/I_0$ levels, with Δt in the order of ms. Again a decrease in Δt and $\Delta I/I_0$ is found with $|V_{bias}|$. Also note that the PTM'ed variants are distinguishable from the unmodified PepD on the basis of their $\Delta I/I_0$ levels.

From the main work regarding PTM detection on PepD, it was concluded that O-GlcNAc and PO_3^{2-} variants of PepD can be distinguished from each other. Substantial evidence was also presented for the decrease of PO_3^{2-} 's contribution to $\Delta I/I_0$ at lower NaCl concentrations. In this measurement series, high salt concentrations were also probed, with the aim of increasing $\Delta I/I_0$ of the PO_3^{2-} variant. The results were summarized in Figure B.2.

Salt concentrations of 3M, 2M and 1M NaCl have been measured and plotted in Figures B.2a, b and c respectively. Overall, we observe that the average Δt decreases with higher NaCl concentration. From the theory described in Section 2.4.3, we also know that the EOF strength also decreases with increasing NaCl concentration. This results in a seemingly contradictory situation, where less EOF generates faster translocation. Another mechanism must therefore be at play to explain the observed behaviour.

We therefore speculate that the increased ionic strength provides stronger shielding of the negative charges in the constriction, which leads to a weaker interaction with the negatively charged Glu-residues on the terminal end of PepD. As a result, the Peptide is stalled less efficiently, leading to shorter Δt on average.

C

Ribonuclease A and B characterization

C.0.1. Mass spectra

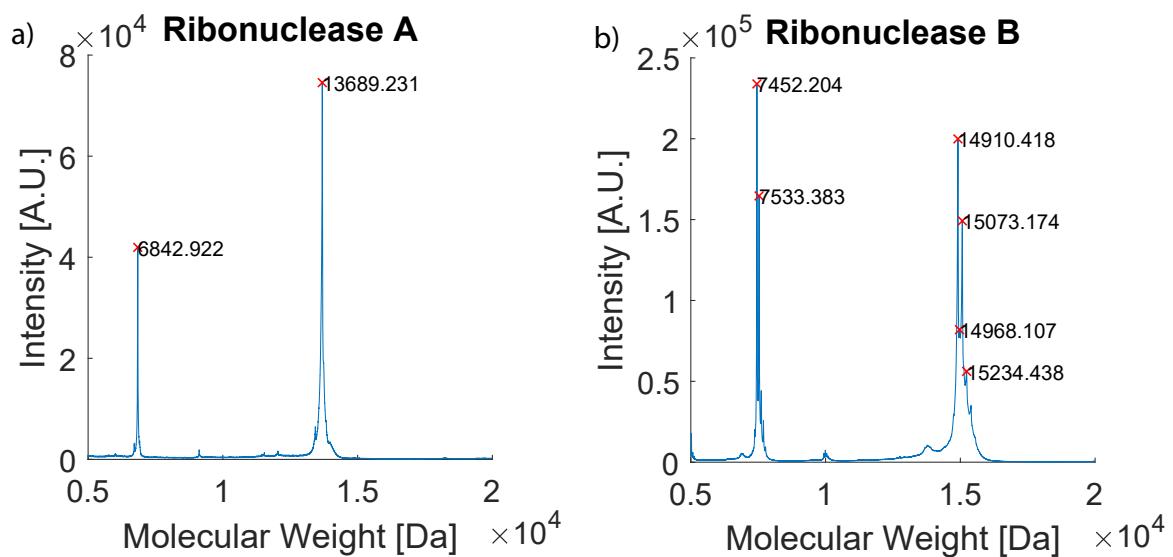


Figure C.1: Mass spectra of Ribonuclease A and B, with expected weights of 13.7kDa and 15kDa respectively.

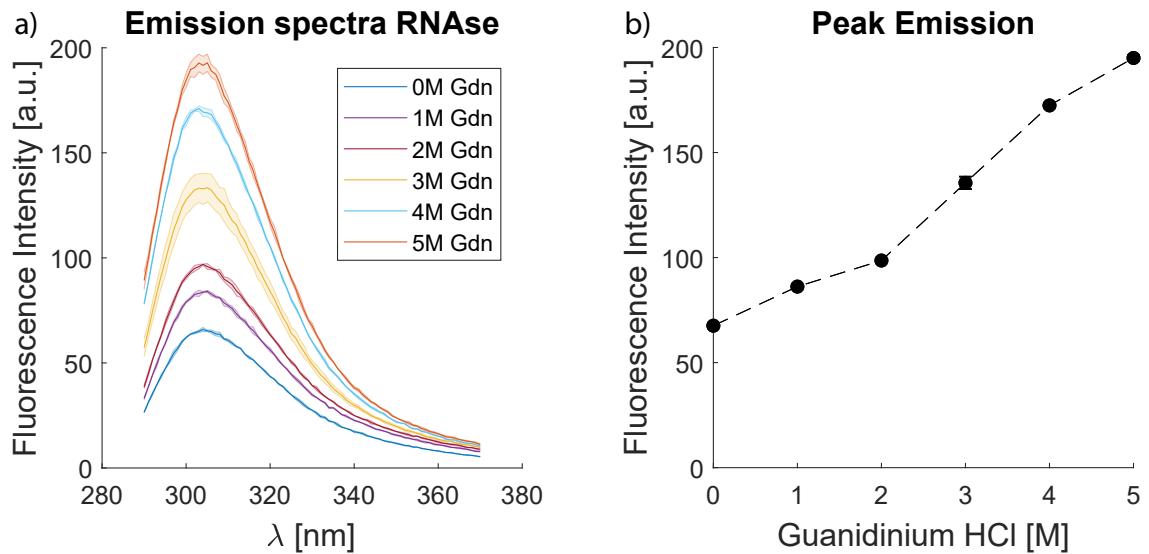
C.0.2. Unfolding Ribonuclease A in Guanidinium HCl

Figure C.2: Results of unfolding RNase A in Guanidinium HCl. a) Unfolding spectra of RNase A for different concentrations of Guanidinium HCl. b) Peak emission for different concentration Guanidinium HCl.

Bibliography

- [1] A. W. Adamson. Physical Chemistry of Surfaces. *J. Electrochem. Soc.*, 124(5):192C, 1977. ISSN 00134651. doi: 10.1149/1.2133374. URL <http://jes.ecsdl.org/cgi/doi/10.1149/1.2133374>.
- [2] Ruedi Aebersold, Jeffrey N. Agar, I. Jonathan Amster, Mark S. Baker, Carolyn R. Bertozzi, Emily S. Boja, Catherine E. Costello, Benjamin F. Cravatt, Catherine Fenselau, Benjamin A. Garcia, Ying Ge, Jeremy Gunawardena, Ronald C. Hendrickson, Paul J. Hergenrother, Christian G. Huber, Alexander R. Ivanov, Ole N. Jensen, Michael C. Jewett, Neil L. Kelleher, Laura L. Kiessling, Nevan J. Krogan, Martin R. Larsen, Joseph A. Loo, Rachel R. Ogorzalek Loo, Emma Lundberg, Michael J. Maccoss, Parag Mallick, Vamsi K. Mootha, Milan Mrksich, Tom W. Muir, Steven M. Patrie, James J. Pesavento, Sharon J. Pitteri, Henry Rodriguez, Alan Saghatelian, Wendy Sandoval, Hartmut Schlüter, Salvatore Sechi, Sarah A. Slavoff, Lloyd M. Smith, Michael P. Snyder, Paul M. Thomas, Mathias Uhlén, Jennifer E. Van Eyk, Marc Vidal, David R. Walt, Forest M. White, Evan R. Williams, Therese Wohlschlager, Vicki H. Wysocki, Nathan A. Yates, Nicolas L. Young, and Bing Zhang. How many human proteoforms are there? *Nat. Chem. Biol.*, 14(3):206–214, 2018. ISSN 15524469. doi: 10.1038/nchembio.2576. URL <http://dx.doi.org/10.1038/nchembio.2576>.
- [3] Y Ai, M K Zhang, S W Joo, M A Cheney, and S Z Qian. Effects of Electroosmotic Flow on Ionic Current Rectification in Conical Nanopores. *J. Phys. Chem. C*, 114(9):3883–3890, 2010. ISSN 1932-7447. doi: Doi10.1021/Jp911773m. URL <ccc:000275045600026>.
- [4] Aleksij Aksimentiev and Klaus Schulten. Imaging α -Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability, and the Electrostatic Potential Map. *Biophys. J.*, 88(6):3745–3761, jun 2005. ISSN 00063495. doi: 10.1529/biophysj.104.058727. URL <http://dx.doi.org/10.1529/biophysj.104.058727> <http://linkinghub.elsevier.com/retrieve/pii/S0006349505734299>.
- [5] Bruce Alberts, Dennis Bray, Julian Lewis, Alexander D. Johnson, Roberts Keith, and Karen Hopkins. *Essential Cell Biology*. Norton & Co, 3 edition, 2009. ISBN 9780815344551. URL <http://ftp.dsma.dp.ua/211/ENG/Other/Alberts-EssentialCellBiology3rded.pdf>.
- [6] T. Albrecht, Thomas Gibb, and P. Nuttall. *Ion Transport in Nanopores*. Elsevier Inc., first edit edition, 2013. ISBN 9781437734737. doi: 10.1016/B978-1-4377-3473-7.00001-7. URL <http://dx.doi.org/10.1016/B978-1-4377-3473-7.00001-7>.
- [7] Tim Albrecht and Joshua B. Edel. Introduction. *Eng. Nanopores Bioanal. Appl. A Vol. Micro Nano Technol.*, 2013. doi: 10.1016/B978-1-4377-3473-7.00011-X.
- [8] Ibraheem Ali, Ryan J. Conrad, Eric Verdin, and Melanie Ott. Lysine Acetylation Goes Global: From Epigenetics to Metabolism and Therapeutics. *Chem. Rev.*, 118(3):1216–1252, feb 2018. ISSN 0009-2665. doi: 10.1021/acs.chemrev.7b00181. URL <http://pubs.acs.org/doi/10.1021/acs.chemrev.7b00181>.
- [9] Pavel Yu Apel, Irina V. Blonskaya, Oleg L. Orelovitch, Patricio Ramirez, and Bozena A. Sartowska. Effect of nanopore geometry on ion current rectification. *Nanotechnology*, 22(17), 2011. ISSN 09574484. doi: 10.1088/0957-4484/22/17/175302.
- [10] Alina Asandei, Mauro Chinappi, Jong Kook Lee, Chang Ho Seo, Loredana Mereuta, Yoonkyung Park, and Tudor Luchian. Placement of oppositely charged aminoacids at a polypeptide termini determines the voltage-controlled braking of polymer transport through nanometer-scale pores. *Sci. Rep.*, 5 (November 2014):1–13, 2015. ISSN 20452322. doi: 10.1038/srep10419. URL <http://dx.doi.org/10.1038/srep10419>.
- [11] Hagan Bayley. Nanopore sequencing: From imagination to reality. *Clin. Chem.*, 61(1):25–31, 2015. ISSN 15308561. doi: 10.1373/clinchem.2014.223016.

- [12] Augusto Bellomio, Koldo Morante, Ariana Barlič, Ion Gutiérrez-Aguirre, Ana Rosa Viguera, and Juan Manuel González-Mañas. Purification, cloning and characterization of fragaceatoxin C, a novel actinoporin from the sea anemone *Actinia fragacea*. *Toxicon*, 54(6):869–880, 2009. ISSN 00410101. doi: 10.1016/j.toxicon.2009.06.022.
- [13] B S Berlett and E R Stadtman. Protein oxidation in aging, disease, and oxidative stress. *J. Biol. Chem.*, 272(33):20313, 1997. ISSN 0021-9258. doi: 10.1074/jbc.272.33.20313.
- [14] Klaus Biemann. Contributions of mass spectrometry to peptide and protein structure. *Biol. Mass Spectrom.*, 16(1-12):99–111, 1988. ISSN 10969888. doi: 10.1002/bms.1200160119.
- [15] Sudipta Biswas, Weisi Song, Chad Borges, Stuart Lindsay, and Peiming Zhang. Click Addition of a DNA Thread to the N-Termini of Peptides for Their Translocation through Solid-State Nanopores. *ACS Nano*, 9(10):9652–9664, 2015. ISSN 1936086X. doi: 10.1021/acsnano.5b04984.
- [16] Steven Bloom, Chun Liu, Dominik K. Kölmel, Jennifer X. Qiao, Yong Zhang, Michael A. Poss, William R. Ewing, and David W.C. Macmillan. Decarboxylative alkylation for site-selective bioconjugation of native proteins via oxidation potentials. *Nat. Chem.*, 10(2):205–211, 2018. ISSN 17554349. doi: 10.1038/nchem.2888.
- [17] Jerry L Brown and Walden K Roberts. Evidence that approximately eighty per cent of the soluble Proteins from Ehrlich Ascites Cells Are Nalpha-Acetylated. *J. Biol. Chem.*, 251(4):1009–1976, 1976. ISSN 0021-9258.
- [18] Chutikarn Butkinaree, Kyoungsook Park, and Gerald W. Hart. O-linked β -N-acetylglucosamine (O-GlcNAc): Extensive crosstalk with phosphorylation to regulate signaling and transcription in response to nutrients and stress. *Biochim. Biophys. Acta - Gen. Subj.*, 1800(2):96–106, 2010. ISSN 03044165. doi: 10.1016/j.bbagen.2009.07.018. URL <http://dx.doi.org/10.1016/j.bbagen.2009.07.018>.
- [19] J. Cervera, B. Schiedt, and P. Ramírez. A poisson/nernst-planck model for ionic transport through synthetic conical nanopores. *Europhys. Lett.*, 71(1):35–41, 2005. ISSN 02955075. doi: 10.1209/epl/i2005-10054-x.
- [20] Javier Cervera, Birgitta Schiedt, Reinhard Neumann, Salvador Mafá, and Patricio Ramírez. Ionic conduction, rectification, and selectivity in single conical nanopores. *J. Chem. Phys.*, 124(10), 2006. ISSN 00219606. doi: 10.1063/1.2179797.
- [21] Javier Cervera, Patricio Ramírez, José A. Manzanares, and Salvador Mafé. Incorporating ionic size in the transport equations for charged nanopores. *Microfluid. Nanofluidics*, 9(1):41–53, 2010. ISSN 16134982. doi: 10.1007/s10404-009-0518-2.
- [22] C. Chad Harrell, Zuzanna S. Siwy, and Charles R. Martin. Conical nanopore membranes: Controlling the nanopore shape. *Small*, 2(2):194–198, 2006. ISSN 16136810. doi: 10.1002/smll.200500196.
- [23] Joy Chakraborty, Valentina Basso, and Elena Ziviani. Post translational modification of Parkin. *Biol. Direct*, 12(1):6, 2017. ISSN 17456150. doi: 10.1186/s13062-017-0176-3. URL <http://www.ncbi.nlm.nih.gov/pubmed/28222786>. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5319146/>.
- [24] L. Chang and M. Karin. Mammalian MAP kinase signalling cascades. *Nature*, 410(6824):37–40, 2001. ISSN 00280836. doi: 10.1038/35065000.
- [25] Todd J. Cohen, Jing L. Guo, David E. Hurtado, Linda K. Kwong, Ian P. Mills, John Q. Trojanowski, and Virginia M.Y. Lee. The acetylation of tau inhibits its function and promotes pathological tau aggregation. *Nat. Commun.*, 2(1):252–259, 2011. ISSN 20411723. doi: 10.1038/ncomms1255. URL <http://dx.doi.org/10.1038/ncomms1255>.
- [26] Robert W. Cowgill. Fluorescence and the structure of proteins. *Arch. Biochem. Biophys.*, 104(1):84–92, jan 1964. ISSN 00039861. doi: 10.1016/S0003-9861(64)80038-2. URL <http://linkinghub.elsevier.com/retrieve/pii/S0003986164800382>.

- [27] Helen Davies, Graham R Bignell, Charles Cox, Philip Stephens, Sarah Edkins, Sheila Clegg, Jon Teague, Hayley Woffendin, Mathew J Garnett, William Bottomley, Neil Davis, Ed Dicks, Rebecca Ewing, Yvonne Floyd, Kristian Gray, Sarah Hall, Rachel Hawes, Jaime Hughes, Vivian Kosmidou, Andrew Menzies, Catherine Mould, Adrian Parker, Claire Stevens, Stephen Watt, Steven Hooper, Rebecca Wilson, Hiran Jayatilake, Barry A Gusterson, Colin Cooper, Janet Shipley, Darren Hargrave, Katherine Pritchard-Jones, Norman Maitland, Georgia Chenevix-Trench, Gregory J Riggins, Darell D Bigner, Giuseppe Palmieri, Antonio Cossu, Adrienne Flanagan, Andrew Nicholson, Judy W C Ho, Suet Y Leung, Siu T Yuen, Barbara L Weber, Hilliard F Seigler, Timothy L Darrow, Hugh Paterson, Richard Marais, Christopher J Marshall, Richard Wooster, Michael R Stratton, and P Andrew Futreal. 6-Mutations of the BRAF gene in human cancer. *Nature*, 417(6892):949–54, 2002. ISSN 0028-0836. doi: 10.1038/nature00766. URL <http://www.ncbi.nlm.nih.gov/pubmed/12068308>.
- [28] Angus Depledge, Daniel P and Kalanghad Puthankalam, Srinivas and Sadaoka, Tomohiko and Beady, Devin and Mori, Yasuko and Placantonakis, Dimitris and Mohr, Ian and Wilson. Native RNA sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen, 2018. URL <https://www.biorxiv.org/content/early/2018/07/20/373522>.
- [29] Tejal A. Desai, Derek J. Hansford, Lawrence Kulinsky, Amir H. Nashat, Guido Rasi, Jay Tu, Yuchun Wang, Miqin Zhang, and Mauro Ferrari. Nanopore Technology for Biomedical Applications. *Biomed. Microdevices*, 2(1):11–40, 1999. ISSN 13872176. doi: 10.1023/A:1009903215959.
- [30] Bruno Domon and Ruedi Aebersold. Mass spectrometry and protein analysis. *Science (80-.).*, 312(5771):212–217, 2006. ISSN 00368075. doi: 10.1126/science.1124619. URL <http://www.ncbi.nlm.nih.gov/pubmed/16614208>.
- [31] Hester A. Doyle and Mark J. Mamula. Post-translational protein modifications in antigen recognition and autoimmunity. *Trends Immunol.*, 22(8):443–449, 2001. ISSN 14714906. doi: 10.1016/S1471-4906(01)01976-7.
- [32] Adrian Drazic, Line M. Myklebust, Rasmus Ree, and Thomas Arnesen. The world of protein acetylation. *Biochim. Biophys. Acta - Proteins Proteomics*, 1864(10):1372–1401, 2016. ISSN 18781454. doi: 10.1016/j.bbapap.2016.06.007. URL <http://dx.doi.org/10.1016/j.bbapap.2016.06.007>.
- [33] Guangyou Duan and Dirk Walther. The Roles of Post-translational Modifications in the Context of Protein Interaction Networks. *PLoS Comput. Biol.*, 11(2):1–23, 2015. ISSN 15537358. doi: 10.1371/journal.pcbi.1004049.
- [34] Monifa A. Fahie and Min Chen. Electrostatic Interactions between OmpG Nanopore and Analyte Protein Surface Can Distinguish between Glycosylated Isoforms. *J. Phys. Chem. B*, 119(32):10198–10206, 2015. ISSN 15205207. doi: 10.1021/acs.jpcb.5b06435.
- [35] M Faller, Michael Niederweis, and G E Schultz. The Structure of a Mycobacterial Outer Membrane Channel. *Science (80-.).*, 303(February):1189–1192, 2004.
- [36] Matthias Firnkes, Daniel Pedone, Jelena Knezevic, Markus Döblinger, and Ulrich Rant. Electrically facilitated translocations of proteins through silicon nitride nanopores: Conjoint and competitive action of diffusion, electrophoresis, and electroosmosis. *Nano Lett.*, 10(6):2162–2167, 2010. ISSN 15306984. doi: 10.1021/nl100861c.
- [37] Kevin J. Freedman, Syed Raza Haq, Michael R. Fletcher, Joe P. Foley, Per Jemth, Joshua B. Edel, and Min Jun Kim. Nonequilibrium capture rates induce protein accumulation and enhanced adsorption to solid-state nanopores. *ACS Nano*, 8(12):12238–12249, 2014. ISSN 1936086X. doi: 10.1021/nn5062645.
- [38] Christophe Grangeasse, Jarg Stalke, and Ivan Mijakovic. Regulatory potential of post-translational modifications in bacteria. *Front. Microbiol.*, 6(MAY), may 2015. ISSN 1664-302X. doi: 10.3389/fmicb.2015.00500. URL <http://www.frontiersin.org/books/Regulatory%20Potential%20of%20Post-translational%20Modifications%20in%20Bacteria/646http://www.frontiersin.org/Microbial%20Physiology%20and%20Metabolism/10.3389/fmicb.2015.00500/full>.

- [39] Alexander Y. Grosberg and Yitzhak Rabin. DNA capture into a nanopore: Interplay of diffusion and electrohydrodynamics. *J. Chem. Phys.*, 133(16), 2010. ISSN 00219606. doi: 10.1063/1.3495481.
- [40] J. S. Gutkind. Regulation of Mitogen-Activated Protein Kinase Signaling Networks by G Protein-Coupled Receptors. *Sci. Signal.*, 2000(40):re1-re1, 2000. ISSN 1945-0877. doi: 10.1126/stke.2000.40.re1. URL <http://stke.sciencemag.org/cgi/doi/10.1126/stke.2000.40.re1>.
- [41] J. E. Hall. Access resistance of a small circular pore. *J. Gen. Physiol.*, 66(4):531–532, 1975. ISSN 0022-1295. doi: 10.1085/jgp.66.4.531. URL <http://www.jgp.org/cgi/doi/10.1085/jgp.66.4.531>.
- [42] Samir M. Hanash, Sharon J. Pitteri, and Vitor M. Faca. Mining the plasma proteome for cancer biomarkers. *Nature*, 452(7187):571–579, 2008. ISSN 14764687. doi: 10.1038/nature06916.
- [43] Farzin Haque, Jinghong Li, Hai Chen Wu, Xing Jie Liang, and Peixuan Guo. Solid-state and biological nanopore for real-time sensing of single chemical and sequencing of DNA. *Nano Today*, 8(1):56–74, 2013. ISSN 1878044X. doi: 10.1016/j.nantod.2012.12.008.
- [44] Greg T. Hermanson. *Bioconjugate Techniques*. Elsevier Inc., 3rd edition, 2013. ISBN 978-0-12-382239-0. doi: <https://doi.org/10.1016/C2009-0-64240-9>.
- [45] Bertil Hille. Pharmacological Modifications of the Sodium Channels of Frog Nerve. *J. Gen. Physiol.*, 51(2):199–219, 1968. doi: 10.1085/jgp.51.2.199.
- [46] Stefan Howorka, Stephen Cheley, and Hagan Bayley. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nat. Biotechnol.*, 19(July):636, 2001.
- [47] Gang Huang, Kherim Willems, Misha Soskine, Carsten Wloka, and Giovanni Maglia. Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores. *Nat. Commun.*, 8(1):935, dec 2017. ISSN 2041-1723. doi: 10.1038/s41467-017-01006-4. URL <http://dx.doi.org/10.1038/s41467-017-01006-4> <http://www.nature.com/articles/s41467-017-01006-4>.
- [48] Barbara Imperiali and Sarah E. O'Connor. Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Curr. Opin. Chem. Biol.*, 3(6):643–649, 1999. ISSN 13675931. doi: 10.1016/S1367-5931(99)00021-6.
- [49] Polymer Principles In. Protein structure and stability. 1991.
- [50] Deanpen Japrung, Jakob Dogan, Kevin J. Freedman, Achim Nadzeyka, Sven Bauerdrick, Tim Albrecht, Min Jun Kim, Per Jemth, and Joshua B. Edel. Single-molecule studies of intrinsically disordered proteins using solid-state nanopores. *Anal. Chem.*, 85(4):2449–2456, 2013. ISSN 00032700. doi: 10.1021/ac3035025.
- [51] Ole Nørregaard Jensen. Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.*, 8(1):33–41, 2004. ISSN 13675931. doi: 10.1016/j.cbpa.2003.12.009.
- [52] Heidi C. Joao, Ian G. Scragg, and Raymond A. Dwek. Effects of glycosylation on protein conformation and amide proton exchange rates in RNase B. *FEBS Lett.*, 307(3):343–346, 1992. ISSN 00145793. doi: 10.1016/0014-5793(92)80709-P.
- [53] Jung Yeul Jung, Punarvasu Joshi, Leo Petrossian, Trevor J. Thornton, and Jonathan D. Posner. Electromigration current rectification in a cylindrical nanopore due to asymmetric concentration polarization. *Anal. Chem.*, 81(8):3128–3133, 2009. ISSN 00032700. doi: 10.1021/ac900318j.
- [54] Sezgin Kara, Sergii Afonin, Oleg Babii, Anton N. Tkachenko, Igor V. Komarov, and Anne S. Ulrich. Di-phytanoyl lipids as model systems for studying membrane-active peptides. *Biochim. Biophys. Acta - Biomembr.*, 1859(10):1828–1837, 2017. ISSN 18792642. doi: 10.1016/j.bbamem.2017.06.003. URL <http://dx.doi.org/10.1016/j.bbamem.2017.06.003>.
- [55] Buddini Iroshika Karawdeniya, Y. M. Nuwan D. Y. Bandara, Jonathan W. Nichols, Robert B. Chevalier, and Jason R. Dwyer. Surveying silicon nitride nanopores for glycomics and heparin quality assurance. *Nat. Commun.*, 9(1):3278, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05751-y. URL <http://www.nature.com/articles/s41467-018-05751-y>.

- [56] Rohit Karnik, Chuanhua Duan, Kenneth Castelino, Hirofumi Daiguji, and Arun Majumdar. Rectification of ionic current in a nanofluidic diode. *Nano Lett.*, 7(3):547–551, 2007. ISSN 15306984. doi: 10.1021/nl0628060.
- [57] George A. Khoury, Richard C. Baliban, and Christodoulos A. Floudas. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Sci. Rep.*, 1:1–5, 2014. ISSN 20452322. doi: 10.1038/srep00090.
- [58] D G Knorre, N V Kudryashova, and T S Godovikova. Chemical and functional aspects of posttranslational modification of proteins. *Acta Naturae*, 1(3):29–51, 2009. ISSN 2075-8251. URL <http://www.ncbi.nlm.nih.gov/article/fcgi?artid=3347534&tool=pmcentrez&rendertype=abstract>.
- [59] Michala Kolarova, Francisco García-Sierra, Ales Bartos, Jan Ricny, and Daniela Ripova. Structure and pathology of tau protein in Alzheimer disease. *Int. J. Alzheimers. Dis.*, 2012, 2012. ISSN 20900252. doi: 10.1155/2012/731526.
- [60] S. Kornfeld, R., & Kornfeld. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.*, 54(1):631–664, 1985.
- [61] I. D. Kosińska, I. Goychuk, M. Kostur, G. Schmid, and P. Hänggi. Rectification in synthetic conical nanopores: A one-dimensional Poisson-Nernst-Planck model. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 77(3):1–10, 2008. ISSN 15393755. doi: 10.1103/PhysRevE.77.031131.
- [62] Michelle L. Kovarik, Kaimeng Zhou, and Stephen C. Jacobson. Effect of conical nanopore diameter on ion current rectification. *J. Phys. Chem. B*, 113(49):15960–15966, 2009. ISSN 15206106. doi: 10.1021/jp9076189.
- [63] Stefan W Kowalczyk, Alexander Y Grosberg, Yitzhak Rabin, and Cees Dekker. Modeling the conductance and DNA blockade of solid-state nanopores. *Nanotechnology*, 22(31):315101, aug 2011. ISSN 0957-4484. doi: 10.1088/0957-4484/22/31/315101. URL <http://stacks.iop.org/0957-4484/23/i=8/a=088001?key=crossref.07752d07c079b830c64bd4f4f6efad78>
- [64] H.A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940. ISSN 00318914. doi: 10.1016/S0031-8914(40)90098-2. URL <http://linkinghub.elsevier.com/retrieve/pii/S0031891440900982>.
- [65] Besnik Krasniqi and Jeremy S. Lee. RNase a does not translocate the alpha-hemolysin pore. *PLoS One*, 9(2), 2014. ISSN 19326203. doi: 10.1371/journal.pone.0088004.
- [66] Nadanai Laohakunakorn and Ulrich F. Keyser. Electroosmotic flow rectification in conical nanopores. *Nanotechnology*, 26(27), 2015. ISSN 13616528. doi: 10.1088/0957-4484/26/27/275202.
- [67] Hari Leontiadou, Alan E. Mark, and Siewert J. Marrink. Molecular Dynamics Simulations of Hydrophilic Pores in Lipid Bilayers. *Biophys. J.*, 86(4):2156–2164, 2004. ISSN 00063495. doi: 10.1016/S0006-3495(04)74275-7. URL [http://dx.doi.org/10.1016/S0006-3495\(04\)74275-7](http://dx.doi.org/10.1016/S0006-3495(04)74275-7).
- [68] W. K. Lim, J. Rosgen, and S. W. Englander. Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proc. Natl. Acad. Sci.*, 106(8):2595–2600, 2009. ISSN 0027-8424. doi: 10.1073/pnas.0812588106. URL <http://www.pnas.org/cgi/doi/10.1073/pnas.0812588106>.
- [69] Hening Lin and Kate S. Carroll. Introduction: Posttranslational Protein Modification. *Chem. Rev.*, 118(3):887–888, 2018. ISSN 15206890. doi: 10.1021/acs.chemrev.7b00756.
- [70] Harvey F. Lodish. Post-translational modification of proteins. *Enzyme Microb. Technol.*, 3(3):178–188, jul 1981. ISSN 01410229. doi: 10.1016/0141-0229(81)90084-3. URL <http://onlinelibrary.wiley.com/doi/10.1002/9780470123133.ch3/summary>

- [71] Adam P. Lothrop, Matthew P. Torres, and Stephen M. Fuchs. Deciphering post-translational modification codes. *FEBS Lett.*, 587(8):1247–1257, 2013. ISSN 00145793. doi: 10.1016/j.febslet.2013.01.047. URL <http://dx.doi.org/10.1016/j.febslet.2013.01.047>.
- [72] James I. Macdonald, Henrik K. Munch, Troy Moore, and Matthew B. Francis. One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. *Nat. Chem. Biol.*, 11(5):326–331, 2015. ISSN 15524469. doi: 10.1038/nchembio.1792. URL <http://dx.doi.org/10.1038/nchembio.1792>.
- [73] Giovanni Maglia, Andrew J. Heron, David Stoddart, Deanpen Japrung, and Hagan Bayley. *Analysis of Single Nucleic Acid Molecules with Protein Nanopores*, volume 475. Elsevier Inc., 1 edition, 2010. ISBN 9780123814821. doi: 10.1016/S0076-6879(10)75022-9. URL [http://dx.doi.org/10.1016/S0076-6879\(10\)75022-9](http://dx.doi.org/10.1016/S0076-6879(10)75022-9).
- [74] Silvina Matysiak, Alberto Montesi, Matteo Pasquali, Anatoly B. Kolomeisky, and Cecilia Clementi. Dynamics of polymer translocation through nanopores: Theory meets experiment. *Phys. Rev. Lett.*, 96(11):1–4, 2006. ISSN 00319007. doi: 10.1103/PhysRevLett.96.118103.
- [75] Céline Merstorf, Benjamin Cressiot, Manuela Pastoriza-Gallego, Abdelghani Oukhaled, Jean Michel Betton, Loïc Auvray, and Juan Pelta. Wild type, mutant protein unfolding and phase transition detected by single-nanopore recording. *ACS Chem. Biol.*, 7(4):652–658, 2012. ISSN 15548929. doi: 10.1021/cb2004737.
- [76] G. L. Millhauser, E. E. Salpeter, and R. E. Oswald. Diffusion models of ion-channel gating and the origin of power-law distributions from single-channel recording. *Proc. Natl. Acad. Sci.*, 85(5):1503–1507, 1988. ISSN 0027-8424. doi: 10.1073/pnas.85.5.1503. URL <http://www.pnas.org/cgi/doi/10.1073/pnas.85.5.1503>.
- [77] Dmitry Momotenko and Hubert H. Girault. Scan-rate-dependent ion current rectification and rectification inversion in charged conical nanopores. *J. Am. Chem. Soc.*, 133(37):14496–14499, 2011. ISSN 00027863. doi: 10.1021/ja2048368.
- [78] M. Montal and P. Mueller. Formation of Bimolecular Membranes from Lipid Monolayers and a Study of Their Electrical Properties. *Proc. Natl. Acad. Sci.*, 69(12):3561–3566, 1972. ISSN 0027-8424. doi: 10.1073/pnas.69.12.3561. URL <http://www.pnas.org/cgi/doi/10.1073/pnas.69.12.3561>.
- [79] M. Munchbach, M. Quadroni, G. Miotto, and P. James. Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety. *Anal. Chem.*, 72(17):4047–4057, 2000. ISSN 00032700. doi: 10.1021/ac000265w.
- [80] M. Muthukumar. Theory of capture rate in polymer translocation. *J. Chem. Phys.*, 132(19), 2010. ISSN 00219606. doi: 10.1063/1.3429882.
- [81] José Luis Neira and Manuel Rico. Folding studies on ribonuclease A, a model protein. *Fold. Des.*, 2(1):R1–R11, 1997. ISSN 13590278. doi: 10.1016/S1359-0278(97)00001-1.
- [82] Jeff Nivala, Logan Mulroney, Gabriel Li, Jacob Schreiber, and Mark Akeson. Discrimination among protein variants using an unfoldase-coupled nanopore. *ACS Nano*, 8(12):12365–12375, 2014. ISSN 1936086X. doi: 10.1021/nn5049987.
- [83] Melinda Noronha, João C. Lima, Emanuele Paci, Helena Santos, and António L. Maçanita. Tracking local conformational changes of ribonuclease A using picosecond time-resolved fluorescence of the six tyrosine residues. *Biophys. J.*, 92(12):4401–4414, 2007. ISSN 00063495. doi: 10.1529/biophysj.106.093625.
- [84] Yoshiya Oda, Takeshi Nagasu, and Brian T. Chait. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.*, 19(4):379–382, 2001. ISSN 10870156. doi: 10.1038/86783.
- [85] R. E. Oswald, G. L. Millhauser, and A. A. Carter. Diffusion model in ion channel gating. Extension to agonist-activated ion channels. *Biophys. J.*, 59(5):1136–1142, 1991. ISSN 00063495. doi: 10.1016/S0006-3495(91)82328-1. URL [http://dx.doi.org/10.1016/S0006-3495\(91\)82328-1](http://dx.doi.org/10.1016/S0006-3495(91)82328-1).

- [86] Abdelghani Oukhaled, Laurent Bacri, Manuela Pastoriza-Gallego, Jean Michel Betton, and Juan Pelta. Sensing proteins through nanopores: Fundamental to applications. *ACS Chem. Biol.*, 7(12):1935–1949, 2012. ISSN 15548929. doi: 10.1021/cb300449t.
- [87] G. Oukhaled, J. Mathé, A. L. Biance, L. Bacri, J. M. Betton, D. Lairez, J. Pelta, and L. Auvray. Unfolding of proteins and long transient conformations detected by single nanopore recording. *Phys. Rev. Lett.*, 98(15):98–101, 2007. ISSN 00319007. doi: 10.1103/PhysRevLett.98.158101.
- [88] C. N. Pace, G. R. Grimsley, J. A. Thomson, and B. J. Barnett. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.*, 263(24):11820–11825, 1988. ISSN 00219258.
- [89] C. Nick Pace, Douglas V. Laurents, and James A. Thomson. pH Dependence of the Urea and Guanidine Hydrochloride Denaturation of Ribonuclease A and Ribonuclease Tl. *Biochemistry*, 29(10):2564–2572, 1990. ISSN 15204995. doi: 10.1021/bi00462a019.
- [90] Oliver Pagel, Stefan Loroch, Albert Sickmann, and René P. Zahedi. Current strategies and findings in clinically relevant post-translational modification-specific proteomics. *Expert Rev. Proteomics*, 12(3):235–253, 2015. ISSN 17448387. doi: 10.1586/14789450.2015.1042867.
- [91] Daniel Pedone, Matthias Firnkes, and Ulrich Rant. Data analysis of translocation events in nanopore experiments. *Anal. Chem.*, 81(23):9689–9694, 2009. ISSN 00032700. doi: 10.1021/ac901877z.
- [92] S. Pianegonda, M. C. Barbosa, and Y. Levin. Charge reversal of colloidal particles. *Europhys. Lett.*, 71(5):831–837, 2005. ISSN 02955075. doi: 10.1209/epl/i2005-10150-y.
- [93] Calin Plesa and Cees Dekker. Data analysis methods for solid-state nanopores. *Nanotechnology*, 26(8):084003, feb 2015. ISSN 0957-4484. doi: 10.1088/0957-4484/26/8/084003. URL <http://stacks.iop.org/0957-4484/26/i=8/a=084003?key=crossref.210b60b9ca7efad8cf4e9bbcbf3ae04d>.
- [94] T H Plummer and C. H. W. Hirs. On the Structure of Bovine Pancreatic Ribonuclease B. *J. Biol. Chem.*, 239(8):2530–2538, 1964.
- [95] Olaf Pongs. Influences of pH and Substrate Analogs on Ribonuclease T1 Fluorescence. *Biochemistry*, 9(11):2316–2322, 1970. ISSN 15204995. doi: 10.1021/bi00813a015.
- [96] Elena A. Ponomarenko, Ekaterina V. Poverennaya, Ekaterina V. Ilgisonis, Mikhail A. Pyatnitskiy, Arthur T. Kopylov, Victor G. Zgoda, Andrey V. Lisitsa, and Alexander I. Archakov. The Size of the Human Proteome: The Width and Depth. *Int. J. Anal. Chem.*, 2016, 2016. ISSN 16878779. doi: 10.1155/2016/7436849.
- [97] S Raina. Making and Breaking. *Annu. Rev. Microbiol.*, 51:179–202, 1997.
- [98] Ronald T. Raines. Ribonuclease A. *Chem. Rev.*, 98(3):1045–1066, 1998. ISSN 0009-2665. doi: 10.1021/cr960427h. URL <http://pubs.acs.org/doi/abs/10.1021/cr960427h>.
- [99] Laura Restrepo-Pérez, Shalini John, Aleksei Aksimentiev, Chirlmin Joo, and Cees Dekker. SDS-assisted protein transport through solid-state nanopores. *Nanoscale*, 9(32):11685–11693, 2017. ISSN 20403372. doi: 10.1039/c7nr02450a.
- [100] Nader Rifai, Michael A. Gillette, and Steven A. Carr. Protein biomarker discovery and validation: The long and uncertain path to clinical utility. *Nat. Biotechnol.*, 24(8):971–983, 2006. ISSN 10870156. doi: 10.1038/nbt1235.
- [101] Felipe Rivas, Osama K. Zahid, Heidi L. Reesink, Bridgette T. Peal, Alan J. Nixon, Paul L. DeAngelis, Aleksander Skardal, Elaheh Rahbar, and Adam R. Hall. Label-free analysis of physiological hyaluronan size distribution with a solid-state nanopore sensor. *Nat. Commun.*, 9(1):1–9, 2018. ISSN 20411723. doi: 10.1038/s41467-018-03439-x.
- [102] Joseph W. F. Robertson and Joseph E. Reiner. The Utility of Nanopore Technology for Protein and Peptide Sensing. *Proteomics*, 1800026:1800026, 2018. ISSN 16159853. doi: 10.1002/pmic.201800026. URL <http://doi.wiley.com/10.1002/pmic.201800026>.

- [103] Joseph W.F. Robertson. *Biological Pores on Lipid Bilayers*. Elsevier Inc., first edit edition, 2013. ISBN 9781437734737. doi: 10.1016/B978-1-4377-3473-7.00004-2. URL <http://dx.doi.org/10.1016/B978-1-4377-3473-7.00004-2>.
- [104] Christian B. Rosen, David Rodriguez-Larrea, and Hagan Bayley. Single-molecule site-specific detection of protein phosphorylation with a nanopore. *Nat. Biotechnol.*, 32(2):179–181, 2014. ISSN 10870156. doi: 10.1038/nbt.2799. URL <http://dx.doi.org/10.1038/nbt.2799>.
- [105] L. Renee Ruhaak, Gege Xu, Qiongyu Li, Elisha Goonatilleke, and Carlito B. Lebrilla. Mass Spectrometry Approaches to Glycomics and Glycoproteomic Analyses. *Chem. Rev.*, page acs.chemrev.7b00732, 2018. ISSN 0009-2665. doi: 10.1021/acs.chemrev.7b00732. URL <http://pubs.acs.org/doi/10.1021/acs.chemrev.7b00732>.
- [106] James F. Rusling, Challa V. Kumar, J. Silvio Gutkind, and Vyomesh Patel. Measurement of biomarker proteins for point-of-care early detection and monitoring of cancer. *Analyst*, 135(10):2496–2511, 2010. ISSN 00032654. doi: 10.1039/c0an00204f.
- [107] K. Sakabe, Z. Wang, and G. W. Hart. -N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl. Acad. Sci.*, 107(46):19915–19920, 2010. ISSN 0027-8424. doi: 10.1073/pnas.1009023107. URL <http://www.pnas.org/cgi/doi/10.1073/pnas.1009023107>.
- [108] J Santoro, C Gonzales, M Bruix, J L Neira, J L Nieto, J Herranz, and M Rico. High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy, 1993. ISSN 00222836.
- [109] Tadashi Satoh, Takumi Yamaguchi, and Koichi Kato. Emerging Structural Insights into Glycoprotein Quality Control Coupled with N-Glycan Processing in the Endoplasmic Reticulum. *Molecules*, 20(2):2475–2491, jan 2015. ISSN 1420-3049. doi: 10.3390/molecules20022475. URL <http://www.mdpi.com/1420-3049/20/2/2475>.
- [110] Franz X. SCHMID. A Native???Like Intermediate on the Ribonuclease A Folding Pathway: 1. Detection by Tyrosine Fluorescence Changes. *Eur. J. Biochem.*, 114(1):105–109, 1981. ISSN 14321033. doi: 10.1111/j.1432-1033.1981.tb06179.x.
- [111] Sangita Seshadri, Keith A. Oberg, and Anthony L. Fink. Thermally Denatured Ribonuclease A Retains Secondary Structure As Shown by FTIR. *Biochemistry*, 33(6):1351–1355, 1994. ISSN 15204995. doi: 10.1021/bi00172a010.
- [112] D. Shental-Bechor and Y. Levy. Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci.*, 105(24):8256–8261, 2008. ISSN 0027-8424. doi: 10.1073/pnas.0801340105. URL <http://www.pnas.org/cgi/doi/10.1073/pnas.0801340105>.
- [113] André M.N. Silva, Rui Vitorino, M. Rosário M. Domingues, Corinne M. Spickett, and Pedro Domingues. Post-translational modifications and mass spectrometry detection. *Free Radic. Biol. Med.*, 65:925–941, 2013. ISSN 08915849. doi: 10.1016/j.freeradbiomed.2013.08.184. URL <http://dx.doi.org/10.1016/j.freeradbiomed.2013.08.184>.
- [114] Zuzanna Siwy, Pavel Apel, Dagmar Baur, Dobri D. Dobrev, Yuri E. Korchev, Reinhard Neumann, Reimar Spohr, Christina Trautmann, and Kay Obbe Voss. Preparation of synthetic nanopores with transport properties analogous to biological channels. *Surf. Sci.*, 532-535:1061–1066, 2003. ISSN 00396028. doi: 10.1016/S0039-6028(03)00448-5.
- [115] Zuzanna S. Siwy. Ion-current rectification in nanopores and nanotubes with broken symmetry. *Adv. Funct. Mater.*, 16(6):735–746, 2006. ISSN 1616301X. doi: 10.1002/adfm.200500471.
- [116] R. M. M. Smeets, U. F. Keyser, N. H. Dekker, and C. Dekker. Noise in solid-state nanopores. *Proc. Natl. Acad. Sci.*, 105(2):417–421, 2008. ISSN 0027-8424. doi: 10.1073/pnas.0705349105. URL <http://www.pnas.org/cgi/doi/10.1073/pnas.0705349105>.

- [117] Langzhou Song, Michael R. Hobaugh, Christopher Shustak, Stephen Cheley, Hagan Bayley, and J. Eric Gouaux. Structure of Staphylococcal alpha -Hemolysin, a Heptameric Transmembrane Pore. *Science* (80-.), 274(5294):1859–1865, dec 1996. ISSN 0036-8075. doi: 10.1126/science.274.5294.1859. URL <http://www.sciencemag.org/cgi/doi/10.1126/science.274.5294.1859>.
- [118] Vincent Tabard-Cossa. Instrumentation for Low-Noise High-Bandwidth Nanopore Recording. *Eng. Nanopores Bioanal. Appl. A Vol. Micro Nano Technol.*, pages 59–93, 2013. doi: 10.1016/B978-1-4377-3473-7.00003-0.
- [119] David S. Talaga and Jiali Li. Single-molecule protein unfolding in solid state nanopores. *J. Am. Chem. Soc.*, 131(26):9287–9297, 2009. ISSN 00027863. doi: 10.1021/ja901088b.
- [120] Koji Tanaka, Jose M.M. Caaveiro, Koldo Morante, Juan Manuel González-Mañas, and Kouhei Tsumoto. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nat. Commun.*, 6(1):6337, dec 2015. ISSN 2041-1723. doi: 10.1038/ncomms7337. URL <http://www.nature.com/doifinder/10.1038/ncomms7337> <http://www.nature.com/articles/ncomms7337>.
- [121] Maureen E. Taylor and Kurt Drickamer. *Introduction to Glycobiology*. Oxford University Press Inc., Oxford, third edit edition, 2011. ISBN 978-0-19-956911-3.
- [122] K. Tejaswi Naidu and N. Prakash Prabhu. Protein-surfactant interaction: Sodium dodecyl sulfate-induced unfolding of ribonuclease A. *J. Phys. Chem. B*, 115(49):14760–14767, 2011. ISSN 15205207. doi: 10.1021/jp2062496.
- [123] Aidan P. Thompson. Nonequilibrium molecular dynamics simulation of electro-osmotic flow in a charged nanopore. *J. Chem. Phys.*, 119(14):7503–7511, 2003. ISSN 00219606. doi: 10.1063/1.1609194.
- [124] David J. Thornton, Karine Rousseau, and Michael A. McGuckin. Structure and Function of the Polymeric Mucins in Airways Mucus. *Annu. Rev. Physiol.*, 70(1):459–486, 2008. ISSN 0066-4278. doi: 10.1146/annurev.physiol.70.113006.100702. URL <http://www.annualreviews.org/doi/10.1146/annurev.physiol.70.113006.100702>.
- [125] Nobuo Ui. Isoelectric points and conformation of proteins. II. Isoelectric focusing of α -chymotrypsin and its inactive derivative. *BBA - Protein Struct.*, 229(3):582–589, 1971. ISSN 00052795. doi: 10.1016/0005-2795(71)90273-X.
- [126] Jeffrey D. Uram, Kevin Ke, and Michael L. Mayer. Noise and bandwidth of current recordings from submicrometer pores and nanopores. *ACS Nano*, 2(5):857–872, 2008. ISSN 19360851. doi: 10.1021/nn700322m.
- [127] Philippe Van Den Steen, Pauline M. Rudd, Raymond A. Dwek, and Ghislain Opdenakker. Concepts and principles of O-linked glycosylation. *Crit. Rev. Biochem. Mol. Biol.*, 33(3):151–208, 1998. ISSN 10409238. doi: 10.1080/10409239891204198.
- [128] Iris van Uitert, Séverine Le Gac, and Albert van den Berg. The influence of different membrane components on the electrical stability of bilayer lipid membranes. *Biochim. Biophys. Acta - Biomembr.*, 1798(1):21–31, 2010. ISSN 00052736. doi: 10.1016/j.bbamem.2009.10.003. URL <http://dx.doi.org/10.1016/j.bbamem.2009.10.003>.
- [129] Bala Murali Venkatesan and Rashid Bashir. Nanopore sensors for nucleic acid analysis. *Nat. Nanotechnol.*, 6(10):615–624, 2011. ISSN 17483395. doi: 10.1038/nnano.2011.129. URL <http://dx.doi.org/10.1038/nnano.2011.129>.
- [130] Christopher T. Walsh, Sylvie Garneau-Tsodikova, and Gregory J. Gatto. Protein posttranslational modifications: The chemistry of proteome diversifications. *Angew. Chemie - Int. Ed.*, 44(45):7342–7372, 2005. ISSN 14337851. doi: 10.1002/anie.200501023.
- [131] Yanan Wang, Jiaxi Wang, Mingxia Gao, and Xiangmin Zhang. A novel carbon material with nanopores prepared using a metal-organic framework as precursor for highly selective enrichment of N-linked glycans. *Anal. Bioanal. Chem.*, 409(2):431–438, 2017. ISSN 16182650. doi: 10.1007/s00216-016-9796-1. URL <http://dx.doi.org/10.1007/s00216-016-9796-1>.

- [132] Meni Wanunu. Nanopores: A journey towards DNA sequencing. *Phys. Life Rev.*, 9(2):125–158, 2012. ISSN 15710645. doi: 10.1016/j.plrev.2012.05.010. URL <http://dx.doi.org/10.1016/j.plrev.2012.05.010>.
- [133] Henry S. White and Andreas Bund. Ion current rectification at nanopores in glass membranes. *Langmuir*, 24(5):2212–2218, 2008. ISSN 07437463. doi: 10.1021/la702955k.
- [134] R. L. Williams, S. M. Greene, and A. McPherson. The crystal structure of ribonuclease B at 2.5-A resolution. *J. Biol. Chem.*, 262(33):16020–16031, 1987. ISSN 00219258.
- [135] Carsten Wloka, Natalie Lisa Mutter, Misha Soskine, and Giovanni Maglia. Alpha-Helical Fragaceatoxin C Nanopore Engineered for Double-Stranded and Single-Stranded Nucleic Acid Analysis. *Angew. Chemie - Int. Ed.*, 55(40):12494–12498, sep 2016. ISSN 15213773. doi: 10.1002/anie.201606742. URL <http://doi.wiley.com/10.1002/anie.201606742>.
- [136] C. T.A. Wong and M. Muthukumar. Polymer capture by electro-osmotic flow of oppositely charged nanopores. *J. Chem. Phys.*, 126(16), 2007. ISSN 00219606. doi: 10.1063/1.2723088.
- [137] Worthington Biochemical Corporation. Ribonuclease A, 2018. URL <http://www.worthington-biochem.com/rnase/>.
- [138] Ronen Zangi, Ruhong Zhou, and B. J. Berne. Urea's action on hydrophobic interactions. *J. Am. Chem. Soc.*, 131(4):1535–1541, 2009. ISSN 00027863. doi: 10.1021/ja807887g.
- [139] Yanan Zhao, Brian Ashcroft, Peiming Zhang, Hao Liu, Suman Sen, Weisi Song, Jongone Im, Brett Gyarfas, Saikat Manna, Sovan Biswas, Chad Borges, and Stuart Lindsay. Single-molecule spectroscopy of amino acids and peptides by recognition tunnelling. *Nat. Nanotechnol.*, 9(6):466–473, 2014. ISSN 17483395. doi: 10.1038/nnano.2014.54. URL <http://dx.doi.org/10.1038/nnano.2014.54>.
- [140] Yingming Zhao and Ole N. Jensen. Modification-specific proteomics: Strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics*, 9(20):4632–4641, 2009. ISSN 16159853. doi: 10.1002/pmic.200900398.